

CAPEX Bias and Adverse Incentives in Incentive Regulation Issues and Solutions



CAPEX Bias and Adverse Incentives in Incentive Regulation

Issues and Solutions

Andrew Smith Phill Wheat Jean-Christophe Thiebaud Alex Stead University of Leeds, United Kingdom

Working Group Paper 2019

The International Transport Forum

The International Transport Forum is an intergovernmental organisation with 60 member countries. It acts as a think tank for transport policy and organises the Annual Summit of transport ministers. ITF is the only global body that covers all transport modes. The ITF is politically autonomous and administratively integrated with the OECD.

The ITF works for transport policies that improve peoples' lives. Our mission is to foster a deeper understanding of the role of transport in economic growth, environmental sustainability and social inclusion and to raise the public profile of transport policy.

The ITF organises global dialogue for better transport. We act as a platform for discussion and prenegotiation of policy issues across all transport modes. We analyse trends, share knowledge and promote exchange among transport decision-makers and civil society. The ITF's Annual Summit is the world's largest gathering of transport ministers and the leading global platform for dialogue on transport policy.

The Members of the Forum are: Albania, Armenia, Argentina, Australia, Austria, Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Canada, Chile, China (People's Republic of), Croatia, Czech Republic, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Hungary, Iceland, India, Ireland, Israel, Italy, Japan, Kazakhstan, Korea, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Mexico, Republic of Moldova, Montenegro, Morocco, the Netherlands, New Zealand, North Macedonia, Norway, Poland, Portugal, Republic of North Macedonia, Romania, Russian Federation, Serbia, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Tunisia, Turkey, Ukraine, the United Arab Emirates, the United Kingdom and the United States.

International Transport Forum 2 rue André Pascal F-75775 Paris Cedex 16 contact@itf-oecd.org www.itf-oecd.org

ITF Working Group Papers

ITF Working Group Papers present findings on specific aspects within the area of inquiry of an ITF working group. They accompany and complement the group's high-level findings presented in its main output, published as an ITF Research Report, with additional detail and analysis. ITF working groups bring together international experts over a period of usually one to two years, and are vetted by the ITF's Transport Research Committee. Any findings, interpretations and conclusions expressed herein are those of the authors and do not necessarily reflect the views of the International Transport Forum or the OECD. Neither the OECD, ITF nor the authors guarantee the accuracy of any data or other information contained in this publication and accept no responsibility whatsoever for any consequence of their use. This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Cite this work as: Smith, A. et al. (2019), "CAPEX Bias and Adverse Incentives in Incentive Regulation: Issues and Solutions", Working Group Paper, International Transport Forum, Paris.

Foreword

Transport infrastructure is a major enabler of economic development. In the drive to refurbish or build, governments worldwide have turned to the private capital market for financing. The primary narrative behind this push is the huge stocks of private capital that are available, while public financing capabilities are said to be limited and insufficient.

The almost exclusive vehicle of private investment in transport infrastructure, including social infrastructure, is Public-Private Partnerships (PPPs). In the context of PPPs, two important aspects have received little attention.

First, sufficient attention has not been given to the role of suppliers. The focus of governments and Intergovernmental Organisations has been on resolving the challenges to private investment from the viewpoint of investors: reducing the uncertainty they face and enabling them to price risk more efficiently by establishing infrastructure as an asset class.

However, looking only at investors gives an incomplete view of the total cost of the risk transferred from the public to the private sphere. In PPPs, investors transfer some of the major risks they are not comfortable bearing to design, construction, maintenance, and operations contractors.

Suppliers, too, face uncertainties and are unable to efficiently evaluate price risk. In such cases, the base cost of the initial investment – and of subsequent services – may be much higher than they might have been, and not just the cost of their financing.

Uncertainty arises from the difficulties to accurately estimate the cost of construction, maintenance, operations, and financing. But it also stems from "unknown unknowns" (the so-called Knightian uncertainty). For instance, changes in weather patterns or paradigmatic technological shifts, the timing and impact of which are unclear, will influence what infrastructure is needed and where.

So what can policy makers do to reduce the cost of inefficient risk pricing of suppliers? Where does this put PPPs? How can public decision makers reconcile long-term uncertainty with private investment in infrastructure? Who should bear long-term uncertainty in projects: the public or the private sector?

These were some of the guiding questions for a Working Group of 33 international experts convened by the International Transport Forum (ITF) In September 2016. The group, which assembled renowned practitioners and academics from areas including private infrastructure finance, incentive regulation, civil engineering, project management and transport policy, examined how to address the problem of uncertainty in contracts with a view to mobilise more private investment in transport infrastructure. As uncertainty matters for all contracts, not only those in the context of private investment in transport infrastructure, the Working Group's findings are relevant for public procurement in general.

The synthesis report of the Working Group was published in June 2018. The report is complemented by a series of 19 topical papers that provide a more in-depth analysis of the issues. A full list of the Working Group's research questions and outputs is available in Appendix 1.

Acknowledgements

The authors would like to acknowledge the comments made by Julien Brunel at SNCF Réseau and more widely the funding provided by SNCF Réseau to enable the study to take place. The authors are also grateful to Dejan Makovšek at the ITF for motivating the initial topic, helping to secure the funding, and for his comments on earlier drafts. Finally, the authors would like to acknowledge the comments made by all the members of the Working Group on the presentations given in relation to this work.

Table of contents

Acronyms and abbreviations	6
Executive summary	7
Introduction	9
Theoretical and regulatory literature10	0
OPEX and CAPEX substitutability 12 Regulatory systems and the A-J effect 13 Capital bias within the RPI-X regulatory system 14 Mitigation of capital bias 15 Summary and concluding remarks 22	2 3 5 9 2
The system of regulation in case studies: Rail, roads and water	2
Rail infrastructure	2 4 6
Evidence for capital bias	0
Introduction 30 Rail infrastructure in Great Britain 31 The water industry in England and Wales 31 Road sector in England 41	0 1 7 5
The effectiveness of efficiency benchmarking in transport 44	8
The purpose of top-down econometric benchmarking.49Key challenges associated with CAPEX benchmarking52Consistency of data definitions52Capital projects are lumpy and intertemporal53Intermediate versus final outputs54The treatment of enhancements55Combining approaches and the role of the regulator63Summary64	9 2 3 5 8 9 2 4
Conclusions	6
Notes	9
References	1
Appendix 1. Research questions and outputs of the Working Group on Private Investment in Infrastructure	5

Acronyms and abbreviations

A-J effect	Averch-Johnson effect
BOTEX	TOTEX less enhancements expenditure
CAPEX	Company capital expenditure
CMA	Competition and Markets Authority
СР	(Regulatory) Control Period (e.g. CP4)
CQC	Cost-Quality-Customer
CIS	Capital incentive scheme
DEA	Data Envelopment Analysis
DfT	Department for Transport (UK)
HLOS	High Level Output Specification
HMEP	Highways Maintenance Efficiency Programme
LHAs	Local Highway Authorities
LHN	Local Highway Network
LRIC	Long-run incremental cost
MFPVS	Marginal future saving in present value terms
M&R	Maintenance and renewal
MUFC	Marginal up-front cost
NPIF	National Productivity Investment Fund
ODIs	Outcome delivery incentives
OFGEM	Office of Gas and Electricity markets (energy regulator, UK)
OFWAT	Office for Water Services (economic regulator of the water sector in England
	and Wales)
OPEX	Company operating expenditure
ORR	Office of Rail and Road (formerly Office of Rail Regulation)
PAYG	Pay-as-you-go
PCs	Performance commitments
PPP	Public-Private Partnership
PR	Periodic Review (e.g. PR14 is the Periodic Review in 2014)
RAB	Regulatory asset base
RCM	Revenue correction mechanism
RCV	Regulated Capital Value
RISs	Road Investment Strategies
RMMS	Rail Market Monitoring Survey
RPI-X	Retail prices index (RPI) – X (productivity factor)
S&Cs	Switches and Crossings
SOFA	Statement of Funds Available
SRN	Strategic Roads Network
TOEX	Measure of company total cash costs (operating expenditure, including maintenance,
	plus capital expenditure (renewals and enhancements)
UIC	International Union of Railways
WACC	Weighted average cost of capital
WaSCs	Water and sewage companies in England and Wales
WOCs	Water only companies in England and Wales

Executive summary

What we did

This paper addresses the question of to what extent benchmarking XXX can deliver YYY. Three case studies from the United Kingdom provide the material. They look at rail infrastructure, local authority roads and the water sector.

Two concerns are at the centre of the analysis. The first is that there may be a capital bias in rate of return regulation systems (the Averch-Johnson 1962 effect) where the allowed rate of return is greater than the cost of capital. Other sources of capital biases may occur within systems of economic regulation based on retail price inflation (RPI) minus X inefficiency (RPI-X). These relate to the differential treatment of company operating expenditure (OPEX) and company capital expenditure (CAPEX) in the benchmarking and wider regulatory approaches. Such biases may cause costs to be above efficient levels in regulated systems, thus raising the question as to whether benchmarking can be effective at overcoming them.

The second concern is whether benchmarking can work more generally for transport investment. Particular focus here is on the specific problems that may emerge in benchmarking CAPEX expenditure (renewals and enhancements).

What we found

The conclusions of this paper are that, historically, CAPEX bias has been overplayed in the literature. However, there are reasons why CAPEX biases may exist and UK economic regulators have been alert to these. Demonstrating such bias with available data is hard, however. The data at hand does not indicate CAPEX bias across the three sectors considered.

Some regulators have nevertheless sought to benchmark aspects of OPEX and CAPEX together and to adjust other aspects of the regulatory regime to counter the possibility of CAPEX bias.

The possibility that OPEX bias in nationalised industries faces short-term funding constraints has also been highlighted. This could be an issue in transport infrastructure investment in the absence of an independent regulator with a focus on funding levels and efficiency and/or a multi-annual agreement that respects long-term planning and whole-life costing.

The purpose of top-down econometric benchmarking is to produce a set of high-level efficiency scores for use in setting efficiency targets. Its particular value is that the nature of technology (e.g. economies of scale and density) is controlled for prior to arriving at an efficiency score. Further, there is transparency over the results, which gives confidence in the resulting efficiency scores.

The results of an econometric model represent one step in the process of setting efficiency targets and delivering efficiency savings. Regulators may make adjustments and apply their judgement. Bottom-up benchmarking – where specific initiatives for improving efficiency are identified – can be powerful in corroborating top-down methods and help ensure that realistic targets are set. In some contexts bottom-up methods can support learning across regions, once top-down benchmarking has established where to look for good and potentially poor practice.

Regulators have to guard against becoming too involved in the micro-management of regulated firms, however. This raises the question of who is running the enterprise and what freedom to optimise is being given to managers. Where several private firms are being benchmarked, this is particularly important. With a single infrastructure manager that is owned by the state (as is the case for all rail infrastructure providers in Europe), however, it is less clear that the regulator can take a hands-off approach. Nevertheless, management and reputational incentives can still motivate managers of state-owned firms to perform well. If regulators become too closely involved in the detail, there is danger of regulatory capture.

There are particular challenges to CAPEX benchmarking, given CAPEX's lumpy and temporal nature. Approaches to deal with such issues are not perfect but can be used to support a strong benchmarking framework. Examples of their implementation are provided.

Enhancements are more challenging, particularly where the level of spending on enhancements and the choice of schemes are set by government. For enhancements in transport, and to an extent, renewals, a focus on project level data to assess the unit cost of delivering projects is a useful way forward in benchmarking. Maintenance and renewals expenditure may then be benchmarked separately from enhancements. The term BOTEX benchmarking has been used for this situation.

Overall, regulatory benchmarking within an RPI-X model, including CAPEX, has been proven to deliver significant benefits across several sectors. While not without problems, it can be a useful alternative to PPPs. Transport applications potentially face greater challenges than other sectors. In part this reflects the state ownership in these sectors, however. In networks, such as road or rail, there may be only one firm responsible for the entire network. Yet the approach can work even in these more challenging contexts, provided regional benchmarking of such operators can be implemented and a strong independent regulator exists.

Finally, care always needs to be taken with regard to the comparability of data. A significant investment in terms of time is needed to ensure comparable definitions. Here international benchmarking is particularly challenging.

What we recommend

Make top-down econometric benchmarking part of the regulators' toolkit to set efficiency targets for transport infrastructure

Economic benchmarking can be applied to CAPEX, although it needs to be carefully tailored. Regulatory judgement and bottom-up approaches are important complementary approaches. Where only one infrastructure manager exists in a country, regional internal and international benchmarking become particularly important tools. The latter is particularly challenging. Concerted efforts by regulators and governments across Europe may be required for making relevant data available.

Consider regulatory benchmarking as an alternative to Public-Private Partnerships

The regulated model can be an effective alternative to PPPs in the provision of transport infrastructure if supported by appropriate benchmarking approaches. It is not without problems but, with an independent and capable regulator, it can provide an alternative for delivering transport infrastructure.

Introduction

A key issue in economic regulation, at least in theory, is the problem of capital bias. This may be split into two separate but related problems:

- 1. The traditional Averch-Johnson effect (Averch and Johnson, 1962) where the allowed rate of return is greater than the regulated firm's cost of capital may create an incentive to over-capitalise, leading to an inefficient capital/labour ratio and costs being too high overall.
- 2. A more recent source of capital bias within economic regulatory systems (especially UK) where different regulatory benchmarking and incentive regimes are applied to operating costs as compared to capital expenditure may have led to re-balancing of expenditure towards capital expenditure.

Given this issue it is pertinent to ask whether the benchmarking framework, within the context of an economic regulatory framework, can ensure cost efficient delivery of transport infrastructure operations, maintenance, renewals and enhancements. Ultimately, if benchmarking can work then the regulated model can be seen as a good alternative to using PPPs for delivering and financing transport infrastructure.

For a framework of cost efficiency benchmarking to work it therefore needs to be able to overcome the two possible sources of capital biases noted above, namely the traditional Averch-Johnson effect and the further bias that may result from the adoption of separate approaches for operating and capital expenditure (as has been common within the UK regulatory approach). So called "TOTEX" solutions, whereby operating and capital expenditure are benchmarked together have been proposed and tried to different degrees in the regulatory approaches applied to rail, water and roads in the UK. A further relevant distinction concerns the treatment of different types of capital expenditure that is designed to enhance the capability or extent of the network in some way (enhancement capital expenditure). It is important to note, however, that to date the evidence base for the existence of capital bias is relatively weak.

Further, in order for a transport benchmarking framework to be successful there are numerous other issues that need to be addressed in addition to the capital bias issue:

- how to address the heterogeneity between networks in terms of, for example, performance, speed, age of the network, traffic mix
- how to address the heterogeneity in cost reporting and capital/operating cost distinctions
- how to deal with the cyclical characteristics of renewals costs which could make one transport infrastructure provider appear relatively inefficient (or efficient) in a particular year
- how best to treat this category of expenditure within the regulatory framework, given the bespoke and potentially lumpy nature of enhancement expenditure
- how the econometric approach to benchmarking, which is conducted at a high level, can be combined with a bottom-up approach aimed at understanding the reasons for gaps in performance

• how econometric results, given by a top-down approach, can be communicated to stakeholders and developed into operational recommendations, with actions to improve productivity.

The purpose of this position paper is to address the following high-level question: to what extent can benchmarking work in transport? We aim to answer this question through three UK case studies: rail infrastructure; local authority roads; and the water sector. In addition to the rail and roads cases, the water case was chosen as a good example of a comparator network industry. These cases offer highly relevant case studies through which to study the issues under consideration in this paper and the project team have been closely involved in working with regulators/policy makers/providers to develop the regulatory approaches developed in each case.

It should be noted that in answering this question we focus in particular on the specific issues relating to the comparison of CAPEX expenditure between organisations. There are accepted means to benchmark operations expenditures and these have been used by numerous regulators. CAPEX benchmarking is however inherently more problematic as will be explained further in this report. There have been instances of benchmarking of such expenditures in the regulatory model (discussed later), but there are some unique challenges associated with it.

The paper is structured as follows. In the next section, the relevant theoretical and regulatory literature is briefly summarised, highlighting the link between the system of economic regulation and the two sources of capital bias. The possibility of an "OPEX" bias within state-owned systems is also discussed. The third section provides a brief overview of the system of regulation in each of three sectors (rail, local authority roads and water), emphasising the treatment of the different categories of expenditure (operations, maintenance, renewals and enhancements). The fourth section considers whether there is any evidence of a capital bias in the three sectors based on a combination of data and desk-based research. The final section addresses the question of whether efficiency benchmarking can be a powerful tool in driving efficiency improvements in transport, viewed through the lens of the three cases, rail, roads (and water). The issues and challenges are set out, together with examples of how they have been resolved to date, and the strengths and weaknesses of these approaches. Emphasis is also placed on the role of the regulator, and in particular how regulators use top-down benchmarking, alongside other studies (including bottom-up approaches), to reach an overall assessment of the potential for efficiency savings in a regulated sector. As noted, the focus is on the particular challenges of CAPEX benchmarking.

Theoretical and regulatory literature

The last thirty years has seen a transformation in the ownership, operation and performance of transport and network services internationally. Across transport and the utilities sectors, much of what was formally provided by the State (at least outside the US) has been transferred to the private sector, often in combination with reforms to introduce competition where possible; where not, regulatory institutions have been set up to protect the public interest, and also ensure that firms are protected from government intervention. Reforms have progressed at different speeds and to different extents in different parts of the world, with the UK being widely acknowledged as a leader of this type of reform and the associated regulatory mechanism that have developed since. A key driver behind the UK's desire to restructure and privatise network industries was the poor performance of nationalised companies (Pollitt, 1999). One feature of state-owned industries in the UK and more widely internationally has been the tendency for conflicting objectives, too much government intervention and weak monitoring to lead to inefficient outcomes (Rees, 1984). This problem is inextricably linked with the typically short-term political cycles affecting government policy in respect of state owned companies that persist around the world. Shortages of funding tend to lead to under-investment, poor quality and over reliance on short-term operating (OPEX) expenditure (patch and mend); which could be seen to be a kind of "OPEX" bias within the state owned system.

One objective of privatisation was therefore to improve efficiency whilst also harnessing private capital to provide much needed investment. Of course this then raises a pre-commitment problem – with investors needing some form of regulatory reassurance before committing to investment. Independent regulators have been set up to offer this reassurance, and the rate of return regulatory approach is one means (widely adopted in the US) to dealing with the pre-commitment problem. However, this method of regulation suffers from providing weak incentives generally to improve efficiency, and also from the specific over-capitalisation bias referred to in the literature as the Averch-Johnson (A-J) effect. The RPI-X method introduced in the UK (Beesley and Littlechild, 1992) was intended to overcome these problems, providing high powered incentives to improve efficiency by exposing companies to greater risk and challenge in respect of cost control within five-year control periods.

The RPI-X approach has been widely credited with delivering very substantive efficiency improvements amongst UK regulated industries, with continued pressure on costs being achieved through the application of benchmarking techniques, combined with incentives to outperform on efficiency targets set by the regulator. This approach has continued to deliver improvements in efficiency many years after privatisation (see Table 1).

Control period	Savings per year (%)
First	2.2
Second	6.8
Third	6.3
Fourth	3.4
Fifth	2.6

Table 1. Real Cost Savings Achieved in UK Regulated Industries

Source: Oxera (2008).

As will be discussed in more detail below, despite the successes, RPI-X regulation has not been without its problems. In particular, significant challenges have emerged as regulators have tried to maintain pressure on capital investment costs, whilst also incentivising new investment. Capital expenditure (CAPEX), given its cyclical and potentially lumpy nature, makes it harder to benchmark against other companies than OPEX. This problem is exacerbated when the particular capital programme is highly bespoke in nature (capital investments aimed at enhancing the scale and capability of the network are particularly problematic in this regard as there may be few, if any, comparable projects). There may even be questions as to whether the investment is needed at all.

The problems associated with capital benchmarking have meant that it has been challenging for regulators to set credible efficiency targets in respect of capital expenditure. As a result, RPI-X regulation

may have created a new form of capital bias in which firms have an incentive to prefer capital intensive solutions, because of the different treatment of OPEX and CAPEX in the regulatory system. Since the original RPI-X mechanism introduced by the Littlechild report (1983), regulators have been gradually evolving the regulatory mechanism to try to overcome these problems.

The remainder of this section is structured as follows. We first highlight the fundamental substitutability between OPEX and CAPEX which is central to the debate. The A-J effect and wider problems inherent within the rate-of-return regulatory system are then set out, followed by a brief discussion of the pre-commitment problem. The "new" problem of CAPEX bias that results from the differing ways in which OPEX and CAPEX are treated within the RPI-X regulatory framework is then introduced and discussed. Finally, mitigation strategies for addressing this latter type of bias are noted. Further details of the regulatory approaches and how they have been implemented in rail, roads and water are provided in the sections "The system of regulation..." and "Evidence for capital bias".

OPEX and CAPEX substitutability

Network industries are often characterised as capital intensive industries due to the fact that they rely on an infrastructure to deliver their outputs. However this initial assessment neglects the fact that network operators face trade-offs between the amount of future day-to-day operations and the initial value of the investment. For instance, in the water industry, regulator Ofwat highlights that in order to match supply and demand, companies may choose between building a reservoir (i.e. a new asset) or relying on bulk water trading. Regarding waste water, companies also have access to various technologies to deal with excess water. They can choose to increase the size of sewers which will require limited maintenance or use alternative approaches such as sustainable drainage systems.

Trade-offs are crucial to both the initial investment and the care and operation of the infrastructure. Network operators rely on different type of works which are usually classified in three categories:

- Operations costs, which consists of the day-to-day costs of operating the network (for example, signalling operations in rail)
- Maintenance, which sustains the condition and capability of existing infrastructure, but does not involve significant replacement of assets
- Renewals, consisting mainly of projects where existing infrastructure is replaced with new assets, normally without enhancement of performance
- Enhancements are projects resulting in a change to network outputs, usually involving construction, which improves network capacity or capability.

Maintenance is considered an operating expenditure (OPEX), while renewals and enhancement are capital expenditure (CAPEX). Yet it is worth noting that the objective of both maintenance and renewal expenditures is to maintain a network in its original state and value, while enhancements should increase the value of the network.

Importantly, the different types of works are, in nature, substitutes to some degree. For instance, Gaudry et al. (2016), point out that in railways continuous maintenance will delay the periodic regeneration of the asset while similar observations have been made for road networks by Newberry (1988). In practice, finding an optimal mix is on-going.¹ For instance, as pointed out in the Value for Money Study of Great Britain's Railways (McNulty, 2011), the trade-off was not fully understood by rail infrastructure managers and one the recommendations was to have a: "better selection of the optimum maintenance

approaches, informed by better understanding of assets and better asset condition information to reduce maintenance and renewals effort." Despite this substitutability, it is often considered that the absence of sufficient renewal will increase the cost of maintenance exponentially and thus renewal cannot be delayed indefinitely.

Regulatory systems and the A-J effect

Given the substitutability between operating and capital expenditures, it is important that regulators supervising network industries set adequate incentives in order to foster allocative efficiency (in terms of input selection) as well as technical efficiency. Misallocation of inputs has been a concern since the seminal paper by Averch and Johnson (1962) who pointed out that rate of return regulation may lead to an inefficient use of inputs, as the profit-maximising firm under regulatory constraint will tend to use a capital-labour ratio different from that which minimises cost for its output level. This result is driven by the assumption that the allowed rate of return is greater than the true cost of capital,² in which case the capital-labour ratio of the regulated firm will be larger than the one that minimises costs, creating a capital bias.

The A-J effect is illustrated in Figure 1 below. The fact that the allowed rate of return is higher than the true cost of capital is reflected in an effective reduction in the price (cost) of capital (the dotted line in Figure 1). In Figure 1, *r* is the price of capital, and *w* the price of labour. Thus firms would optimise (from their perspective) based on the new "price ratio", thus following an inefficient expansion path (expansion path₂ in place of the efficient expansion path₁). At all levels of output along the inefficient expansion path too much capital is used than is (socially) optimal, and costs will be higher overall as a result. The evidence (or not) for the existence of the A-J effect is discussed in section "Evidence for capital bias" below.





Of course, the application of RPI-X regulation should mitigate against the A-J effect, at least to some extent. During the control periods in the UK regulatory model, firms are exposed to cost risk for an exogenously set period (five years) and therefore should have stronger incentives to improve technical efficiency than the rate of return model; and this incentive should also apply to allocative efficiency. Any increased cost due to the existence of an inefficient capital structure would reduce firm profitability

during the control period. It should be said, as discussed further below, that since rate of return still plays a role in the RPI-X regulatory model, through the way in which investment in capital assets are remunerated, the A-J effect could still occur even in the RPI-X model. It would however be expected to be less prevalent and will depend on precisely how firms are remunerated for capital expenditure.

It should be noted that whilst the provision of an allowed rate of return within a RPI-X regulatory system could create a capital bias, such an approach is of course designed to avoid the opposite "pre-commitment" problem, where incentive regulation may create uncertainty and restrict investment. Indeed in network industries, investments are often irreversible, that is the physical capital once set in place cannot be redeployed or taken back (Evans and Guthrie, 2005). And as pointed out by Levy and Spiller (1994), an infrastructure manager will keep on operating a network as long as it covers its operating expenditure even if it cannot cover its capital expenditure. Regulated firms become vulnerable to future decisions made by a regulator subject to public or political pressure as they do not know how regulation will evolve in the following time period and the life cycle of regulated assets will usually exceed the five year laps of a regulation period (Joskow, 2008).

A solution to foster long-term investment was therefore to introduce regulators that could credibly commit to granting a return on capital expenditure made by firms. In particular, this meant promoting independent regulatory agencies, which are seen since Levy and Spiller (1994) as an effective mean to achieve credible commitment. Therefore, as pointed out by Stern (1997), the regulator's role is to reassure an infrastructure manager in a context of sunk investment. Independent regulation within regulatory systems such as that in Great Britain offers this degree of security for private investors; but problems of pre-commitment in developing countries can be more problematic, where such institutional arrangements may not be in place. As will be discussed in sections "The system of regulation..." and "Evidence for capital bias", the issue of regulation of a company in receipt of government subsidy (e.g. Network Rail) may lead to greater problems than in other contexts, since the government has an even greater interest in regulatory decisions as it will impact government spending. A process has been put in place to try to address this issue in Britain (see next section).

The concept of independent regulation of private firms can also be seen as an antidote to the problem of too much government interference in the model of state-ownership, which can be seen as a key source of inefficiency in nationalised industries. Yet it is often the case that the financing of network industries partly relies on public subsidies, in which case an independent agency does not have the role or means to commit the government and parliament on future subsidies. For instance, in the case of rail the McNulty report does mention a "plan B", that is a smaller network in case the railway industry fails to deliver certain improvements. In this case, part of the assets of the infrastructure manager would cease to be of use.

The introduction of incentive regulation could also mean a decrease in quality because an easy way for a firm to increase its profit over the period of regulation might be to decrease its investments in quality enhancing projects. Yet in practice, Ai, Martinez and Sappington (2004) find that in the telecom industry in the US, incentive regulation³ led to an increase in quality when it had a demand enhancing effect. Alternatively, the regulator might want to set targets in terms of quality in order to avoid any decrease. In which case, we assume that quality can be controlled by the firm and observed by the regulator. In the case of electricity distribution, a description of such mechanisms is given by Growitsch et al. (2010). In general the regulatory process in the UK includes incentives (and penalties) for good (bad) performance and there is strong evidence that these have led to quality improvements in many cases.

Tensions remain with respect to the fact that regulators desire to give companies the certainty that they need to invest and innovate, whilst also encouraging technical and allocative efficiency. It is possible that

regulators could push too far in either direction. Arguably rate of return regulation should ensure that firms have the incentive to invest, but there are limited pressures to improve technical efficiency, and a possible capital bias through the A-J effect. RPI-X regulation may create strong incentives to improve both technical and allocative efficiency, but could encourage under-investment, leading to quality problems (in the short and longer terms). The different ways in which UK economic regulators have sought to deal with these issues are further discussed below.

Capital bias within the RPI-X regulatory system

Since RPI-X regulation was introduced in the UK, differences have emerged in the treatment of OPEX and CAPEX in the regulatory framework. This dual approach can be explained by the desire to mitigate under-investment risks under an incentive-based regime and also by difficulties associated with efficiency targets for CAPEX. This duality is deemed to have caused a further source of capital bias. As noted in Ofwat (2013), there are three inter-related policy decisions taken by regulators that could contribute to this problem:

- Cost recovery this reflects the proportion of current operating and investment expenditure that should be paid for by existing versus future customers
- Cost benchmarking framework, and whether there are differences in the treatment of OPEX and CAPEX
- Cost performance incentives how firms are incentivised to improve their efficiency over time and how risk might be shared with customers, for example via the design of a regulatory menu.

Before turning to discuss these, we first briefly describe the broad regulatory framework that applies in the United Kingdom. Independent regulators have, in general, set in place a framework that distinguishes between operating and capital expenditure. The most common approach is the so-called building block approach: which determines the allowed revenues of a regulated firm as the sum of three main blocks which are operating expenditure, depreciation and the return on capital, such that:

Allowed revenues = OPEX + depreciation + RAB*WACC

where RAB is the regulatory asset base and WACC⁴ is the weighted average cost of capital. At the start of each control period the regulator must review and set the appropriate value for these components, which in turn will have an influence over the decisions of regulated firms and raises the question of a possible CAPEX bias. This methodology can be summed up by the following taken from Great Britain's railway regulator: The difference between the treatment of OPEX and CAPEX is immediately clear in this regulatory framework regulated; firms are remunerated on a "pay-as-you-go"⁵ basis for OPEX, but CAPEX is capitalised and added to the RAB⁶, so that the firm is funded gradually over time through a depreciation charge, together with an allowed rate of return on the RAB. The gap between revenues and expenditure (if not dealt with via direct subsidy) becomes a source of debt for the firm. Although the allowed rate of return should in theory cover the cost of debt, it can create a risk for the long-term sustainability of the regulated firm.

Cost recovery

In terms of cost recovery, such an approach means that the cost of new investment is shared across current and future consumers over time, which may be seen as more equitable. However, it is possible that the approach outlined in Figure 2 could impact firms' choices regarding the mix between OPEX and

CAPEX. The approach potentially reduces incentives to improve efficiency on CAPEX as compared to OPEX, because cost savings in the latter translate straight through into improved profitability, whereas the effect for capital is more muted because it comes through in the form of a lower depreciation charge on new CAPEX (which will only be a small proportion of the actual capital spend, given the long asset lives in network industries). Further, to the extent that the allowed rate of return is higher than the cost of capital, there could be A-J related capital biases entering the system – though these again should be weakened to some degree by the implied losses resulting from inefficiency, at least during the five-year control period.





The pay-as-you-go approach has been applied to some aspects of CAPEX in some industries; indeed prior to 2001, renewals were treated in that way in the regulation of the rail infrastructure manager (Railtrack) by ORR. Indeed, ORR noted that a pay-as-you-go approach should also provide "a strong incentive for Railtrack to improve the efficiency of its renewals programme and to understand the condition of its assets and the relationship between renewals, maintenance condition and performance."⁷ ORR later switched from pay-as-you-go mechanisms to capitalisation of expenditure because of potential "unanticipated changes in renewals requirements (e.g. because of the lack of information about the current state of the assets⁸)." This point is related to the approaches to benchmarking and also efficiency incentives, to which we now turn.

Cost benchmarking framework

The framework to compute the allowed revenue of a regulated firm is not the only factor influencing allocative efficiency. The way in which the efficiency assessment (cost benchmarking) is carried out for each of the elements of expenditure also has an influence on a firm's choices. On the one hand, in the past, UK economic regulators have typically applied top-down econometric benchmarking techniques to OPEX cost categories (maintenance), with OPEX set to an efficient level by comparing with similar firms as suggested by the Schleifer (1985); referred to as yardstick competition.

Source: ORR Periodic Review Final (2013).

On the other hand, renewals and enhancements were usually dealt with differently, with econometric techniques generally playing a less prevalent role. Expenditure allowances for CAPEX are usually determined via a mix of bottom-up challenge of business plans, perhaps supplemented by some form of relatively simple quantitative cost analysis. Depending on the context, econometric work on CAPEX often focuses on comparing the cost of units across projects, rather than judging the efficient level of activity. There are potentially good reasons for taking this approach. The level of renewals that are needed in a given year will depend on the stage of replacement cycle that the company is operating in, which could differ from that of comparators, thus distorting efficiency comparisons in a given year. This problem could be mitigated though if a long time series of data is available, but often this is not the case. The problem could be exacerbated by past under-investment (including during the period before privatisation). Enhancement investments are even more problematic – depending on the context, they could be highly bespoke and hard to compare against other benchmarks.

It might then be said that if a firm is inefficient with respect to its OPEX, it will be found out, and targeted with tough efficiency targets, based on the yardstick competition approach underpinned by an econometric model. If the firm is inefficient in respect of its renewals and enhancement spend it may be much more difficult for regulators to determine a firm's relative performance. Hence companies may have incentives to implement capital-intensive solutions to move expenditure into a less challenging benchmarking regime. In this way firms will achieve a good ranking in the OPEX comparisons, whereas any inefficiency in CAPEX will be much harder to detect.

The bias created by regulation may not only lead to an inefficient choice of inputs, but might also translate in accounting manipulation to increase allowed revenues. Indeed firms might try to game the system because of a blurred difference between OPEX and CAPEX in some cases. The Australian Energy Regulator (Australian Competition and Consumer Commission, 2012) pointed out that as some OPEX can be capitalised, such as wage and installation costs for new assets, regulated utilities may have an incentive to capitalise as much OPEX as possible in order to inflate the regulatory asset base (if the allowance for capital cost is subject to rate-of-return regulation, rather than external benchmarking). Therefore, reviewing the regulatory asset base is a challenge for the regulator, as firms might try and inflate it.

In addition to any attempts at accounting manipulation, the absence of norms with regard to the treatment and reporting of different aspects of costs (or indeed explanatory variables in a cost modelling framework) may also hamper the comparability of firms. For instance in its PR13 efficiency benchmarking of Network Rail, ORR notes that some expenditure on signalling, telecoms, civils, plant and machinery and other cost categories are considered as renewals by Great Britain's infrastructure manager while other countries had been classifying these as enhancements. This may be a problem that is most prevalent in international benchmarking, but even domestic benchmarking in the UK has been hampered by problems of data comparability and definitions.

Cost performance incentives

Finally, cost performance incentives may differ between OPEX and CAPEX. Once the benchmarking analysis has been done (bottom-up and top-down), and efficiency gaps between firms established, the regulator faces a number of choices in how to translate the results into a set of efficiency targets. UK regulators typically use glide paths to give firms time to reach the efficiency frontier, and have also instituted measures to ensure that firms have incentives to continue improving efficiency throughout the five-year control periods. The latter is achieved through allowing firms to keep the benefits of savings for, say, five years, irrespective of whether the savings are delivered in year one or year five of the

control period. Ofgem has used this approach, referring to it as "rolling efficiency mechanism" (Ofgem, 2009). This approach has also been used in Australian energy regulation, where it is referred to as an Efficiency Benefit Sharing Mechanism (EBSS) (Electranet, 2015).

Another related issue with respect to CAPEX in particular, is the degree of certainty that firms (and regulators) may have with regard to the level of activity needed and the unit cost of such activity. Regulators have therefore instituted various forms of what might be called "sliding-scale regulation" that essentially share the rewards (and risks) of exceeding or (falling short) of efficiency targets; possibly supplemented by a menu-based approach – discussed further below. As a result, there could be weaker incentives to improve efficiency in CAPEX than OPEX, a problem noted by Ofgem (2008)⁹: "DNOs bear the full cost if they spend GBP 1 of additional OPEX but only 29p to 40p if they spend GBP 1 of additional CAPEX" thus creating a strong imbalance in terms of incentives.

Overall, the above three policy decisions (with respect to cost recovery, cost benchmarking and cost performance incentives) tend to suggest that incentives to improve efficiency of CAPEX are likely to be weaker than OPEX and that it will be more difficult to identify efficiency differences between firms for CAPEX than OPEX. These factors could encourage firms to implement more capital intensive solutions than is optimal, or even game the definitions of the two cost categories; first because they consider that they can earn a return on capital expenditure through the RAB (if the allowed rate of return is higher than the cost of capital); second because an increase in OPEX can lead to a decrease in relative efficiency when operational expenditure are benchmarked, which can lead to bad publicity in terms of ranking; and third because firms may prefer a weaker set of incentives, with the ability to share risks with consumers.

It should of course be noted that these arguments are not totally clear-cut; firms that can easily implement cost reductions may prefer stronger incentives as it creates opportunity for profits. It is clear though that this issue has been identified by regulators as a serious problem, requiring a set of solutions (discussed below). As pointed out by Ofwat (2011), the fact that a regulator had "measured and incentivised OPEX and CAPEX separately builds up a world between the two and can limit flexibility."¹⁰ It may also be considered that, due to the engineering culture within firms, there is a preference for asset-based solutions.

The above discussion of potential biases in the system has been supported by some anecdotal evidence – that is, there are OPEX/CAPEX trade-offs that may emerge in practice. In the water sector, considering the choice between building a reservoir and relying on bulk water trading, companies felt that the asset-based solution - a reservoir - was less risky in terms of cost forecasts than relying on trading. And once the expenditure was added to the asset base, the regulated firm would earn a return on the investments, while the OPEX generated by trading could count against them in the future assessment of the regulator. Also companies would build larger sewers rather than use sustainable drainage systems because "the regulatory framework incentivises CAPEX rather than OPEX solutions."¹¹ Of course similar trade-offs exist in transport, for example between maintenance (operating cost) versus renewals (capital cost) on the railways; or making investment in capacity versus the (operating) cost of traffic management. The evidence for capital bias is discussed further below.

The overall challenges facing the RPI-X regulator system are best summed up in a review of incentive regulation of electricity distribution and transmission networks by Joskow (2008): "In the UK, for example, the initial failure of regulators to fully understand the need for a uniform system of capital and operating cost accounts as part of the foundation for implementing incentive regulation mechanisms has placed limitations on their effectiveness and led to gaming by regulated firms (e.g. capitalizing operating costs to take advantage of asymmetries in the treatment of operating and capital costs)."¹²

The shortcomings of the building block approach have prompted the water and the energy regulators in the UK to introduce two new regulatory mechanisms, namely TOTEX regulation and menu regulation, with the objective to limit capital biases and improve efficiency incentives and performance more generally. New approaches to cost recovery have also been implemented.

Mitigation of capital bias

A number of regulators have moved towards adopting some form of what is described as TOTEX-based regulation, instead of dealing separately with capital and operational expenditure. This has three elements: a change in the approach to cost recovery; a change to the benchmarking framework; and the use of menus (operating at a TOTEX level).

Regarding cost recovery, Ofwat¹³ introduced for its 2015-20 price review a new "pay-as-you-go" (PAYG) mechanism, where companies suggest the share of expenditures which are to be recovered over the price review (that is, some share of the pooled capital expenditures and operational expenditures is recovered over the control period). The objective is to reduce the distinction between OPEX and CAPEX. The rest of the expected capital expenditures (outside the PAYG ratio) are added to the regulatory asset base. Interestingly, it is not physical assets that are allocated to one mechanism or the other, but shares of expenditure. This allocation should reflect the share of capital expenditure that benefits current or future consumers; thus achieving a balance between affordability (by users) and the ability of the company to finance its activities (and risk). According to the regulator, this approach should both mitigate the bias towards capital expenditure and simplify the approach (see Figure 3).



Figure 3. Ofwat Regulatory Approach (2015)

Note: RCV - Regulated capital value.

Source: Ofwat PR 14.

Where TOTEX regulation has been adopted, regulators have often adapted the tools to measure efficiency since "any efficiency analysis relevant from an economic perspective should ideally start from an explicit identification of the economic objective and market situation of the actor under evaluation.¹⁴" Therefore regulators have switched to TOTEX benchmarking in order to avoid the problem that only OPEX is properly benchmarked in regulatory efficiency assessments. In its 2010 guidelines, OFGEM¹⁵ stresses that "the ambition [is] to avoid biasing the network company into particular solutions (e.g. CAPEX solutions over OPEX)". At the 2008 Periodic Review, ORR adopted a similar approach, benchmarking maintenance and renewals expenditure together in a single econometric model.

There are multiple benefits of TOTEX benchmarking (Smith, 2008; 2012). To be precise, by TOTEX benchmarking we mean that OPEX and CAPEX costs are added together and benchmarked at this aggregated level. As discussed in more detail below, TOTEX may be broad – including all OPEX and CAPEX, or only some aspects of OPEX and CAPEX (e.g. in rail in Great Britain maintenance and renewal costs were benchmarked together; see Smith, 2012). Firstly, given that there are likely to be capital substitution effects between OPEX and CAPEX, this approach internalises those, taking the overall TOTEX measure as the cost variable to be benchmarked. It also avoids biases in the individual OPEX and CAPEX models if these capital substitution possibilities are not considered (which is hard to do).

Second, there are often definitional differences in the way companies may classify certain cost items, for example, the cut off between maintenance and renewals. This was certainly an issue that emerged during the 2008 review of Network Rail's efficiency, with different countries adopting different approaches to the boundary between maintenance and renewals (and indeed between renewals and enhancements). Whilst such problems should be less of a concern for regulators using domestic benchmarking – where regulators can set common definitions – our discussions with UK economic regulators suggest that achieving common and consistent data between firms and over time is one of the biggest challenges in economic regulation even in such circumstances.

Third, it allows the overall level of capital related activity to be benchmarked within the framework; as compared to separate OPEX and CAPEX approaches, where in many cases only the unit costs of CAPEX activities are benchmarked directly, leaving the level of activity to be reviewed through regulatory challenge of company business plans, which is more subjective.

There are challenges, however. The fundamental problems of lumpy renewals and the problems of comparing bespoke enhancement levels and projects across companies do not go away. Further, aggregate cost categories may be harder to benchmark: separating OPEX and CAPEX could permit a more targeted (and different) set of variables to be included in each of the separate models. In some cases, enhancement expenditures are taken out of the TOTEX models; this approach being referred to as BOTEX¹⁶ benchmarking (where BOTEX = TOTEX – Enhancement CAPEX). These issues are discussed further in the section "The effectiveness of efficiency benchmarking in transport".

Finally, regulators have introduced menus into economic regulation, based on the pioneering theoretical work of Laffont and Tirole (1986). Menus are particularly useful in situations where there is some uncertainty about the true, efficient level of costs for regulated firms. Such uncertainty is likely to be greatest in respect of capital expenditure. Under this approach firms are offered a menu of contracts, rather than a one-size-fits-all approach. The menu should give the correct incentives to encourage firms to choose the menu that best fits their own cost characteristics – giving a truth-telling incentive. Thus, firms that know they are low cost/efficient will choose that part of the menu that gives higher powered incentives for cost reduction; whereas high cost/inefficient firms may want to choose a lower powered mechanism that shares more of the risk with consumers. Depending on the precise details of the process, the menu approach could be used to encourage firms to produce accurate business plans;

though Ofwat determined that there were better ways of achieving that aim (through a specific business plan quality incentive); see Anglian Water (2015).

Importantly, in the water case the menu was set at the TOTEX level, thus creating the same cost efficiency incentives on both aspects of expenditure.

Box 1. Outsourcing of capital work

Beyond accounting differences, the split between maintenance and renewals may also correspond with organisational differences on how the works are carried out. The privatised rail infrastructure company Railtrack initially contracted out all maintenance and capital work. Later maintenance was brought in-house. Currently 70% of renewals are carried out by Network Rail's own High Output team (a team of 1 200 people, of which 50% are sub-contractors), with the remainder contracted out. Enhancement work is contracted out, under the oversight of Network Rail's infrastructure projects team who are responsible for design and delivery management, and procurement; leaving design and construction to be delivered by contractors. There are proposals to increase contestability in this process, whereby third parties could compete with Network Rail to offer the design and delivery management and procurement functions (The Hansford Review, 2017).

While outsourcing may be a way to increase efficiency through competitive tendering, procuring renewals entails challenges for the infrastructure manager. The latter needs to develop a long-term relationship with its contractors either for safety concerns or to foster long-term investments and innovation. As noted, half of Network Rail's High Output team in charge of rapid track renewal of the network are subcontractors; who have to operate complex track renewal trains owned by Network Rail.

A benchmarking exercise by the Swedish Transport Administration on the procurement of rail infrastructure projects details some of the contractual tools Network rail has implemented to increase co-ordination. For instance, in order to enhance early contractor involvement and co-ordination between contractors and consultants, Network Rail has introduced an alliance structure with its own reward system. The emphasis put on co-ordination means that the infrastructure manager does not necessarily aim at setting high powered contracts, but favours risk sharing mechanisms to set remuneration.

More generally, the report highlights that out of the six infrastructure managers studied, those of Great Britain, the Netherlands, Norway and Sweden have shifted towards less rigid forms of contracting when procuring work on the network. This shift translates into new tools being implemented to increase the early involvement of contractors as well as co-ordination. For more complex projects, the financial incentives may be loosened in order to avoid too frequent renegotiations. The two other countries in the benchmark - Germany and Switzerland - favour high-powered contracts and contractors are less involved in the design.

Sources: <u>https://www.networkrail.co.uk/high-output-rapid-renewals-railway/;</u> "Procurement of railway infrastructure projects – A European benchmarking study", Trafikverket (2016); <u>https://trafikverket.ineko.se/</u><u>Files/en-US/15106/Ineko.Product.RelatedFiles/2016 121 procurement of railway infrastructure projects.pdf</u>

A final point should be made. We have been discussing the efficiency incentives at the level of the regulated firm but in many cases regulated entities carry out some work in-house, whilst other aspects are contracted out. Of course the outsourcing decision is one for companies to make, though regulators may form an opinion regarding the extent to which the company is appropriately managing its supply chain and whether its procurement processes are efficient (if not, it would be expected that this may be

picked up through the top-down benchmarking exercise). Box 1 provides some summary information concerning contracting out in the rail and water sector (for reference, given the wider purpose of the Working Group to consider different approaches to the provision of transport infrastructure).

Summary and concluding remarks

It is clear from the above discussion that RPI-X regulation, combined with some form of benchmarking framework, has delivered significant efficiency savings across a number of sectors in the UK. This form of regulation has been widely adopted and is seen to be a high powered incentive regulatory framework, overcoming many of the problems of the rate of return approach to regulation.

However, it is also clear that the problem of capital bias is seen to be a significant problem, and a number of solutions have been considered. Regulators' assessments of capital bias do appear though to be based on considering economic incentives and looking at anecdotal evidence. We are not aware of clear, quantitative evidence of this problem. In the section "Evidence for capital bias", we further explore whether it is possible to find evidence for the A-J effect and capital bias more generally, looking at the wider academic literature, and also by studying data from the three case study countries.

Apart from problems relating to capital bias, there are a number of wider challenges in developing a benchmarking framework to challenge companies and ensure efficiency across all categories of expenditure. A number of approaches have been developed to address those challenges. These issues will be explored in detail via the three case studies in the final section. Through considering the experience of economic regulation, and benchmarking in particular, in the three case studies we aim to answer the question posed in the introduction, namely: to what extent can benchmarking work in transport?

The system of regulation in case studies: Rail, roads and water

In this section we provide a high-level overview of the regulatory framework in the three case studies. The details of the benchmarking framework are set out and discussed in the following section, where we explore the experience and solutions in greater depth.

Rail infrastructure

In Britain the railways were one of the later industries to be privatised. In order to support as much competition as possible, the railways were radically restructured, transforming a vertically-integrated state-owned monopoly into a vertically-separated system. The rail infrastructure was hived off into a legally-separate body, Railtrack, in 1994, which was later privatised by public flotation in 1996. Railtrack was later placed into administration (see the previous section) and replaced by Network Rail which as of

2018 is a state-owned company (after a period from 2002-14 in which it was technically a private firm, though with no shareholders, funded entirely by debt).

As noted previously, Network Rail is regulated by an independent economic regulator, ORR, according to the traditional building block approach (see Figure 2). After the first control period (1996/97 to 2000/01), where renewals was remunerated on a pay-as-you-go basis, renewals are now capitalised and added to the RAB, along with enhancement expenditure. In line with other regulated companies, Network Rail is subject to RPI-X regulation, with efficiency targets set for each element of expenditure (operations, maintenance, renewals and enhancements). Whilst in 2008 a "TOTEX" econometric model was developed that combined maintenance and renewals (M&R) together, bottom-up evidence was also used and ultimately different targets for maintenance and renewals were set. Operations costs and enhancement targets are also set based on a separate approach that is tailored according to the cost category. Benchmarking approaches are described in more detail in the final section of this paper.

There is no direct use of menus/efficiency sharing mechanisms; except with regard to regulatory pressures for Network Rail and train operators to join forces to develop whole-industry cost reducing approaches (then with incentives to share the benefits of these joint initiatives). A key issue in respect of Network Rail has been the problem that there are no shareholders, which reduces the incentives for out-performance or the dis-incentives for under-performance.

A defining, and perhaps unique, element of Network Rail's regulation to date has been the process used to reconcile the aims and funding positions of DfT (and devolved bodies) versus the ORR's independent assessment of the amount of money needed for an efficient infrastructure manager to sustain and develop the network. A problem emerged during the 2003 Interim Review, when ORR determined that Network Rail required a very substantial increase in funding. Thus a position emerged whereby an independent body was essentially deemed to be determining government expenditure levels.





A new process was thus developed to try to address the above problem (see Figure 4). Under this approach the government (DfT and devolved bodies) determine what they want to buy from the railways (the High Level Output Specification [HLOS]) and also how much money they have available (Statement of Funds Available [SOFA]). The ORR then independently assesses how much it should cost an efficient

infrastructure manager to deliver the HLOS. If this amount is greater than the SOFA then either the government needs to provide extra funding, or re-specify what it wants to buy from the railways. In the event that the government does not make that decision, the final decision would fall to ORR.

Of course this approach is potentially powerful and could be seen to address the problems that have emerged historically with state owned railways whereby governments provide less money to railways (due to funding constraints) whilst simultaneously requiring the same outputs to be delivered (though without explicit transparency of this situation). Under the current approach in Great Britain, such an approach would not be possible – government cannot set both price and quality. Of course, government could bring pressure to bear on the regulator, though in theory the regulator is independent (but could be subject to capture by government potentially).

Such an approach is also consistent with the notions of efficiency pressure and financial equilibrium in EU legislation (Recast of the 1st Railway Package) and other railways across Europe are developing approaches that are similar in nature (though typically with less power given to regulators). Germany now has a similar approach to that in Great Britain, with efficiency pressure applied through RPI-X regulation on access charges, but also a multi-annual agreement between government and the infrastructure manager setting out performance targets and agreed funding levels (with a reconciliation mechanism (Nash et. al., 2017)).

It should be noted that from 1 September 2014, Network Rail has been re-classified as a state-owned company, meaning that all of the company's debt is classed as government debt. Formally speaking this does not alter the independent regulatory framework and ORR is still required to scrutinise Network Rail's efficiency. However, the change occurred during control period 5 (starting in 2014/15) and the company's funding was impacted by cash constraints imposed by government within the control period (now more sensitive to Network Rail's borrowing). Going forward, ORR will still play a role in ensuring that Network Rail plans for the longer term and where funding constraints impact on the company's plans, the independent regulator would play a role in ensuring that performance targets are adjusted to reflect the impact of funding constraints.

Finally, in 2017 plans were announced to treat enhancements separately from renewals, with the Department for Transport (DfT) taking responsibility for the regulation of enhancements outside of the regulatory process. Previously DfT (along with devolved government bodies) funded enhancements (as they will continue to do), but the determination of Network Rail's overall, efficient expenditure allowance (including operations, maintenance, renewals and enhancements) for a control period was carried out as part of the regulatory review process, overseen by ORR. The aim is to permit greater flexibility – thus allowing the right amount of scrutiny/approval to take place at the right time, depending on the degree or surety about the costs, rather than being tied down to a five-year regulatory cycle. The proposals have not yet been finalised but one risk is that there is no independent scrutiny by ORR.

Water

Since the privatisation of the former Regional Water Authorities by the Water Act 1989, the water and sewage industry in England and Wales has consisted of a relatively small number of private monopoly companies. The Act created ten large water and sewage companies (WaSCs) – the former Regional Water Authorities – and also allowed the remaining statutory water companies – smaller, privately owned water only companies (WOCS) – to convert to ordinary public or private limited companies, lifting their restrictions on profits and dividends. At the time of privatisation, there were 29 WOCS and

ten WaSCs. Since then, several mergers have taken place, some between a WaSC and a WOC, and others between WOCs, currently reducing the number of companies to eight WOCS and ten WaSCs.

Given the natural monopoly nature of the industry, and the lifting of previous restrictions, the need for some form of regulation of the industry was recognised. Initial price caps were set by the government for a five-year period, and an independent regulator was established, as in telecoms, electricity, and gas supply industries, with responsibility for subsequent price reviews. This was initially the Director General of Water Services, supported by a team referred to as the Office for Water Services (Ofwat). This was subsequently replaced in 2006 by the Water Services Regulation Authority, also known as Ofwat.

The regulatory regime in the water and sewage industry is an adaptation of the RPI-X regulation proposed by the Littlechild report, and is referred to as RPI+K regulation. The K factor, reflecting the maximum allowed price increase above inflation, consists of two components, such that K=Q-X. In keeping with the terminology in other regulated industries, the X factor reflects the regulator's view of potential productivity growth. The Q factor, unique to water regulation, is an allowance reflecting the additional investment needed to increase or improve outputs. The inclusion of this term is explained by the need to make up for historic underinvestment in the pre-privatisation period, and the need to ensure that the industry complies with UK and EU drinking water and environmental standards and regulations.

Ofwat's initial price controls essentially followed a "building blocks" approach as described above, in which firms' revenue allowances are calculated by summing allowed OPEX, and allowance for depreciation, and an allowed return on capital. Allowed OPEX was determined partly through top-down econometric benchmarking exercises, while the allowed return is calculated by multiplying a company's weighted average cost of capital (WACC), determined by Ofwat, by the company's regulatory capital value (RCV). As discussed in detail above, this method potentially incentivises a bias towards CAPEX because of the difference in the way the expenditure categories are treated. The K factor was then calculated as the weighted average increase in the tariff basket that will be needed to attain the revenue allowance. Ofwat's tariff basket consists of the following outputs:

- unmetered water supply
- metered water supply
- unmetered sewerage services
- metered sewerage services
- receiving, treating, and disposing of trade effluent.

Recent years have seen a number of significant regulatory changes in the water industry, as in other regulated industries. Significantly for the issue of capital bias, Ofwat introduced a capital incentive scheme (CIS). Under the CIS, companies receive rewards for outperformance of their CAPEX forecasts from their business plans, or are penalised if they fail to meet these targets. Added to the standard "building blocks" approach outlined above, the CIS should in principle mitigate against any potential CAPEX bias. Another innovation for PR09 was the introduction of a revenue correction mechanism (RCM). The RCM is a reduction or increase in allowed revenue at the next price review by the extent of over- or under-recovery of revenues in the present price review.

From PR14, there separate price controls have been set for wholesale and retail activities. This anticipated the introduction of competition into non-household retail activities from 2017-18 onwards, and the likely further extension of competition into household retail activities (on which, at the request of the government, Ofwat published a cost-benefit assessment¹⁷). This separation follows the same model as in, e.g. the electricity and gas supply markets, in which monopoly distribution companies

charge wholesale rates to suppliers in a potentially competitive retail market. Previous to this, Ofwat had been unusual among UK regulators in setting no separate wholesale price controls, as shown by Table 2.

Sector	Wholesale controls?	Retail controls?	Control period	Form of Control
Water and sewage	No – before PR14 Yes – since PR14	Yes	5 years	RPI+K
Telecoms	Yes	No	Various	LRIC (Long-run incremental cost)
Electricity	Yes	No	Before 2015-16: 5 years From 2015-16: 8 years	Before 2015-16: RPI-X From 2015-16: RIIO
Gas	Yes	No	Before 2013-14: 5 years From 2013-14: 8 years	Before 2013-14: RPI-X From 2013-14: RIIO
Post	Yes	Yes	Various	RPI-X (for retail)
Rail	Yes - track access and station long-term charges	Yes - some fares	5 years	RPI-X
Airports	Yes	No	5 years	RPI-X

Table 2. Comparison of UK price controls

Source: Adapted from Ofwat discussion paper.¹⁸

PR14 also saw a further strengthening of CAPEX incentives (following on from the introduction of the CIS in PR09). Ofwat moved from its traditional building blocks approach to a form of TOTEX regulation. Under this approach, illustrated in Figure 3, benchmarking is based on combined TOTEX, and a proportion of TOTEX is remunerated on a PAYG basis informed by the TOTEX benchmarking. The remainder is then added to the RCV, against which an allowed return is calculated using Ofwat's estimate of the WACC.

Further innovations for PR14 were the introduction of performance commitments (PCs) and outcome delivery incentives (ODIs) – rewards and penalties relative for performance relative to these PCs – to incentivise delivery of outcomes, and the introduction of menu regulation in which the firm is presented with a menu of regulatory contracts which incentivises them to reveal information to the regulator.

Roads

The roads sector in England is interesting because it is undergoing reform, partly to address a long-standing capital investment issue. This has involved the formation of either an economic regulatory structure (in the case of strategic roads) or "quasi-economic regulation" as discussed below (in the case of local roads). Both structures necessitate a role for top-down benchmarking. In England, the roads infrastructure can be grouped into two categories:

- Strategic Road Network (SRN) covering motorways and "trunked" A roads (A roads of national importance) operated and owned by Highways England
- Local Highway Network (LHN) operated and owned by 152 local authorities.

As shown in Table 3, the SRN is a relatively select length of network relative to the total road length in England. The SRN does operate at a substantially higher traffic density than the average density on the local road network, although on the local road network averages mask the fact that some sections have high traffic densities, particularly on urban corridors.

Road class	Road type		Route length (miles)	Route length (km)
Strategic Road	Motorways		1 600	2 600
Network	Smart Motorways		200	300
	Trunk Roads		2 600	4 200
	Total for SRN	4 400		7 100
Local Highway	Principal Roads	A-Roads	17 500	28 200
Network	Minor Roads	B-Roads	95 900	154 300
		C and U-Roads	70 700	113 800
	Total for LHN	184 084		296 300
England total		188 500		303 400

Table 3. Road lengths statistics in England

Source: Department for Transport (2017).

Given the different ownership models for the SRN and LHN, difference governance and regulatory structures exist for each. Each is described below.

Strategic Road Network

Highways England is a government-owned company limited by guarantee. As such it is at "arm's length" from government. This is similar to the situation of Network Rail between 2002 and 2014 (see section "Rail infrastructure"). Highways England was formed in 2015 and before that was known as the Highways Agency and was a direct agency (arm) of government.



Figure 5. Regulatory and governance structure for Highways England

Source: Office of Rail and Road (2015).

The new arrangements for Highways England necessitate oversight of the company (Figure 5). Broadly speaking there is a similar arrangement for the SRN. In rail in the United Kingdom, there is the funder, the Department for Transport, the infrastructure manager, Highways England, and an independent "Monitor", Highways Monitor (in addition there is a user representative, Transport Focus). Highways Monitor is part of the Office of Rail and Road. It is not a strict economic regulator, as it only advises Government on issues such as efficiency, however it discharges similar activities. Importantly for this study, the Highways Monitor is expected to monitor the efficiency performance of Highways England and undertake comparative benchmarking.

The Road Investment Strategies (RISs) are five-year funding and key output specifications (including specification of enhancements). These are similar to control periods in other regulated sectors. As such, at least from the perspective of the need for efficiency benchmarking the SRN infrastructure manager can be thought of as subject to incentive regulation, similar to the rail and water sectors in Britain.

Local Highway Network

There are 152 local authorities with responsibilities for the LHN in England. The majority of funding comes directly from central government.¹⁹ The sources of funding are summarised in Figure 6 and Table 4 summarises the key sources of funding for the LHN.





Note: NPIF – National Productivity Investment Fund.

Source: Kemp (2017, p. 15).

Historically, there have been limited direct efficiency incentives specifically targeted at the highways functions of local government. However, it should be noted that because the "Needs" category of funding (see Table 4) is not ring fenced for highways, then the general budgetary pressures faced by local

government are relevant to incentivising better practice. However, the general principle that local government is responsible for choosing the services they provide and ensuring value for money drove a hands-off approach (from the perspective of central government) at least with respect to expenditure other than "major schemes", i.e. enhancements.

Fund name	Description	How allocated	Expenditure type	Ring fenced
Needs Fund	Accounts for network size to distribute the DfT funding for Local Highway Authorities (LHAs) circa GBP 800 million per annum (see Figure 6)	Needs formula	Maintenance and renewal	Νο
Incentive Fund	Self-assessment with 3 Band outcomes (1, 2 or 3) GBP 578 million in total	Formula (Band determines %)	Maintenance and Renewals	No
Pothole Action Fund	GBO 250 million from 2016 to 2021 to target potholes	Formula	Maintenance	Outputs (potholes filled) have to be delivered
Challenge Fund	Challenge Excludes lighting scheme but otherwise used to bid for projects not possible to fund using Needs Fund		Major renewals and enhancements	Yes
NPIF	GBP 1.1 billion - aims to focus on housing, growth and tackling barriers that impede growth of productivity	Bidding process	Enhancements	Yes – However can cover non-highway expenditure if included in the bid

Table 4. Explanation of funding sources for the Local Highway Network

Source: Mayat (2017) analysis of Kemp (2017).

In the most recent decade, however, there has been a push to greater scrutiny of local authority highways expenditure and efforts to incentivise better practice. The broader context was a funding environment for Local Government which was reduced in real terms following the 2008 recession. Two key developments can be identified:

- 1) The Highways Maintenance Efficiency Programme (HMEP) was a central government instigated and initially funded initiative from 2011 and aimed at sharing and enhancing practice in the highway sector. As part of the work programme, HMEP undertook two pilot studies on top-down statistical benchmarking for the LHN.
- 2) From 2016/17 an element of the central government funding "Needs" fund was replaced by the "Incentive" fund and by 2018/19 the equivalent of 15% of the Needs fund is to be replaced with the Incentive fund. The defining feature of the Incentive fund, compared to the Needs fund, is that local authorities have to demonstrate they are undertaking good practice in their activities, through submitting evidence via a self-assessment process. As part of the self-assessment, authorities are required to demonstrate that they are benchmarking their performance and learning from other local authorities. Authorities are graded into three bands (1-3) and each band attracts an increasing proportion of the available funding.

Overall the introduction of the incentive fund has led to quasi-economic regulation of local authorities' highway spending, as an element of maintenance and renewals funding now depends on local authorities demonstrating that they are making improvements.

In order to facilitate benchmarking between local highway authorities, the top-down statistical benchmarking pilot in the HMEP project has been continued, under the CQC Efficiency Network. That initiative is funded through subscriptions from participating local authorities. As of 2017/18 year of analysis, there are 90 local authorities (comprising over 75% of the local road network in England) participating.

Evidence for capital bias

Introduction

As noted in the section "Theoretical and regulatory literature", the A-J effect has been established as a potential (theoretical) problem that may emerge in systems of rate of return regulation if the allowed rate of return is above the firms' cost of capital. Indeed as noted, even within the RPI-X system, the use of a RAB, combined with an allowed rate of return to re-base prices at five yearly intervals, means that there is some (though more limited) scope for the AJ effect to pose a theoretical problem in that context as well. The AJ problem is stated (theoretically) in economics textbooks (see, for example, Viscusi et. al., 2005) where it is also noted that to the extent that a capital-intensive production technology is associated with technological progress, the overall effect of the AJ effect could even be beneficial.

Long-quoted (supposed) examples of the AJ effect that are often referred to include:

- the tendency for US airlines to use larger aircraft than needed (this being corrected by the de-regulation of the industry in the 1980s)
- the provision of capacity to meet peak demand by individual electric utilities in the United States, where such capacity could be shared between firms.

It should be noted that questions have also been raised as to whether firms really act in such a sophisticated way in practice, given the natural objective of regulators to try to prevent such activity.

However, Law (2014) sets out a much more comprehensive attempt to consider the evidence for the AJ effect. This paper points out a number of challenges in providing conclusive evidence on the existence of the AJ effect internationally based on a review of 192 empirical studies (since 1962). Key problems that they highlight include:

- the fact that models are built using a single period of data, rather than data over a number of years; whereas effects might be considered to impact on firm behavior in complex ways over time, and taking a single period snapshot could be misleading
- assuming exogeneity of input prices which could be unrealistic in regulated industries

- failing to explicitly model regulatory constraints and how they impact on the production technology and the regulated firm's cost function. The assumed regulatory constraint may not fully capture the full nature of the interaction between the firm and regulator (including actions that regulators may take to limit the AJ effect for example)
- the problem that at least post-investment in capital, capital and labour may actually be complementary rather than being substitutes.

More generally the paper concludes that detection of the AJ effect is challenging from an empirical perspective – and indeed that there are likely to be many other incentives and outcomes arising from complex regulatory mechanisms; thus leading to the overall finding that the AJ effect should certainly not be assumed and that further investigation of its existence is probably not a fruitful line of research.

We thus take the above literature as our starting point. As discussed above, additional reasons for capital bias have been added to the AJ effect, resulting from the particular way in which OPEX and CAPEX are treated within the RPI-X regulatory framework. Therefore, below we present data on the evolution of OPEX and CAPEX for the three case studies, providing a brief commentary on the data in each case. Overall, given the past evidence, and the discussion above, it could be seen to be more plausible that the latter form of capital bias could be more important than the AJ effect, particularly as incentive-based regulation is generally considered to weaken AJ-related biases.

Before proceeding it is worth noting a possible complication in the definition of CAPEX and OPEX in network industries. In network industries, network replacement (or renewal) which is ultimately capitalised and added to the RAB will involve the use of energy, materials, plant and labour. Thus measuring the observed CAPEX and OPEX expenditure lines in a company accounts is not the same as observing the balance between use of labour and plant in the "production process" of implementing the renewal. What is at issue here is what we consider to be the inputs and outputs. If we consider the delivery of the renewal project to be the (intermediate) output, then the inputs are labour, energy, materials and plant (capital), where the plant may refer, for example, to the machinery used in the renewal process.

However, the wider regulatory process actually sees the outputs of a network industry to be, for example, the number of customers served, or train-km moved over the system. In that context, the inputs can be thought of as the capital (including plant, but also the capital cost of renewing and enhancing the network itself), combined with expenses required to operate and maintain the capital. Thus our analysis of the data will reflect the OPEX and CAPEX reporting categories used in each case study, as set out below. As noted above, the treatment of some aspects of CAPEX, particularly renewals, in the regulatory system will also affect a company's incentives with respect to its choice between the two different types of expenditure.

Rail infrastructure in Great Britain

Figure 7 below shows the evolution of infrastructure expenditure in Great Britain with costs divided into four categories: operations costs (which includes signalling and day-to-day operations), maintenance of the infrastructure, renewals and enhancements (where the latter either refers to new lines or where the capability or capacity of the infrastructure is enhanced in some way). The vertical dotted lines on the chart refer to the (usually) five-year control periods.



Figure 7. Rail infrastructure expenditure in Great Britain

Source: Regulatory Financial Statements, Network Rail.

A number of points can be noted from this chart. First of all, the data covers the period post-privatisation and needs to be considered against the backdrop of the general concern that rail (and other formerly nationalised industries) had struggled to obtain sufficient funds from government to invest in the network. As an indicator of renewal activity, Figure 8 charts a longer-term series of the volume of track renewals. It is apparent that a considerable amount of track renewal activity took place in the 1970s, followed by a sharp fall in the 1980s and in particular in the run-up to privatisation in the mid-1990s. It was therefore expected that in the early years after privatisation there would be some need for an increase in renewals activity. Indeed, as noted above, initially after privatisation, rail renewals were remunerated on a pay-as-you-go basis, so during that early period there would have been no A-J-related capital biases and the regulatory framework would have encouraged efficiency savings in both OPEX and (renewals) CAPEX alike.

The evolution of renewals expenditure over the subsequent years was driven most significantly by the reaction to an accident at a place called Hatfield in October 2000. An intercity train de-railed, with four people losing their lives, as a result of track that was in very poor condition. This accident highlighted a wider problem that the privately-owned rail infrastructure manager, Railtrack, did not have good knowledge of its assets. A fundamental change in the maintenance and renewals programme resulted, and the subsequent very large increase in costs, made worse by over-runs on a major enhancement programme on the West Coast Mainline, led to Railtrack being placed into administration, and being replaced by a not-for-profit company, Network Rail. For further details see Smith (2006) and Kennedy and Smith (2004). Initially Network Rail was technically a private company, limited by guarantee, and funded entirely by debt (with no equity shareholders).

From 2001/02, the regulator (ORR) decided to change the renewals accounting basis from pay-as-you-go to one in which renewals would be capitalised and added to the RAB. Whilst ORR recognised that this might reduce incentives to improve efficiency with respect to renewals it was deemed necessary because of the uncertainty created by the aftermath of the Hatfield accident. ORR thus switched from pay-as-you-go mechanisms to capitalisation of expenditure because of potential "unanticipated changes in renewals requirements (e.g. because of the lack of information about the current state of the assets²⁰)."

As a result, during the period from 2001/02 (when Network Rail was formed) up until 2003/04, the normal RPI-X regulatory framework was essentially put on hold and it is generally recognised that during

that time cost control was very weak. As is clear from Figure 7, all categories of costs, operations, maintenance and renewals costs rose sharply during this period. Enhancement costs were in general relatively low, largely because of the very large increase in expenditure needed to sustain the existing network.

An interim review was concluded by ORR in December 2003 with the regulator seeking to disentangle the rises in costs that had occurred between inefficiency and genuine cost rises, resulting from the need to address the poor quality of the network. Efficiency targets were set at that point, though, as noted by Smith (2012), these were based on very limited evidence.

The next few years saw reductions in cost as Network Rail delivered improvements in efficiency (across all categories of expenditure) over the control period, in line with the targets set in the 2003 review, based on various benchmarking studies (it must be remembered that costs were being reduced from a very high base, having approximately doubled in the previous years). However, an international econometric cost study published in 2008 as part of the 2008 Periodic Review (Smith, 2008; 2012), reported a 37% efficiency gap between Network Rail and its European peers in respect of maintenance and renewals expenditure. Bottom-up engineering studies commissioned by ORR provided corroboration to the results of the econometric assessment. As a result, Network Rail was tasked with closing this efficiency gap, albeit over a period of ten years. The targets for maintenance and renewals for the five-year control period starting in 2009/10 was set at 18% and 24% respectively, with the balance to be achieved in the following control period (Smith et al., 2010). Similar, but slightly lower targets of 16% were set for operations expenditure.





Source: Smith (2006).

In broad terms the data then shows Network Rail starting to make some significant cost reductions, particularly in the first few years of the control period. However, by the end of the control period (2013/14) it was clear that the company was falling behind against its initial targets, though only by 3% (delivering 18% across its operations, maintenance and renewals cost categories, as compared to 21%; [ORR, 2013])²¹. ORR in turn set Network Rail a further target of 20% across those cost categories for the

control period starting 2014/15, thus staying true to the original ten-year assessment made in the 2008 Periodic Review.

Since 2014/15, far from improving efficiency, Network Rail's efficiency performance deteriorated. This was driven by a number of factors, but the failure of the company to deliver enhancement projects on time and to budget, led to a re-think about future enhancements (in particular electrification). Efficiency performance in respect of renewals also deteriorated. Figure 7 shows that renewals expenditures have fallen since 2013/14 but, as shown in Figure 9, this fall has been associated with a sharp reduction in activity, as evidenced by the reduction in track renewal volumes (thus revealing a substantial rise in unit costs).





Source: Own analysis, based on Regulatory Financial Statements and Annual Returns, Network Rail

ORR (2017) identifies a number of underlying factors driving this deterioration. In part Network Rail was under-prepared to deliver the planned volume of renewals and falling renewal volumes has knock-on implications for productivity. Further, management ability to focus on resolving the problem was hampered by the major problems that developed at the same time in respect of the enhancement programme. It also appears that Network Rail's efficiency improvement aspirations over this period were not founded on clear plans, though that does not necessarily explain the deterioration in efficiency that has occurred. Obtaining access to the network to do maintenance and renewals work in an efficient and orderly way has also been an issue, with apparent conflicts between the short-term needs of train operators and access to do maintenance and renewal work.

A further change during this time period was the re-classification of Network Rail as a state-owned company (from 1 September 2014). Since that date, in addition to the pressure emanating from the regulatory framework, Network Rail has been subject to cash limits which have also had an impact on its expenditure decisions. In particular, re-planning and profiling of renewals to fit within cash limits imposed by government impacted on productivity. Perversely, in addition, increased devolution of powers to Network Rail's regions (Routes) – see the final section of this paper – led to local enhancements being added into renewals schemes at the regional level but without the necessary budgetary control (enhancement budgets were not managed at the Route level (Bowe, 2015).

As can be seen from Figure 7, a further clear trend is that enhancement expenditure, particularly after 2008/09 has been rising sharply, as costs for operations, maintenance and renewals were finally starting

to fall, and ORR and DfT had agreed the HLOS/SOFA process to agree how the government's priorities for new investment would be operationalised and subject to appropriate regulatory scrutiny. As noted in the previous section, enhancements have always been regulated in a different way to the other expenditure categories and the latest 2017 proposals are for an even greater separation, with enhancements to be dealt with entirely separately by DfT, separate from the five-year regulatory control period.

The general picture that emerges in respect of Great Britain's railways is therefore a complex one – whilst RPI-X regulation, supported by top-down (and bottom-up) benchmarking exercises has been successful during some control periods, this has not been true in all cases; recently Network Rail's measured efficiency performance appears to have deteriorated markedly. Figure 10 summarises the OPEX/CAPEX split within maintenance and renewal (M&R) over the period. This shows some change over the period, driven by the changes in renewals and also increased enhancements discussed above – with the percentage of CAPEX being generally higher after the first control period (post 2000/01), though with some fluctuations and a recent reduction following the deterioration in renewals efficiency and the resulting lower renewals volumes.





Source: Own analysis, based on Regulatory Financial Statements and Annual Returns, Network Rail

It appears that, as indicated from the wider literature review carried out by Law (2014), there are many factors other than the AJ effect that have been impacting the OPEX/CAPEX split. It does not appear that AJ effects or other incentives that might bias towards CAPEX resulting from the regulatory regime have had much impact on rail expenditure patterns; these being much more driven by underlying factors. That said, during the period of Network Rail's stewardship of the network, and prior to its re-classification of a state-owned company, the ability of the company to borrow without the debt appearing on the government's balance sheet, may have ensured that Network Rail was able to select an appropriate level of renewals without constraints. It remains to be seen whether, under the company's new status, a more OPEX-oriented approach will emerge over the longer term.

Finally, we compare the split of maintenance versus renewals across European countries for the period 2005-14 (see Table 5). A number of points might be observed from this. First, to the extent that any of the railways in the sample are assumed to be optimising (though they may not be), it suggests that there is no, one optimal maintenance/renewal split. This is particularly so when looking at expenditure over a short time horizon, because renewals may naturally be expected to follow a cyclical pattern, thus making
it hard to establish the underlying, steady-state level of renewal in each country. Thus it may be expected that it is very hard to establish the presence of a capital bias using data of this kind.

The data does suggest, however, that Network Rail has a higher proportion of renewals than other countries. Given that infrastructure is essentially owned by the state in all of the other comparators, this finding could suggest that there is if anything an OPEX bias in the other countries – whereby companies are prevented from selecting an optimal level of renewals because of cash constraints. Further biases (within CAPEX) may occur in state-owned contexts. For example, where there is strong political involvement, politicians may prefer enhancements to renewals. As noted, Great Britain (with an independent regulator and up until recently with Network Rail's debt being treated as private debt), has not faced this pressure.

Table 5. Ratio of renewals expenditure in M&R aci	ross European countries (2005-14)
---	---------------------------	----------

Country	Average
Austria	57%
Belgium	58%
Denmark	56%
Finland	49%
France	43%
Germany	68%
Ireland	48%
Italy	41%
Netherlands	36%
Norway	49%
Poland	51%
Slovakia	41%
Slovenia	32%
Sweden	32%
United Kingdom	71%

Source: RMMS Reports, European Commission.

It is worth noting that Germany has a regulatory framework in which the regulator (as of 2018) plays a role in setting efficiency targets, but where a multi-annual agreement exists between government and the infrastructure manager, to determine the amount of funding needed (and a process to reconcile efficiency targets with expenditure needs at least in respect of maintenance (Nash et. al., 2017)). Table 5 shows that Germany has a similar OPEX/CAPEX split to Great Britain, which could therefore relate to the regulatory regime. It should be noted that we are not in a position to verify the accuracy of the data on which Table 5 is based, so we do need to treat any conclusions here with some caution. However, if that is the case, then the argument is more about avoiding an OPEX bias, than suggesting the existence of an inefficiency-inducing CAPEX bias. It is also known that France has seen relatively low renewals volumes for many years, but Figure 11 shows that this is now increasing; again, reinforcing the fact that renewal cycles, combined with cash constraints (or government priorities as between renewals and

enhancements) imposed by government, can substantially change the profiles of expenditures over many years. It is expected that renewals expenditure in France should continue to increase up to 2019 for the core network, before returning back to a steady state level.²²





Overall, our conclusion from this section is that, as expected from the past literature, it is challenging to identify OPEX/CAPEX biases relating to AJ or other regulatory incentives in company data. Many other factors impact costs and the balance between expenditure categories. In rail, the problem of past under-investment prior to privatisation, and stewardship problems in the early days of Railtrack, have combined to present a complex and challenging picture where renewals had to increase sharply, particularly after 2000. Regulatory pressure through benchmarking, within an RPI-X regulatory framework, has shown some successes in bringing all categories of expenditure down, but significant problems have emerged post 2013/14 in particular relating to CAPEX costs and Network Rail's ability to deliver renewals and enhancements. It is also clear that Network Rail has a higher proportion of renewals compared to maintenance than its European peers; but this may in part reflect cash constraints in other countries that Network Rail, pre-reclassification as a state-owned company, did not face. Different countries may also be in different positions on the renewals cycle, as indicated by increases in renewals in France in recent years.

Finally, it is clear that enhancements have been, and perhaps have to be, dealt with differently in the regulatory framework, and the level of spending in Great Britain depends on what government wants to buy from the railways. Changes in this category are therefore probably not influenced by CAPEX bias.

The water industry in England and Wales

Before presenting data with respect to the water industry in England Wales, the water regulator has considered there to be some anecdotal evidence for CAPEX bias, which has led to the introduction of regulatory changes (TOTEX benchmarking; the introduction of common efficiency incentives across OPEX and CAPEX, through the use of menus; and changes to cost recovery approaches). Ofwat (2011) also considers a number of possible ways of identifying CAPEX bias:

Source: French railway regulatory body (Arafer).

- By examining changes in the proportion of OPEX and CAPEX over time. If there is CAPEX bias (in either direction), we may expect to see changes in the ratio of CAPEX to OPEX over time, and especially as a result of any regulatory or environmental changes that we might expect to affect such bias.
- By comparing OPEX and CAPEX performance in terms of the extent to which firms have been under- or out-performing targets and forecasts for each expenditure category. The existence of CAPEX bias implies that there are incentives other than those intended by the regulator to under-perform or out-perform in certain areas. As a result, we may expect to see, for example, companies out-performing their OPEX targets while under-performing with respect to CAPEX.
- Evidence that the companies are developing low levels of potential OPEX -based solutions or if an OPEX -based solution is the best one for customers, evidence of a reluctance to carry out that solution such as choosing to increase sewer capacity over sustainable drainage systems. This solution would rely on anecdotal evidences and a detailed knowledge of the available alternatives.
- The level of transfers required between the two expenditure categories at price reviews or during the control period. A CAPEX bias should manifest itself in the companies' choice of capitalisation policy, preferring to classify expenditures as CAPEX in order to increase their revenues. If it is the case, the regulator should be able to catch part of this accounting game given their ability to challenge the classification of costs. And we should observe at price reviews, a transfer triggered by the regulator from CAPEX to OPEX. In the case of Ofwat, this transfer amounted to GBP 145 million of OPEX and a further transfer of GBP 1 billion of capital at the PR09.

In the following section, we mainly focus on the first two possibilities, and discuss the support (or lack thereof) for the existence of CAPEX bias in the reported data. The data used in the following charts are taken from a variety of sources, including the June Returns formerly collected by Ofwat until 2010-11, the Industry Facts and Figures published by Water UK (an organisation representing the WaSCs and WOCs in the UK) from 2011-12, the regulatory accounts published up to 2014-15, and their replacements, the Annual Performance Reports, published in 2015-16 and 2016-17. We were able to gather consistent data on OPEX and TOTEX, broken down into water and wastewater activities, covering all English and Welsh WaSCs over the 18-year period from 1997-98 to 2014-15, and more limited data for 2015-16 and 2016-17. This represents the vast majority of the industry in England and Wales in terms of customers.

Figure 12 shows the trend in TOTEX summed up across all WaSCs, broken down into OPEX and CAPEX, and additionally by service, over the entire period. Shaded bars represent the start of a new price review period. The chart covers the last two years of the PR94 control – Ofwat's first price control, following the earlier caps set by the Government in 1989 – the entirety of the PR99, PR04, and PR09 periods, and the first year of the PR14 price control. The figures are deflated by the RPI to 2016-17 prices. As can be seen, there is a considerable fall in real TOTEX overall in the first years. Following generous allowed price increases set post-privatisation, Ofwat imposed stringent real price cuts in PR94 and PR99, which helps to explain the reduction in real TOTEX observed in the first years shown in the chart. Thereafter, in PR04 and PR09, price caps were loosened somewhat.



Figure 12. Totex from 1997-98 to 2013-14, English and Welsh WaSCs

Source: Ofwat June Returns (1997-98 – 2010-11), Water UK Industry Facts and Figures (2011-12 – 2014-15), Regulatory Accounts (2011-12 – 2013-14).

It is immediately noticeable that CAPEX is more volatile than OPEX, and that this accounts for most of the time-variation in TOTEX. This can be explained in terms of the investment cycle and the inherent lumpiness of capital investments in the industry, and underlines the difficulty of identifying CAPEX bias in this way. Figure 13 shows the same data in percentage terms; i.e. the components of TOTEX as proportions of the whole. We see no dramatic change in the composition of TOTEX between 1997-98 and 2014-15. Again, there is some fluctuation in the shares, which may be cyclical – perhaps partly reflecting a regulatory cycle – but the share of CAPEX is somewhere between the 40%-50% mark throughout the period, and in fact is slightly lower at the end of the period than at the beginning, which clearly does not suggest a systematic bias towards CAPEX.

In Figure 14, similar data are shown for the years 2015-16 and 2016-17. Note the change in reporting from 2011-12 in Figures 12 and 13, so that OPEX is split into wholesale water and wastewater OPEX categories along with a third retail OPEX category as opposed to the two water and wastewater categories used previously; the totals remain comparable. However, in Figure 14 we can see an analogous change in reporting of CAPEX, however in this case there is an implicit retail CAPEX category that we are missing, limiting comparability with the 1997-98 to 2014-15 data.

We can, however, see from Figure 14 that there was little change in the composition of TOTEX (less retail CAPEX) from 2015-16 to 2016-17. It also seems reasonable to assume that there was no dramatic change in the missing retail CAPEX between 2014-15 and 2015-16. This is significant due to the fact that these are the first two years of the PR14 price control, which was the first to employ the new methodology, i.e. TOTEX benchmarking and PAYG recovery of TOTEX rather than OPEX alone (and the earlier PR09 control had seen the introduction of the CIS). Given that we might have expected these innovations to mitigate any CAPEX bias present, this may be an indication that there was no significant bias in the first place.



Figure 13. OPEX and CAPEX Shares 1997-98 to 2013-14, English and Welsh WaSCs

Source: Ofwat June Returns (1997-98 – 2010-11), Water UK Industry Facts and Figures (2011-12 – 2014-15), Regulatory Accounts (2011-12 – 2013-14).

It may be that, rather than affecting the shares of OPEX and CAPEX, there are biases affecting the nature of the CAPEX undertaken, e.g. that is a bias within CAPEX expenditure towards renewals and away from maintenance. Since enhancements may be driven by a range of considerations, including pressure to improve water quality, it might be that CAPEX bias reveals itself more in respect of a tendency to impact the split between OPEX and CAPEX renewals. That said, as noted earlier, Ofwat has identified substitution possibilities between OPEX and enhancement CAPEX (e.g. building a new reservoir).



Figure 14. OPEX and CAPEX Shares 2015-16 and 2016-17, English and Welsh WaSCs

Note: WaSC - Water and Sever Companies.

Source: Annual Performance Reports (2015-16 and 2016-17).

Ofwat (2011) identify several additional explanations for CAPEX bias, including "risk of failure and penalty strength" and "wider requirements and incentives". In the first case, it is argued that the existence of minimum standards could lead risk-averse companies to underperform in relation to CAPEX, particularly if the minimum standards are attached to strong penalties. In the second case, Ofwat acknowledge that secondary regulatory mechanisms set by Ofwat or other bodies may bias companies towards certain categories of spending. Both of these arguments underline the difficulties posed by the existence of multiple mechanisms and standards, which may in some cases create conflicting incentives. English and Welsh WaSCs and WOCs are all regulated not only by Ofwat, but also by the Drinking Water Inspectorate, the Environment Agency, and others. There are also the add-on mechanisms introduced by Ofwat themselves, such as the ODIs introduced for PR14.

To the extent that drinking water and environmental standard set by the DWI and the Environment Agency may be associated with severe penalties for underperformance (understandably, given the public health implications in many cases), a risk-averse company may be encouraged to overcapitalise. It is also possible that a risk-averse regulator may be reluctant to challenge such "gold plating". The ODIs may have a similar effect, but in addition, since these are based on performance commitments made by the company, they could be used as a way of rationalising preferred types of spending. Aside from the prospect of pecuniary gain or loss, companies may be motivated by the desire to avoid reputational risk. That is, given the informational asymmetry between the firm on the one hand, and the customers and regulators on the other, it may make sense from a reputational point of view to over-spending in certain areas rather than risk underperformance against key performance or quality metrics.





All of these are potential explanations, not only for CAPEX bias, but also for bias towards certain categories of CAPEX. We now look further at decomposition of TOTEX into OPEX and CAPEX categories. Figure 15 shows shares of two OPEX categories and five CAPEX categories in TOTEX, and how these have changed over time. We limit our attention here to wastewater services. These more detailed data could only be obtained for the shorter period between 2004-05 and 2010-11, covering PR04 and the first two years of PR09 (and in the latter case, the introduction of the CIS). For the sake of simplicity, OPEX relating

Source: June Returns.

to enhancements – to quality, to customer service, or to the demand/supply balance – are aggregated into a single enhancements OPEX category.

The most notable feature of the above chart is the reduction in the share of CAPEX enhancements – again, we refer to the sum of quality, customer service, and demand/supply enhancements – and particularly of CAPEX quality enhancements. OPEX enhancements are likewise reduced over the period. This paints a more nuanced picture than the less disaggregated trends shown earlier. While the shares of OPEX and CAPEX remain roughly the same, it seems that enhancements spending in both categories had diminished over the period (and of course that base OPEX, non-infrastructure maintenance, and infrastructure renewals take up a larger share). This, if anything, suggests a bias against enhancements in favour of maintenance and renewals. In the context of rail, ORR offers one potential explanation for such a bias, stating that:

"... [A] five-yearly cycle fits better with regular activities such as maintenance and renewals than it does with enhancement projects, as these involve bespoke engineering solutions that may not be ready for assessment or have robust and mature estimates at the time at which the periodic review is determined. In addition, they may take more than five years to develop and deliver, or be subject to significant changes in their timing and/or scope as they go through the development process."²³

The above may dissuade firms subject to short regulatory cycles from spending on enhancements. On the other hand, it may simply be that the need for enhancements spending decreased over the period. Again, this highlights the inherent difficulty in identifying biases simply on the basis of observed trends.

In respect of any bias that may have occurred over the period in respect of the split between OPEX and CAPEX renewals, again Figure 15 does not indicate such a change – indeed the share of both categories within TOTEX increased over the period, with the OPEX share increasing by the most.





Source: June Returns.

Returning to Ofwat's suggestions outlined at the beginning of this section, a second option is to compare performance against targets. For the water and sewage industry, we are able to retrieve data for each of the expenditure categories and compare these to values from the Final Determinations documents published by Ofwat for each price review. Again, these cover the years 2004-05 to 2010-11, and are deflated by RPI. For each category, actual spending are compared to the targets from the final

determinations, and the difference is shown both in absolute terms and as a percentage of the target value.

Figure 16 compares actual and target base OPEX. In Figure 15, we saw an increase in base OPEX as a proportion of TOTEX, and here we can see that, with the exception of 2010-11, WaSCs have tended to underperform somewhat in relation to this category. This underperformance has lessened over the period, from being 14% above the target in 2004-05, to around 4% and finally 3.62% under target in the final year.



Figure 17. Wastewater enhancements OPEX 2004-05 to 2010-11, English and Wales WaSCs

Source: June Returns.

On the other hand, we see from Figure 17 that WaSCs have generally outperformed their enhancements OPEX targets. This outperformance is proportionally greatest in the last three years of the period, with actuals more than 40% under the targets in the final year.



Figure 18. Wastewater enhancements CAPEX 2004-05 to 2010-11, English and Wales WaSCs

Source: June Returns.

A similar story emerges when we look at enhancements CAPEX. Again, each year the industry has outperformed the target values, with the difference in this case being particularly marked early in the

period. It seems clear on the basis of Figures 16-18 that the WaSCs have generally been consistently outperforming their targets for OPEX and CAPEX enhancements spending. Together with the decline in the share of enhancements in TOTEX seen earlier, one possible explanation is that there is a bias against enhancements spending in favour of other types of OPEX and CAPEX. We may therefore expect to see underperformance in other non-enhancement spending categories, as we did in Figure 16 relating to base OPEX.





Source: June Returns.

Figure 19 compares actual and target values of wastewater infrastructure renewals. Again, in keeping with the idea that there may be bias against enhancements spending, we see underperformance in most years, although given notable out-performance in the first two years, there is more ambiguity here than in other cases. Finally, Figure 20 looks at performance in terms of wastewater non-infrastructure maintenance, a component of CAPEX. As with infrastructure renewals, we note underperformance from 2006-07 on, which fits with the idea that there is a bias in favour of non-enhancements spending.



Figure 20. Wastewater non-infrastructure maintenance CAPEX 2004-05 to 2010-11, English and Wales WaSCs

Source: June Returns.

To summarise this section, we analysed trends in TOTEX and its components, and discussed the extent to which there is evidence for CAPEX bias in the UK water and sewage industry. In terms of trends in overall OPEX and CAPEX, there does not seem to be any clear indication of a substantial change in the composition of TOTEX that could indicate a CAPEX bias. Looking at a further split of OPEX and CAPEX into maintenance, renewals, and enhancements components yielded a more interesting and nuanced picture. In particular, the shares of the enhancements components of wastewater OPEX and CAPEX, which includes spending on quality improvements, the demand/supply balance (i.e. reliability), and customer service enhancements, have decreased. We also find that WaSCs have on average been outperforming targets relating to wastewater enhancements while, on balance, underperforming in terms of wastewater maintenance and renewals. Finally, in respect of the split between OPEX and renewals-related CAPEX we see no evidence of a capital bias.

The reduction in enhancements spending could be seen to be surprising at first sight, since as Ofwat point out, below-ground enhancements are not depreciated, but are treated as permanently increasing the asset base. On this basis, we perhaps ought to expect to see, if anything, bias in favour of enhancements. However, Ofwat have also observed general outperformance in terms of enhancements, stating:

"... historically we have observed high outperformance for capital enhancement projects. This effect was perceived to be an issue relating to scoping and information asymmetry. Our response to this was to introduce the CIS at PRO9. We are yet to see the impact from this mechanism. For OPEX, the observed levels of outperformance are significantly different across the companies. Although it is difficult to conclude anything from the data, this observation may be an indicator of wider influences on bias, including company culture."

We also note ORR's suggestion that this may be driven by the long timescales involved with enhancements projects relative to the five-year investment plans that coincide with the five-year price reviews in the industry. We tentatively conclude that this may be an indicator of bias against enhancements spending, but that further investigation is needed. However, it should be noted that the cycle of enhancements will also be determined by a range of other factors, for example the need (or not) to achieve quality-related targets, so we should not read too much into the current data without looking at a longer time period.

Road sector in England

Overall there does not appear to be evidence of capital bias in the water and rail sector. The Strategic Roads Network (SRN) and Local Highway Network (LHN) are considered in turn.

SRN maintenance, renewal and enhancement expenditure

One of the key motivations for reform of the SRN infrastructure manager was to remove volatility in funding from year to year (Cook, 2011). This in turn allows for a more coherent maintenance, renewals and investment strategy. Such arrangements are necessary to address perceived historic underinvestment (enhancements) in roads over a number of decades during which road traffic has increased dramatically (Figure 21). Thus the new organisational arrangements and supporting regulation is motivated by a need to deliver a step change in road investment.



Source: DfT (2015).

The Road Investment Strategy 1 (RIS1) sets out the funding available and key enhancements to be undertaken (and other performance outputs) for the period 2015/16-2019/20. The implied spending on maintenance, renewals and enhancements is shown in Figure 22. Overall the plan for the period is roughly static maintenance and renewals expenditure but a step change in enhancement expenditure.



Figure 22. Planned maintenance, renewals and enhancement expenditure for the SRN in England

Source: Data from DfT (2015).

Thus there does not seem to be evidence of capital bias. Rather, the new arrangements (which require an economic regulatory regime) have been chosen to facilitate a large programme of enhancement. Further the stability of maintenance and renewals expenditure over time does not indicate any bias towards one or the other (unfortunately comparable historic data is very difficult to obtain).

LHN maintenance, renewal and enhancement expenditure

Figure 6 reports past and future planned expenditure allocations to local highways authorities. In the LHN, capital bias is unlikely to be present. Indeed the bigger issue is OPEX bias driven by sufficing under the pressure of short-term budget constraints rather than long-term focused asset management which minimises whole life cost.

	England (ex-London)		London	
	2012	2016	2012	2016
Total shortfall in annual carriageway maintenance budget	£627m	£623m	£89m	£86.7m
Estimated time to clear carriageway maintenance backlog	11 years	14 years	9 years	16 years
Estimated one-time catch- up cost per authority	£73m	£88.8m	£20m	£22.1m
Average number of potholes filled per authority last year	12,392	16,616	3,083	4,099
Total spent filling potholes in past year	£80.6m	£110.6m	£5.6m	£11.4m
Amount paid in road user compensation claims	£16.7m	£8.9m	£3.2m	£4.1m

Figure 23. Backlog in local road renewals

Source: UK Parliament (2016) original source: AIA, Annual Local Authority Road Maintenance Survey 2016, 23 March 2016, p3.

This is perhaps best illustrated by Figure 23 which shows the estimated time to clear the estimated carriageway renewals backlog. Note that the table refers to a maintenance backlog. However "maintenance" is a term used for both maintenance and renewals in this sector. Overall this highlights a major problem with backlogs in asset renewals which results in both poorer condition of the road network for users and a sub optimal move towards reactive maintenance which is more costly than an appropriate mix of reactive and more preventative maintenance/renewal activity. Such under-investment is typically of state-owned operators.

Summary

Following recent reforms to both the Strategic Road Network infrastructure manger and the funding for local authorities with respect to local roads, roads in England are subject to economic regulation in some form. Given the relatively new nature of reforms there is little evidence of the impact of reforms (and therefore the impact of regulation). However the motivations for the reforms, at least for the strategic road network are to facilitate more capital spending in the sector. As such any future evaluation needs to be careful not to confuse correlation of higher CAPEX and reform with evidence of capital bias. If Highways England delivered these substantial increases in enhancements then that would be seen as successful. ORR is in the process of developing a benchmarking framework which may include a TOTEX-type approach.

In local roads the main problem is lack of funding resulting in less-than-optimal asset management. This could be viewed as OPEX bias, but again that arises from the constrained funding regime. Local authority roads are not regulated in the traditional way; however, as will be highlighted in the following section, a TOTEX benchmarking framework has been developed.

The effectiveness of efficiency benchmarking in transport

Having considered the theoretical and regulatory literature with respect to capital bias and examined the available (perhaps limited) evidence for its existence, the purpose of this section is to address the question of whether efficiency benchmarking can be a powerful tool in driving efficiency improvements in transport, viewed through the lens of the three cases of rail, roads and water.

These cases offer highly relevant case studies through which to study the issues under consideration in this paper and the project team have been closely involved in working with regulators/policy makers/providers to develop the regulatory approaches developed in each case.

This section of the report considers two key issues: the purpose and broad approach of top-down econometric benchmarking; and the specific issues relating to the comparison of CAPEX expenditure between organisations.

It should be noted that there are wider issues impacting on the effectiveness of top-down benchmarking in general that are not covered in this report. Broadly there are accepted means to benchmark operations expenditures and these have been used by numerous regulators. CAPEX benchmarking is very relevant to the remit of this working group as the group is focused on private sector financing of investment in infrastructure. There have been instances of benchmarking of such expenditures in the regulatory model (as we will show), but there are some unique challenges associated with it.

Hence this report focuses on the question of how to benchmark CAPEX, the most problematic aspect of regulatory benchmarking. The chapter is structured as follows. We first motivate the purpose of top-down econometric benchmarking in the economic regulatory framework and explain the overarching efficiency benchmarking approach at a conceptual level. We then outline and summarise the

key challenges associated with CAPEX benchmarking before dealing with each of these in the subsequent sections: data definitions, lumpy CAPEX, and intermediate versus final outputs. Following this, we draw particular attention to the similarities and differences between renewals and enhancements CAPEX. We then discuss how regulators can use top-down econometric benchmarking to set efficiency targets, including how such approaches can be combined with other methods to arrive at a final determination that is realistic and achievable. This discussion naturally leads to one about the appropriate role of an economic regulator.

The purpose of top-down econometric benchmarking

The material in this section draws partly on Smith and Nash (2014).

Studies of transport efficiency usually have one of two motivations. Firstly, they may aim to identify which transport operators and/or infrastructure managers are efficient and which are not, in order to draw lessons as to the level of improvement that may be required. An example of the benchmarking framework developed on behalf of the British rail regulator in deciding on the financial requirements of Network Rail, the infrastructure manager, is discussed below (e.g. Smith et al., 2010). Secondly, studies may seek to draw policy conclusions about which policies regarding industry structure, competition and regulation will be most beneficial.

In both cases the importance and motivation of efficiency analysis, and the reforms and policy interventions which may occur based on such analysis, is the delivery of efficiency savings (and productivity gains more generally) with the ultimate aim of delivering either or both lower prices to users of transport services, or reduced taxpayer support.

In sectors of the economy in which markets are a reasonable approximation to perfect competition, a reasonable measure of overall efficiency may simply be the profitability of the firm. Under perfect competition, prices are not influenced by the individual firm and therefore the more profitable the firm, the more it has been able to minimise costs of production and to produce the most valuable combination of goods in the eyes of consumers.

However, transport provision is in most cases a long way from being a perfectly competitive industry. In some sectors of rail operations (e.g. coal, commuters) for example, operators still have considerable monopoly power, whilst for this and other social reasons rail prices are often regulated by governments, who also play a key role in specifying passenger sector outputs. These factors must be allowed for if the aim is to examine the efficiency of railway management. To the extent that railway managers (at least in the European passenger sector) have limited control over their outputs, the key issue is whether they produce them at minimum cost.

Further, in the absence of competitive pressure, regulatory pressure is needed to ensure that transport firms operate in an efficient manner. Given the standard problem of asymmetric information – firms know more about their costs than regulators – the latter need to arm themselves with additional information to overcome this asymmetry. Benchmarking – which involves comparing the efficiency performance of the regulated firm with other transport firms – is thus an important, indeed crucial part of the regulatory process. The data that is used in benchmarking could be drawn from other domestic transport firms, where possible, international benchmarks (which can be more problematic), or internal benchmarks (for example comparing the performance of different regions of an infrastructure manager). Data at the level of individual projects could also be appropriate as discussed further below.

Economists are used to distinguishing between technical efficiency and allocative efficiency. Technical efficiency is measured by whether output is maximised for a given level of inputs (or conversely inputs are minimised for a given output). The standard economic approach to examine this is to estimate a production function using econometric methods, although non-parametric methods – such as data envelopment analysis (DEA) – have also been used. Allocative efficiency considers whether the correct mix of inputs is used to minimise cost for a given level and quality of output. Cost efficiency, which is the product of technical and allocative efficiency, thus takes both technical and allocative efficiency into account. Cost efficiency measures are typically obtained by estimating a cost function or frontier; though such measures can also be obtained via DEA. Arguably cost efficiency is the most relevant concept from a regulatory perspective since it is through cost reductions that prices can be reduced to users, or the burden of subsidies lowered. That said, in situations where it may be unrealistic for firms to optimise the capital input – as may be the case in regulated network industries – and/or where there is a lack of good and comparable data on costs or input prices, technical efficiency could be a more relevant measure.

Before contemplating advanced techniques for measuring the relative efficiency of transport firms, it is first worth asking why it is important to go beyond simple partial productivity measures - such as (in a rail context) cost per train-km, cost per passenger-km or cost per route or track-km. That is, in principle, benchmarking could simply proceed by collecting comparable data across firms and using these to compare unit costs across companies – thus raising the question as to whether advanced techniques are needed at all.





However, unit cost measures are only partial measures of efficiency performance and raise multiple questions. Firstly, such measures may not cover all costs and thus would not give an overall assessment of performance. For example, staff costs per train-km could be distorted in comparison between firms if different firms take different out-sourcing decisions. Further, capital substitution possibilities may be ignored. Even if the measure of costs cover all costs there is then a of how to characterise the outputs in industries where output may be described in terms of, for example, passenger-km, train-km or even

represented by the size of the network, making cost per track-km another candidate measure. Further disaggregation is also possible as between freight and passenger traffic, or even further traffic-type disaggregation within the generic categories.

Typically, unit cost comparisons with different denominators can give very different rankings. Similar problems emerge in other regulated industries, such as the water sector, where unit costs may be calculated per customer or per unit of water delivered or per length of mains, with potentially very different results. It is thus problematic, since it is not always clear which is the preferred measure, and different rankings and efficiency scores result from using different measures.

An approach that simultaneously takes account of multiple denominators/drivers of costs is therefore needed - this being a key advantage of the econometric approach, as multiple cost drivers can be included in the model at the same time. DEA can also handle multiple inputs and outputs, sometimes via a two-stage process. A second key advantage of adopting an econometric framework (and to some extent also a DEA approach), as compared to simple unit cost measures, can be illustrated in Figure 24.

Simple unit cost comparisons make an implicit assumption of constant returns to scale, which is very unlikely to hold in transport applications. Thus in the left hand panel of Figure 24, firm A appears to be inefficient because it has higher unit costs. However, once the presence of increasing returns to scale is permitted – as represented by the right hand panel of Figure 24 – it becomes apparent that the reasons for firm A's higher unit costs is that it operates at small scale. Given its scale – which will usually not be under the control of management – firm A turns out to be operating on the cost frontier and is thus an efficient operator. By using an econometric approach, the shape of the cost frontier can be estimated as part of the modelling framework. Indeed the assumption of constant returns to scale can be tested directly.





Figure 24 has shown cases where firms have different unit costs but are both efficient in the sense that they cannot reduce cost without sacrificing output (or quality). Figure 25 shows the case where some firms are efficient (C, D and E) as they are operating "on" the minimum cost frontier, whilst A and B are inefficient as they are operating above the minimum cost frontier. A and B can reduce costs and still maintain output by adopting better operational practices. More formally we can define the ratio of

minimum to actual cost as "cost efficiency", sometimes referred to as an efficiency score. A value of 100% implies the firm's minimum cost equals its actual cost and so it is efficient. A value less than 100% implies scope for cost savings as the firm is inefficient.

By using an econometric framework, the shape of the cost frontier can be estimated as part of the efficiency modelling approach and the assumption of constant returns to scale tested as noted. Likewise DEA methods can be adapted to take account of the possibility of non-constant returns. Having controlled for the underlying technology, efficiency comparisons can then be more accurately assessed.

It is worth noting that there are many other important factors that need to be taken into account in benchmarking and which more advanced methods – that is, going beyond simple unit cost measures of key performance indicators – can deal with at least to some extent. These include the need to take account of:

- variations in quality between firms and over time
- observed (or unobserved) heterogeneity between firms that impacts on costs, but is not related to management performance these are likely to be persistent factors that do not vary over time, and may relate to factors such as the weather conditions or topography of the country or region being benchmarked
- variations in the price of inputs between firms and over time
- dealing with the particular challenges of CAPEX benchmarking.

In the subsequent sections we focus on the latter of these issues.

It is also important to note that within the standard model of RPI-X regulation, the role of an economic regulator is to use benchmarking as part of a range of methods to set efficiency targets for firms that are credible and achievable, and provide appropriate incentives for firms to deliver those targets and even outperform. Regulators need to guard against the temptation to micro-manage the activities of regulated firms. We return to this issue in this section "Combining approaches and the role of regulator", where the particular challenges of dealing with state-owned firms within the regulatory framework are discussed.

Key challenges associated with CAPEX benchmarking

As seen above, it is important to accurately draw the cost frontier for each company. We have to control for the key factors influencing its costs; namely the outputs it needs to produce and the price it faces for inputs, but also the condition of the infrastructure that it manages.

For capital costs, there are several additional complications over more conventional operations cost benchmarking:

- Data between firms needs to be comparable across a broad set of cost categories. This presents challenges as how best to collect and analyse data and provides a motivation for "TOTEX" benchmarking.
- Capital projects are "lumpy" and intertemporal in nature. This means that it can be misleading to look at a single year of data across firms without making some allowances for the current investment position of each in their investment cycles.

- Related to the previous point, a question arises as to whether it is appropriate to benchmark in terms of the relationship between costs and final outputs (e.g. train-km in railways) or intermediate outputs (e.g. km of rail renewed).
- Finally there is a question as to whether enhancements can be benchmarked in the same way as renewals costs and operations costs. In relation to the previous point, how far can in-year enhancement spend be related to current year outputs (as opposed to future outputs)? Further, to what extent does the firm decide which enhancements to proceed with, versus the firm delivering a set of enhancements as specified by government?

Each of these issues is discussed in turn in sections below.

Consistency of data definitions

There are two key aspects of defining and collecting cost data across a number of organisations:

- What "direct costs" are in scope? Essentially this covers which activities are being considered, such as maintenance versus renewal activities, or more nuanced definitions of track renewal versus rail or ballast renewal in railways.
- To what extent is an element of overhead included in costs?

Any benchmarking requires that costs cover the same activities. Costs can be split between activities e.g. in highways: carriageway work, embankments, street lighting, and between different expenditure types such as operations, maintenance, renewals and enhancements.

A more narrow definition of cost allows more appropriate cost drivers to be identified and included within the benchmarking models. For example, the cost drivers of carriageway maintenance may be very different to the costs drivers for structures and embankments. However, a narrower definition has two key drawbacks. Firstly, there may be allocation differences between road authorities of exactly which activities fall into each category. An example from the CQC Efficiency Network is that some organisations find separation of footways from carriageway spending problematic. Secondly, there may be some possibilities to substitute one form of activity for another, such that there are joint costs in provision and thus elements of provision should be considered together.

The same arguments apply to maintenance, renewals and enhancement expenditure categories. In rail, there is evidence that different countries define the boundaries between maintenance and renewal slightly differently, though the differences may be relatively small (Smith, 2012). Network Rail has also argued in its discussions with ORR about international benchmarking that there may be differences in definitions with regard to the boundary between renewals and enhancements. The CQC Efficiency Network has spent considerable resource trying to better define OPEX spend versus CAPEX spend. This has emerged because of specific accountancy reporting rules impacting on local government can often mean that some activities are recorded under capital spend when they are clearly maintenance (OPEX) in nature.²⁴ As such the CQC Efficiency Network defines cost of OPEX and CAPEX by itemising activities that fall within each.

Thus there can be trade-offs between benchmarking categories together (avoiding cost allocation issues) versus separately (achieving a more tailored model specification for each category). Naturally, therefore, it is sensible to carry out both approaches and compare (as has been done, for example in the modelling work to date as part of the 2019 Periodic Review process in the water sector). In rail, models covering maintenance (and renewals) only have been compared to models that include both maintenance and

renewals together. That is, undertaking modelling of separate cost categories (by activity) if data is available is a sensible approach to complement aggregate benchmarking. The choice of the approach, and the weight given to these different modelling approaches, can be informed by qualitative feedback from participants²⁵/firms on differences in the definitions of activities within cost elements and is an area where a regulator may wish to exercise regulatory judgement. We discuss how project level data could also be used in the section "Intermediate versus final outputs".

Turning to the issue of overheads, a key element with any cost data is the apportionment of overheads to direct costs. When benchmarking across organisations such apportionment needs to be understood and achieved on a consistent basis. Ideally it would not be allocated at all (as this is likely to introduce some arbitrary allocation of joint costs). However when organisations have different structures, carry out a range of activities that may be benchmarked, and also have different contracting out policies, it is often not possible to completely eliminate overheads.

For example, area-wide contracting such as that used by Highways England for maintenance - prior to the formation of in house determination of maintenance activity (a recent development) - meant that the contractor price included a mark-up for management overheads (within the contractor) in the charge to Highways England. The contractor price is very different for the revised model implemented by Highways England, namely in house maintenance with the use of contractors for specific jobs. To be comparable the latter arrangement would require both contractor costs and an allocation of Highways England's management costs. This adds to the data collection challenge.

As such it may be sensible to agree a pragmatic definition of costs which may include some overheads. This was the approach in the CQC Efficiency Network where there are indirect costs collected as well as direct costs. Further the CQC Efficiency Network collects qualitative information in cases where the definition of cost provided by a participant does not fully align to the agreed definition, so that the likely impact of this can be assessed at least qualitatively. Of course, in a more formal regulatory situation, this situation might not be deemed acceptable and the regulator may impose data definitional requirements where possible and to the extent that they are achievable at sensible cost to the firms.

Overall regulators need to try to ensure comparability of data, whilst recognising the regulatory burden that this may impose and also make choices about the appropriate levels of disaggregation, taking into account amongst other things the resulting definitional challenges that result from disaggregated data. Where definitions cannot be fully aligned, collecting qualitative/directional evidence on the disparity between reported costs and the definition for use in understanding the resulting efficiency scores can be useful.

Perhaps one of the lessons that may be learned here is that top-down econometric benchmarking using international benchmarking is particularly challenging because it takes considerable time and commitment from a group of countries (or companies) to make the analysis credible and usable. In the 2008 Periodic Review, PR08 ORR had the advantage of a ready-to-go dataset, produced by the International Union of Railways (UIC), and this enabled top-down, econometric international benchmarking to play a more significant role than it has in other regulated sectors in Britain (though even there the data quality has been questioned). If international benchmarking is to work, then it may require concerted efforts by regulators/governments (and/or regulated firms) across Europe working together to establish a common benchmarking framework against which all companies can be compared, thus also implying that data can be requested and audited by regulators and policy makers.

Capital projects are lumpy and intertemporal

Capital expenditure changes over time. This is due to two factors: short-term budget constraints and longer-term asset management. The latter arises because it may be optimal from a life cycle cost perspective to renew a differing volume of assets in any given year due to historical usage and spending patterns. This presents substantial challenges for benchmarking renewal expenditure between organisations, since some organisations may spend relatively large or small amounts because they are simply at a given part of their asset life cycle.

Short-term fluctuations in expenditure due to temporary funding constraints can potentially be overcome by examining performance over a number of years (or if they persist for a number of years, using asset condition measures as discussed below). As noted in earlier sections of this report, independent economic regulation is intended to provide a framework in which firms are given sufficient funding to carry out their duties, and should ensure efficient, long-term planning. However, in some contexts, particularly where regulated firms are state-owned, the problem of OPEX bias can occur.

Dealing with longer term asset life issues is a vexing question in benchmarking. For example, asset lives of highway and railway infrastructure are very long and so it is unrealistic to ever collect data over the whole life cycle of such assets. Instead, several approaches have been suggested:

- Method A: Dispense with a cash measure of renewals and adopt accounting depreciation as the measure of cost this approach necessarily smooths the pattern of "renewals". However it is challenging to define appropriate depreciation measures across organisations in a meaningful way.
- Method B: Adjust costs directly to reflect a long-term average rate of renewals activity.
- Method C: Take a moving average of renewals costs as done by Ofwat.
- Method D: Control for the condition and life expiry of assets directly in the model.

Method B was used by ORR in PR2008 and PR2013 for analysis of Network Rail's renewals costs against other international railways. This was termed a "steady-state adjustment"; although given lumpy renewals volume levels can be optimal from a life cycle cost perspective, given historical patterns of renewal, the term "steady-state" renewal is open to interpretation. ORR adjusted track renewal costs for each railway up (or down) depending on whether track renewals volumes were below (or above) a specified rate (renewal km per network km per annum) e.g. 2.5% of the network renewed per annum (this percentage being determined based on engineering judgement/evidence).

This approach has much to commend it and has a lot in common with the concept of depreciation. However there are a number of challenges:

- As well as data on costs, data on renewals volumes has to be collected. This is not trivial as there is a question as to how to most appropriately measure renewals volume. ORR used rail renewal but there are more track assets in addition to the rail.
- What is the average rate of renewals? Should this be the same across organisations? Indeed, in PR2008, ORR could only apply the steady-state adjustment to Network Rail as there was little evidence to support the same level for other railways.
- There is the question of sub-optimal CAPEX/OPEX ratios. A railway could be "over-renewing" i.e. not cost minimising by choosing renewal rather than maintenance.

• Finally, as a more micro-detailed point, the steady-state adjustment carried out by ORR implicitly assumed constant returns to scale in renewal volume activity. That is, if renewal volumes were double the estimated steady-state level, renewal costs would be reduced by 50% prior to estimation (in practice there could be scale economies or diseconomies in delivering renewal work).

Figure 26 illustrates the potential power of Method B, as used by ORR in 2008 (and in 2013). In the early years of the sample it was deemed (at least after the event) that Railtrack was not spending enough on renewals. In the unadjusted model, using a measure of cost that included maintenance plus CAPEX renewals, this therefore makes Railtrack look efficient (with an efficiency score of unity) during the early years. When renewal volumes later increase sharply, efficiency performance appears to deteriorate very sharply. The adjusted efficiency scores show a less stark picture, and have the effect of moving Railtrack off the efficiency frontier. The deterioration in efficiency is also reduced thereafter, though the results do indicate that even after adjusting renewals to their steady-state level, Network Rail's costs were still much higher than the frontier firms by the end of the sample. The maintenance-only model is a further useful comparator because it does not include lumpy CAPEX; however, there may be issues regarding the classification between maintenance and renewals.





Source: Smith (2012).

Method C is similar to method B in that it is an attempt to smooth out fluctuations in renewals spending and implicitly is making an adjustment to the raw renewals cost data. An issue arises as to what should be the length of the moving average (e.g. two, three, four or five years). In addition there is an issue as to how to treat the first and last years in the sample as renewals cost may not be observed for years before the first year of analysis (to form the moving average based on preceding years data) and also not available after the last year of analysis (to use proceeding years of data in the moving average). This means, particularly if only a small number of time periods observed, the moving average may still be very volatile as the first and last years are computed off one or two data points; or that data has to be dropped in order to retain say a five-year moving average. A further issue in economic regulation is that historical data may be deemed to be of less relevance than more recent data, such that using averages based on data that is three years prior to the current five-year control period – to make assessments about expenditure in the next five-year control period – may raise issues.

Method D has been developed in most econometric benchmarking studies, in the sense that most include at least one measure of the capability or condition of the capital stock. However the direction of the statistical relationship between this measure and cost is often ambiguous. Taking the case where the dependent variable is a summation of maintenance plus a measure of CAPEX (perhaps smoothed to an extent), on the one hand, better capability/condition implies lower ongoing costs as there is a greater past investment in capital. On the other, it costs money to get to such a high level of capability/condition indicating the reverse relationship if an operator is investing to improve asset condition/capability. At the heart of the problem here is the fact that time lags exist between CAPEX and this being reflected in improved condition (and in turn lower maintenance). Thus there can be times when it appears that high expenditure is correlated with poor asset condition; whereas the "longer" term relationship should be that higher condition requires higher overall cost in the long term (assuming firms are efficient).

In roads, the CQC Efficiency Network has developed this further, as this issue is very relevant to local roads which are in relatively poor condition in Great Britain. In that analysis, the asset condition measure (a defect measure) is included in two ways:

- In its level, i.e. the value of asset condition in a given year that has a negative relationship with cost as it captures the impact on cost of having a good or bad network. So the worse the network, the greater cost required to maintain and renew it.
- The change from the previous year we would expect this to have a positive relationship with cost. If road condition improves from one year to the next then that indicates that more cost was incurred to improve the infrastructure's condition.

Figure 27. The trade-off between upfront cost of asset improvement and future present value saving



This approach is innovative because it no longer penalises authorities for investing in their network and recognises that there is up front cost for longer term cost savings, which is what our cost model indicates when we estimate it using the data. Of course this relies on having a measure of road condition which is responsive to investment. Figure 27 illustrates the trade-off between the marginal upfront cost of improving road condition and the marginal saving in future years (measured in net present value terms). For poor road condition the up-front marginal cost of improving road condition is lower than the present value saving which flows from cost savings in the future. This can be expected to diminish over time such that there is an optimum road condition (or more broadly asset condition) at Q* where improving road condition further incurs a greater marginal up-front cost (MUFC) in relation to marginal future saving in

present value terms (MFPVS). Essentially this is the standard trade-off in life cycle cost analysis (between capital cost and ongoing maintenance), but what is innovative here is that the CQC work has been able to embed such trade-offs in the benchmarking model.

Importantly, it should be noted that the nature of statistical approaches (to be discussed in the addendum referred at the beginning of this section) is that they can, to an extent, deal with "noise" in the data, particularly where panel data is available. Further, the way in which frontier-based efficiency analysis methods work means that the frontier may be defined by multiple firms, thus reducing the possibility that the observed frontier results from some freak event in which costs are very low in one year for one firm. Thus, to an extent, frontier-based econometric approaches to benchmarking CAPEX can mitigate some of the issues noted above; though these should be combined with some of the other techniques set out above.

Overall, we consider that statistical benchmarking is feasible for renewals costs and has been used within the CQC network, by ORR and other economic regulators (e.g. OFWAT) as noted above.

Finally, one option that has not explicitly been discussed, is to include the volume of renewal activity as an explanatory variable in the regression model. This approach implies a very different interpretation of the cost function, and leads naturally to a discussion of intermediate versus final outputs in cost models in economic regulation; to which we now turn.

Intermediate versus final outputs

A key problem that has been identified in the academic and regulatory literature concerns the benchmarking of CAPEX costs in economic regulation. As noted above, CAPEX can be lumpy and intertemporal such that estimating an econometric model that regresses CAPEX on measures of final outputs (such as service provided, as measured by train-km) is likely to be problematic. An alternative approach is to regress CAPEX on measures of intermediate outputs, for example the volume of rail renewed during the year. This approach seems more direct, and relates the amount spent in that year to what is delivered (in terms of renewal work completed) during the year.

One disadvantage of this approach is that it takes as given the volume of renewals carried out, which could potentially be too high (or too low); it only considers whether the renewal activity that is being done is being done in an efficient manner (i.e. at an efficient unit cost per intermediate output). This approach could be combined with other approaches aimed at establishing whether the renewal volume is appropriate (see below).

This approach may be applied to renewals or enhancements in principle. The approach (referred to as the cost base approach) has been used in the water sector for both enhancements and renewals (termed capital maintenance in the water sector); see for example Ofwat $(2004)^{26}$. The approach works by utilising comparative data across multiple (repeated) projects within and across companies. Standard unit costs per project are computed for each company from the database of projects implemented and compared with those of other companies in order to identify a benchmark against which to compare all companies and set efficiency targets. In the case of water, this approach has been applied to both renewals (referred to as capital maintenance) and enhancements (see Ofwat, 2004; 1999). The approach does not use econometric techniques in this case and works with project-level data as noted.²⁷

The approach has also been implemented in the rail sector where in this case multiple renewal projects within Network Rail were compared against each other. In the latter case an econometric model was

used to control for other factors (such as whether the work was carried out at night or not and whether the line was electrified)(LEK, 2003).

The approach does not have to use econometric methods; once an appropriate volume measure is identified (such as km of track renewed in railways) then unit cost measures can be computed and compared. However, there are significant advantages in using an econometric method even if only a single variable (track renewal volumes) is included in the model as the model can adjust for possible economies/or diseconomies of scale and also filter out noise from the data. Other factors can also be taken into account, as in the rail case noted, to reflect differences in the conditions under which the activity is taking place.

In addition to benchmarking the unit cost of renewal or enhancement activities, regulators also commission other studies to challenge company's approaches to running their operations and how they undertake project management, and generally scrutinise the business plans submitted by companies as part of the regulatory process. Thus whilst the unit cost approach only looks at whether CAPEX is being implemented efficiently, and not at whether the right CAPEX is being done, other approaches do exist for regulators to assess the latter (though to an extent based on judgement, rather than hard data or econometric evidence).

The treatment of enhancements

In regulation, CAPEX comprises two elements: renewals cost which is associated with activities that replace an asset with an asset of the same capability to service users; and enhancements, that deliver new capability. An example of renewals is a substantial resurfacing of a road's carriageway, whilst an example of an enhancement is a construction of a new road or an upgrade to a railway that improves capacity or line speed.

The nature of the data used in regulatory benchmarking is primarily organisation-wide data (or, in theory, regions within an organisation). For example, the cost of Highways England may be compared against another strategic road operator in Europe. Alternatively, ORR has compared Network Rail against other rail infrastructure managers in Europe (Smith, 2012) and at the same time has benchmarked Network Rail's regions or routes against each other. ORR has also used mixed approaches, in which regional data from a number of countries are combined (Smith and Wheat, 2012). Such data is definitely useful for considering OPEX cost (maintenance), as costs can be related to factors such as (in the case of roads) network length, traffic and road condition. Potentially, as long as suitable care is taken, renewals CAPEX can also be benchmarked with such data in the same way, as it can be described at such a strategic level by such high-level factors. Effectively this is because maintenance and renewals cost are concerned with good asset management which should be a function of the size, quality, usage and the historical condition of the assets.

However, the efficiency of delivering enhancements in transport is very different. The aim is to evaluate whether a set of externally determined projects are delivered at minimum possible cost, given the characteristics of the projects. In transport contexts, these projects are determined through undertaking cost benefit analysis of potential projects and then prioritising projects to fit the government's funding envelope. Therefore, it is the cost of delivering projects that is important for benchmarking, and not the choice of projects being taken forward. It is also the case that enhancement project costs in one year are not likely to be related to outputs in that year, as enhancements typically have impacts in terms of future outputs and indeed enhancement projects can take more than one year to deliver.

Thus it is likely to be hard to relate enhancements CAPEX in an econometric model to high-level metrics such as network length and final outputs such as traffic volume. This fact is reflected in actual practice. ORR's econometric benchmarking framework has focused on maintenance and renewals (with enhancements dealt with outside the econometric framework). Likewise in the water sector, at the 2014 review TOTEX econometric models were used (including enhancements) to relate costs to final outputs, but the analysis mainly focused on what are referred to as BOTEX models, which are essentially comprised of maintenance plus renewals (with enhancements dealt with through a separate approach). Whilst there were thought to be advantages of using a TOTEX approach, in that all activities could be benchmarked together, there were questions of credibility and in developing a robust econometric model that could deal with a cost base that included enhancements). More recent policy is also moving in favour of separating out aspects of enhancements from the models.

Thus another approach is required to relate the costs of enhancements to the intermediate outputs that they create, rather than to final outputs as in the standard econometric framework; and to do this, project level data of some sort may be required. As noted above, in the water sector, project-level data has been used to compare standard unit costs of renewal (capital maintenance) and enhancement projects across companies (with the data drawn from company project databases); the so called "cost base" approach. This compares the cost of delivering projects across companies, rather than trying to relate enhancement costs to final outputs (and in the case of water, did not use econometric methods).

Further, in 2014, Ofwat developed a set of models that are referred to as unit cost models.²⁸ This is an econometric approach, using company-level data to relate costs to intermediate outputs (and thus is based on relatively small samples). One example of the latter approach includes a model for enhancement relating to improving quality (lead standards), where enhancement costs were regressed on the volume of lead pipes replaced (an intermediate output):

Enhancement Costs = Function of (Volume of Lead Pipes Replaced)

Through this approach companies can be benchmarked against each other, not in terms of the scale of the enhancements programme or the choice of projects, but in terms of the unit cost of their delivery. Ofwat separately challenges companies on their business plans and also sets appropriate incentives, as discussed above.

As noted, in rail, ORR has taken a similar approach, with enhancements being treated separately from maintenance and renewals. In rail, enhancements might be considered potentially more challenging and bespoke in nature than at least some of the water enhancement programmes. The approach taken at the 2013 Periodic Review was to start with Network Rail's own cost estimates, which in turn had been developed, and internal efficiency targets set, based on Network Rail's internal project data. ORR then scrutinised the plans and applied some non-rail benchmarks (for example, based on project management) and also removed some risk adjustments that had been included within Network Rail's projections.

Since then there has been a major problem in respect of cost escalation within Network Rail, with the company not meeting the efficiency targets set in 2013; the problems particularly emerging in respect to project delivery on enhancements (especially electrification) and also on renewals where unit costs have risen sharply. Partly for this reason, following the Bowe Review (Bowe, 2015), it was decided not only to treat enhancements separately within the benchmarking framework, but to remove enhancements entirely from the regulatory framework. Many of the reasons outlined for the cost escalation and overruns on Network Rail's enhancements programme will have resonance with respect to the wider discussions of the Working Group (for example):

- poor planning processes given the scale of the enhancements programme
- poor scope definition and scope creep issues
- the fact that it was hard to accurately assess the cost of some aspects of the programme at the time of the regulatory review, because of the immaturity of the projects in terms of their stage of development.

It was therefore considered that it would be advantageous to take enhancements out of the regulatory sphere, with DfT taking charge of this aspect of Network Rail's costs and activity. This change raises a number of issues. At one level it does seem sensible to take enhancements outside of the five-year regulatory cycle to permit greater flexibility – thus allowing the right amount of scrutiny/approval to take place at the right time, depending on the degree or surety about the costs. That said, even under the previous approach, there were cost adjustment mechanisms to allow for changes to project costing during the life cycle of their development.

However, it also raises the possibility of even greater fragmentation, as the amount of enhancement work will impact on the running of the railway generally and there are links between accelerated renewals and enhancements and related definitional issues to contend with. These can no longer easily be internalised because ORR and DfT would be working separately. The proposals would also seem to mean that there will be a lack of independent scrutiny by ORR which is perhaps the greatest concern. This could lead to a situation where Network Rail is asked to deliver an output without adequate funding; and/or putting additional pressure on the operations of the existing railway without adequate protection.

Details on the new process for regulating enhancements appear to be light at present, but pre-requisites for their likely success might include: a clear process to link the approval process and scrutiny of enhancements with ORR's regulatory processes; and evidence that the new process learns from the mistakes of the previous cost over-runs and also from good practice enhancement projects. It would seem useful to explore how regulation of enhancements works in other contexts (e.g. in Germany, where there is also a parent company that supports greater co-ordination between infrastructure and operations and where the regulatory body plays much less of a role). It is also important that a benchmarking framework is established whereby the cost of different enhancement schemes can be compared (within Great Britain and internationally). Though this will be challenging, it will enable the success of the reforms to be measured in terms of final outturn costs compared to other projects as well delivery and cost variability between estimates and final outturns.

Overall, we may conclude that enhancements are challenging to deal with in the standard econometric framework, as compared to maintenance and renewals costs. Thus separate approaches are likely to be needed, though where project data is available, this does not rule out econometric methods. The key difference is that such approaches will inevitably focus on understanding the efficiency of the delivery of enhancements projects, rather than asking whether the enhancements should be implemented in the first place.

We may also conclude that enhancements in rail are probably harder to deal with than in other sectors such as water. They may be more bespoke and complex – in the case of electrification in Britain, such upgrades had not been attempted for many years (though it might be expected that international expertise could have been bought in). There is a further problem of having only one rail infrastructure company (as opposed to around twenty water companies), thus limiting the ability to compare across companies (as occurs in water). Further, the recent evidence in Great Britain is that there have been substantial cost over-runs and problems within Network Rail in respect of project delivery and costs for

both renewals (see Figure 9) and enhancements, such that a major change to enhancements regulation has been proposed, which involves removing enhancements entirely from the regulatory framework.

Finally, it should be noted that it is surely harder to incentivise a nationalised company as compared to a private company in respect of keeping control of project-related costs; though reputational incentives and management performance incentive schemes can play a role. There is also an added complexity in that government (Department for Transport) has a key role in determining what enhancements should be implemented, as well as acting as owner of Network Rail.

Combining approaches and the role of the regulator

It was noted above that within the standard model of RPI-X regulation, the role of an economic regulator is to use benchmarking as part of a range of methods to set efficiency targets for firms that are credible and achievable. In addition, regulators ensure that there are appropriate incentives for firms to deliver those targets and the aim is even for firms to outperform. Importantly, there can be a danger that regulators begin to micro-manage the activities of regulated firms, also potentially duplicating the role of management.

That said, it is also important that regulators set achievable and credible efficiency targets. In part, some of the asymmetric information problems can be resolved through appropriate incentive mechanisms, including the use of menus where possible. Nevertheless, the cost benchmarking information that underpins such mechanisms needs to be credible. Here we briefly address two issues. First, at a high level, the question of how regulators choose the econometric model out of the range of methods tried during the modelling process. Secondly, how such studies may be combined with other evidence.

A feature of efficiency benchmarking in the UK regulatory system is that regulators, through their consultants or in-house teams, test many different model forms before arriving at a final determination. These may have, inter alia, different functional forms (reflecting different assumptions about the nature of the technology and the extent of returns to scale or density), different variables included in the model, and different estimation methods. Regulators then go through a process of evaluating each model according to whether the underlying assumptions of the model are reasonable, how the model performs with respect to standard diagnostic tests (e.g. how well does the model fit the data), and also whether the signs and magnitudes of the coefficients on the variables in the model are in line with prior expectations based on economic theory, engineering and operational considerations.

Indeed this latter point is a key benefit of the econometric approach over other methods. Before studying the efficiency results coming out of the model, it is possible to observe whether the model is showing a sensible relationship between costs and the explanatory variables in the model. For example, at a basic level, are costs shown to be rising in line with outputs and is the size of that effect reasonable. Given that these relationships are transparent to users, they can give reassurance that the model is sensible – and hence that the efficiency scores that result are sensible.

Following this process there will usually remain more than one preferred model. At this point regulators tend to use a process referred to as model averaging or triangulation, in which the efficiency results from the different models are combined, either based on a simple average, or possibly a weighted average if the regulator has more confidence in some models than others (such a weighting can be hard to justify, though a straight average could also be seen to be arbitrary). One of the motivations for this approach is that, having narrowed the choice of models down to a small range of models, regulators often find themselves in a position where there are no clear reasons for choosing one model over the other, and

averaging lowers the risk of getting it "wrong" for individual companies. The following quote from a report by the UK Competition and Markets Authority [CMA] regards Ofwat's approach to benchmarking the water companies in England and Wales: "Ofwat told us that it recognised that there may be different plausible ways and models to use to arrive at an expenditure forecast and that, by using a suite of models, it had mitigated the risk of choosing any single model which, for any given company, may have a large variance between the estimate and the 'correct' answer" (CMA, 2015).

There are other precedents, for example, in the Office of Rail Regulation (ORR) 2008 Periodic review of Network Rail's finances, the efficiency results of a single model were selected (Smith, 2008). However, in this case it was considered that the chosen model was the most appropriate for the dataset – i.e. there was a good reason to choose that model over the others (and the other models produced similar results as well). Later, in 2013, several competing econometric models were developed and this evidence was used by ORR to produce a range of results (rather than expressing a clear preference for one model); see (ORR, 2013).²⁹ It should be noted that in 2013, there was less clear evidence for the dominance of one model over the other.

Thus in the UK regulatory model there is typically a rigorous approach to the development and selection of models, and uncertainty over which model is best is dealt with partly by using the model averaging approach. In addition, where regulators consider that the model does not fit one company very well, special factor adjustments may be used. As noted above, incentive mechanisms can give further flexibility for firms and regulators also typically aim away from the frontier, instead requiring firms to achieve the performance of the upper quartile; firms may also be given time to catch-up to the benchmark (which itself may be assumed to be improving in line with an assumption about technical change).

In addition to the above mechanisms, regulators may look to other evidence. We have previously referred to econometric exercises as top-down approaches; and regulators also use what might be regarded as bottom-up approaches. This term could be seen to capture a broad range of approaches, but in essence it involves the regulator challenging specific plans, processes and operations of the company and, having identified weaknesses, setting an efficiency target based on eliminating those specific weaknesses. In this way the efficiency target is developed from the bottom-up, based on a set of specific efficiency-improving initiatives that the firm should be able to implement based on the regulator's analysis.

Such analysis can be seen to have a key advantage over top-down approaches. In the latter case, a performance gap is identified but there is no clear guidance on how the efficiency gap can be closed; though econometric models do identify the best and worst performers, which can then point the way to finding solutions. Importantly, bottom-up studies are can provide supporting evidence for the results of the econometric model, whilst offering insights into how firms can improve.

There are, however, disadvantages. Such approaches are likely to understate the scope for savings since it will not be possible to identify all possible ways of delivering efficiency gains; and indeed a stretch target, set by a top-down approach, can be useful in that regard. Bottom-up approaches also raise questions about who should be running the firm; that is, if consultants and regulators know how to do it, shouldn't they be doing it? A further issue can be the credibility of such approaches as they are typically based on the judgement of the regulator/consultants, rather than hard evidence. Finally, whilst consultants may base their findings in part on their experience of how, for example, rail renewals are done in other countries, but there could be genuine questions as to how approaches from elsewhere can really be applied in Britain (without upsetting other processes, which may in fact be efficient). Overall, however, it is clear that given the large amounts of money involved, regulators need to exercise considerable due diligence in arriving at their conclusions, and using a range of econometric methods, combined with bottom-up approaches and wider incentive mechanisms can be seen, in general, to be highly effective.

We note below a significant example of the use of bottom-up methods to corroborate the results of top-down approaches (Smith et al., 2010) from the 2008 Periodic Review of Network Rail.

A substantial efficiency gap of the order of 40% had been identified through top-down econometric evidence. The examples in Table 6, based on bottom-up studies by the regulator and its consultants, found evidence that the adoption of best practice could lead to very substantial savings; in line with the econometric evidence. For further details see Smith et al. (2010).

A final point may be made at this stage. It is easier for regulators in general to make efficiency assessments through top-down methods where there are several firms under the same regulator. For transport infrastructure – certainly in rail – there is typically only one firm, making the assessment more challenging for the regulator. Thus regulators may need to place greater emphasis on international benchmarks, which can be challenging because of data comparability issues and gaining the level of commitment needed from international railways to conduct the analysis. There could also be greater issues in introducing practices from one country to another.

However, even within a single rail or road infrastructure provider, there may be several regions or units that can be benchmarked. In Britain, increased power has been/is being devolved to Network Rail's eight routes (broadly these can be seen as regions); and these can be benchmarked against each other. A further advantage here is that the regulator (and/or government as the owner) should be able to put pressure on Network Rail not only to identify performance gaps between routes, but to discuss and search for ways of applying the company's own best practice across the network, tailored for each region/route. In a similar way, the CQC Efficiency Network in local authority roads not only seeks to identify efficiency gaps but, as a benchmarking club, to identify best practice and see it promulgated across the sector.

Summary

This section has explained the purpose of top-down econometric benchmarking, highlighting its potential power for primarily controlling for the nature of technology (economies of scale and density) and then in identifying a set of high-level efficiency scores. These targets can then be used by regulators, with suitable adjustments and regulatory judgement, and combined with bottom-up evidence, to set appropriate efficiency targets (within the context of an incentive-based approach to regulation). Bottom-up benchmarking can give reassurance that the results of top-down approaches are credible and the efficiency targets derived from them achievable. It can also support learning across organisations and this can be particularly useful in contexts where benchmarking clubs are formed, e.g. local authority roads. It can also be useful within organisations, for example, in respect of Network Rail route benchmarking, where such an approach may help support learning across regions/routes. Regulators must also guard against the concern that they become too involved in the micro-management of regulated firms. Ultimately the power of top-down benchmarking is that it gives a data-based assessment of performance differences, and shines the spotlight on where to look for potential efficiency savings.

Asset inspection and asset management	In general best practice European railways undertake fewer track inspections but inspections are generally of higher quality. It is estimated that similar techniques applied in Britain could reduce foot patrolling inspection costs by around 75% and tamping expenditure by 20%
Recycling components	This is common European practice. In Switzerland, for example, rail, point motors, sleepers and signal heads are regularly refurbished then cascaded from higher to lower category routes. Cascaded rail on lines re-laid with steel sleepers could lead to savings. Additionally ballast cleaning (partial renewal) as opposed to traxcavation (complete renewal) could reduce ballast renewal cost in Britain by 40%
High output rail stressing	Stressing continuously welded rail by heating it rather than physically stretching it is a process discontinued in Britain in the 1960s and 1970s. Some European networks (using modern equipment) have re-introduced this method which doubles on-site productivity and, if applied to the renewals re-railing workbank in CP4, could lead to significant annual savings for Network Rail
Formation rehabilitation trains	Modern high output European plant is regularly used to undertake formation and also ballast renewals. If applied to Network Rail's CP4 category 7 and 12 track renewals RailKonsult estimate that it could reduce unit costs for both activities by around 40%
Lightweight station platforms	The use of modular construction polystyrene station platforms in the Netherlands could provide opportunities in Britain, given the substantial CP4 platform extension workbank. Analysis suggests a unit cost saving of around 25% in Britain
Efficient European re-railing techniques.	This particular study brought together many themes from the previous RailKonsult work by focussing upon the Swiss re-railing method. Bespoke plant, high output welding techniques and dedicated teams are applied routinely. Put together for basic re-railing work alone this method is around 40% more efficient than current Network Rail practice
Use of dedicated teams	Contractors are widely used by most continental railways, as they are in Britain. However there is generally a greater degree of specialisation by activity in Europe (such as S&C renewal or tamping). This ensures a highly skilled and productive workforce dedicated to particular tasks in contrast to the situation in Britain where contractors are often not even dedicated to rail.

Table 6. Examples of bottom-up evidence

Source: Smith et al. (2010).

We have focused on the particular challenges that are faced in carrying out CAPEX benchmarking, given their lumpy and temporal nature. Approaches have been developed to deal with such issues and though they are not perfect we consider that they can enable a sensible econometric benchmarking framework to be implemented; and we have shown examples of such implementation. It is clear that enhancements are more challenging however, particularly in the context of a situation where the level of enhancements spend and choice of schemes is set by government. Particularly for enhancements in transport, and to an extent, renewals, the use of intermediate outputs and the unit cost of delivering projects, is a useful way forward in benchmarking (rather than question the schemes themselves, which has a separate appraisal process).

More widely, and not limited to CAPEX, benchmarking always has to overcome potential issues of data comparability and these are not trivial, and commitment from companies (either through a benchmarking club or via regulatory pressure) to achieve common data definitions is a time consuming and important task. In this respect international benchmarking is particularly challenging because it takes considerable time and commitment from a group of countries (or companies) to make the analysis credible and usable. If international benchmarking is to work, then it may require concerted efforts by regulators/governments across Europe working together to establish a common benchmarking framework.

Conclusions

This report has focused on the following areas:

- setting out the conceptual issues relating to capital bias and reviewing the relevant literatures
- explaining the regulatory frameworks in the three cases, and assembling data to assess the evidence for capital bias in those cases
- explaining the purpose of benchmarking within the regulatory framework
- discussing and showing possible solutions for the particular challenges of CAPEX benchmarking, which also reflects the focus of the Working Group.

The conclusions around CAPEX bias indicate that in general this may be seen to be an over-played issue in the literature; however there are reasons why such biases may exist and UK economic regulators have been alert to these. It is hard to demonstrate such bias using available data, but what data we do have does not seem to indicate a CAPEX bias across the three sectors we are considering. Nevertheless, some regulators have sought to benchmark aspects of OPEX and CAPEX together, and also adjust other aspects of the regulatory regime to counter the possibility of CAPEX bias. The possibility of OPEX bias in nationalised industries facing short-term funding constraints have also been highlighted; and in the absence of an independent regulator with a focus on funding levels and efficiency and / or a multi-annual agreement that respects long-term planning and whole-life costing, this could be an issue in transport infrastructure investment.

We have set out the purpose of top-down econometric benchmarking – namely to produce a set of high-level efficiency scores for use in setting efficiency targets; and noted its particular value in that prior to arriving at an efficiency score, the nature of technology (e.g. economies of scale and density) is first controlled for. Further, there is transparency over the results concerning the shape of the technology that can be challenged and, if accepted as reasonable, this process gives confidence in the resulting efficiency scores.

Ultimately, the results of an econometric model (or models) represent only one step in the process of setting efficiency targets and delivering efficiency savings on the ground. Regulators may make suitable adjustments and apply regulatory judgement, and also combine the results of the top-down models with bottom-up evidence. Further, regulators will want to provide an appropriate incentive mechanisms for risk sharing (e.g. via menus) and outperformance. Bottom-up benchmarking – where specific initiatives for improving efficiency are identified – can be and are used powerfully alongside top-down methods to corroborate the latter, and ensure that realistic targets are set. In some contexts, e.g. local authority roads or Network Rail route benchmarking, bottom-up methods may help support learning across regions/routes and thus support the delivery of efficiency savings on the ground (with top-down benchmarking having established evidence on where to look for good and potentially poor practice).

However, importantly, regulators must also guard against the concern that they become too involved in the micro-management of regulated firms, as it raises the question as to who is running the firm, and what freedom is being given to managers to optimise. This is particularly true where there are several, private firms being benchmarked. However, particularly in situations where there is a single infrastructure manager, owned by the state (as is the case for all rail infrastructure providers in Europe), it is less clear that the regulator can stand back and take a hands-off approach. Nevertheless, management incentive schemes and reputational incentives can still provide strong incentives for managers of state-owned firms to perform well; and there is a strong danger of regulator capture if regulators become too closely involved in the detail.

We have focused on the particular challenges that are faced in carrying out CAPEX benchmarking, given their lumpy and temporal nature. We have outlined approaches that have been developed to deal with such issues and though they are not perfect we consider that they can be used to support a strong benchmarking framework; and we have shown examples of such implementation. It is clear that enhancements are more challenging however, particularly in a situation where the level of spending on enhancements and choice of schemes is set by government. Particularly for enhancements in transport, and to an extent, renewals, a focus on project level data – assessing the unit cost (per intermediate output) of delivering projects – is a useful way forward in benchmarking (rather than questioning the schemes themselves, which has a separate appraisal process). Maintenance and renewals expenditure may then be benchmarked separately as a form of TOTEX benchmarking (the term BOTEX benchmarking has been used to refer to this situation).

We have also noted the move to take enhancements entirely out of the regulatory framework for rail in Great Britain, to be dealt with entirely by DfT. Whilst we understand the reasons for this decision, we raise concerns, particularly with regard to the loss of independent regulatory scrutiny on the link between costs and outputs and also the impact of enhancement schemes on the day-to-day railway operations (which the regulator retains oversight of).

Overall, we consider that regulatory benchmarking within an RPI-X model, including of CAPEX, has proven to deliver significant benefits across several sectors. It is not without problems but it can be seen to be a useful alternative to PPPs. We do recognise that transport applications potentially raise greater challenges than other sectors, though in part this reflects the state ownership in these sectors, and also that there may be one firm. Provided regional benchmarking of such operators can be implemented – as is happening in rail for example, and provided there is a strong independent regulator, we believe the approach can work even in these more challenging contexts.

Finally, and not limited to CAPEX, care always needs to be made with regard to the comparability of data and significant time investment is needed to ensure comparable definitions (either through a benchmarking club or via regulatory pressure). Here international benchmarking is particularly challenging, and if it is to work – given that companies in different countries are unlikely to be covered by a common regulator – concerted efforts by regulators/governments across Europe may be required.



1 Beyond the choice between maintenance, renewal and enhancement, network operators also have to decide how works should be achieved. For instance once a renewal decision has been made, the amount of CAPEX might be influenced by the time at which construction is done. Favouring renewal works at night will decrease the impact they have on traffic but will increase the cost of the overall operation. Similarly, small possessions for maintenance will limit their impact on traffic, but will increase the overall cost of renewal operations.

2 According to the authors, if the allowable rate of return is less than the actual cost of capital, the firm withdraws from the market, making it a trivial case.

3 Which in their sample can be price regulation or revenue sharing regulation as opposed to rate of return regulation.

4 The WACC should be equal to the weighted cost of debt and equity. If the former is readily observable, the cost of equity has to be set according to perceived risk. In practice, it stems from negotiations between regulated firms and the regulators and acts as an allowed return on capital. As such the WACC can be set at a high level to trigger investments (as pointed out by Rondi & Cambini), and overestimating the WACC could lead to an AJ effect.

5 By which we mean that costs incurred within the five-year control period are entirely paid for by customers (or government) within the five-year period.

6 Determining the value of assets with long life cycles can be subject to interpretation. Such debates are particularly relevant as ownership changes with privatisation. See for example, McCartney and Stittle (2015) for a discussion regarding the case of Railtrack.

7 ORR (1999) par 2.18 of "The periodic review of Railtrack's access charges: Provisional conclusions on revenue requirements".

8 ORR (1999) par 2.18 of "The periodic review of Railtrack's access charges: Provisional conclusions on revenue requirements".

9 Ofgem (2008) par 4.56 of "Electricity Distribution Price Control Review Policy Paper".

10 "CAPEX bias in the water and sewerage sectors in England and Wales -substance, perception or myth? A discussion paper", p. 10.

11 "CAPEX bias in the water and sewerage sectors in England and Wales --substance, perception or myth? A discussion paper", p. 11.

12 Page 3 of the version prepared for the National Bureau of Economic Research Conference on Economic Regulation, 9-10September, 2005.

13 Ofwat (2013) Section 4.5 of "Setting price controls for 2015-20 - framework and approach: A consultation".

14 Cherchye, L., B. De Rock, A. Estache and M. Verschelde (2015), "Efficiency Measures in Regulated Industries: History, Outstanding Challenges and Emerging Solutions". ECARES Working Papers.

15 Ofgem (2010) par. 8.37 of "Handbook for implementing the RIIO model".

16 Anglian Water (2017) "Water industry cost Modelling. Anglian Water's approach and initial results".

17 Ofwat (2016) "Costs and benefits of introducing competition to residential customers in England".

18 Ofwat (2015) "The form of the price control for monopoly water and sewerage services in England and Wales - a discussion paper".

19 There are some very minor contributions from other sources for instance property developers.

20 ORR (1999) par 2.18 of "The periodic review of Railtrack's access charges: Provisional conclusions on revenue requirements".

21 Office of Rail Regulation, October 2013, Final determination of Network Rail's outputs and funding for 2014-19, p. 133.

22 Cf. the multiannual contract signed by SNCF Reseau and the French State (April 2017).

23 ORR (2016) "Working paper 5: Options for the treatment of enhancements in PR18".

24 Note that this is a separate issue to the data classifications reported in Figure 6 earlier in this report. That data contains the sum of maintenance and renewals costs for Local Highways Authorities in England. To reiterate the conclusions with respect to local roads in England, there is no evidence of a capital bias as evidenced by the very poor condition of the network.

25 It should be noted that this point is made particularly in respect of the CQC network, in which participants are voluntary members. In other contexts, economic regulators may require firms to prepare data to certain definitions, though regulators have to be mindful of what is possible and the cost of the requirements they impose.

26 Appendix 4.

27 It should be noted that Ofwat has used other approaches, termed, "unit cost models" to benchmark capital maintenance (renewals) in water (Ofwat, 2007). The "unit cost model" approach involves computing simple measures of aspects of capital maintenance costs divided by

measures of final outputs (such as number of connected properties). This approach has been taken when econometric models for capital maintenance (see Ofwat, 2004, Appendix 3 for more details on the econometric approach) have not worked. The approach uses overall company level data.

28 Note these are different to the unit cost models noted in Footnote 11, where the approach related renewals costs to final, rather than intermediate outputs, and did not use an econometric approach.

29 See page 336 of the ORR's Final Determination: http://orr.gov.uk/ data/assets/pdf_file/0011/452/pr13-final-determination.pdf

References

Anglian Water (2015), Future use of menus as part of price setting methodology. <u>http://www.anglianwater.co.uk/ assets/media/Future use of menus as part of price setting metho</u> dology - Final.pdf

Ai. C., S. Martinez and D.M. Sappington (2004), "Incentive regulation and telecommunications service quality". *Journal of Regulatory Economics*, Vol. 26/3, pp. 263-285.

Australian Competition and Consumer Commission (2012). Benchmarking OPEX and CAPEX in Energy Networks *Working Paper no.6.*

Averch, H. and L.L. Johnson (1962), "Behavior of the firm under regulatory constraint". *The American Economic Review*, Vol. 52/5, pp. 1052-1069.

Beesley, M.E. and S.C. Littlechild (1992), "The Regulation of Privatized Monopolies in the United Kingdom", in Beesley, M.E. (ed.), *Privatization, Regulation and Deregulation*.

Benedetto, V., A.S. Smith C.A. Nash (2017), "Evaluating the roles and powers of rail regulatory bodies in Europe: A survey-based approach". *Transport Policy*, Vol. 59, pp. 116-123.

Bowe, C. (2015), Report of the Bowe Review into the Planning of Network Rail's Enhancements Programme 2014-2019.

Boycko, M., A. Shleifer and R.W. Vishny (1996), "A theory of privatisation". *The Economic Journal*, pp. 309-319.

Brunekreeft, G. and R. Meyer (2011), Regulation and Regulatory Risk in the Face of Large Transmission Investment. *Competition and Regulation in Network Industries*, Vol. 12/2, pp. 155-172.

Cabral, L.M. and M.H. Riordan (1989), "Incentives for cost reduction under price cap regulation". *Journal of Regulatory Economics*, Vol. 1/2, pp. 93-102.

Cambini, C. and L. Rondi (2010), "Incentive regulation and investment: evidence from European energy utilities". *Journal of regulatory economics*, Vol. 38/1, pp. 1-26.

Competition and Markets Authority (CMA) (2015), Bristol Water plc A reference under section 12(3)(a) of the Water Industry Act 1991 Report, Appendix 2.3.

Cook, A. (2011), A fresh start for the Strategic Road Network Managing our roads better to drive economic growth, boost innovation and give road users more for their money, <u>http://webarchive.nationalarchives.gov.uk/20140324063150/https:/www.gov.uk/government/uploads/s</u> <u>ystem/uploads/attachment_data/file/4378/strategic-road-network.pdf</u> (accessed 31 January 2018).

Department for Transport (DfT) (2015), Road Investment Strategy: for the 2015/16 – 2019/20 Road Period, <u>https://www.gov.uk/government/publications/road-investment-strategy-for-the-2015-to-2020-road-period</u> (accessed 30 January 2017).
Department for Transport, (2017), Road lengths in Great Britain 2016, <u>https://www.gov.uk/government/statistics/road-lengths-in-great-britain-2016</u> (accessed on 18

November 17).

Electranet (2015), AER Expenditure Incentive Schemes. <u>https://www.electranet.com.au/wp-content/uploads/resource/2016/06/2015-Factsheet-ConsumerAdvisoryPanel.pdf.</u>

Evans, L.T. and G.A. Guthrie (2005), Risk, price regulation, and irreversible investment. *International Journal of Industrial Organization*, Vol. 23/1, pp. 109-128.

Friebel, G., M. Ivaldi and C. Vibes (2010), Railway (de) regulation: a European efficiency comparison. *Economica*, Vol. 77/305, pp. 77-91.

Growitsch, C., T. Jamasb, C. Müller and M. Wissner (2010), Social cost-efficient service quality— Integrating customer valuation in incentive regulation: Evidence from the case of Norway. *Energy Policy*, Vol. 38/5, pp. 2536-2544.

Hansford, P. (2017), The Hansford Review – Unlocking rail investment – building confidence, reducing costs: An independent review chaired by Professor Peter Hansford FREng. <u>https://cdn.networkrail.co.uk/wp-content/uploads/2017/07/The-Hansford-Review.pdf.</u>

Joskow, P.L. (2008), Incentive regulation and its application to electricity networks. *Review of Network Economics*, Vol. 7/4.

Kemp, G. (2017), *Local Highways: Moving Britain Ahead March 2017*. <u>http://www.apse.org.uk/apse/assets/File/Gary%20Kemp%20-%20Session%201.pdf</u> (accessed on 21 August 2017).

Kennedy, J. and A.S. Smith (2004), Assessing the efficient cost of sustaining Britain's rail network: Perspectives based on zonal comparisons. *Journal of Transport Economics and Policy (JTEP)*, Vol. 38/2, pp. 157-190.

Laffont, J.J. and J. Tirole (1986), "Using cost observation to regulate firms". *Journal of political Economy*, Vol. 94/3, Part 1, pp. 614-641.

Law, S.M. (2014), "Assessing the Averch-Johnson-Wellisz Effect for Regulated Utilities". *International Journal of Economics and Finance*, Vol. 6/8, pp. 41.

Leaver, C. (2009), "Bureaucratic minimal squawk behavior: Theory and evidence from regulatory agencies". *The American Economic Review*, Vol. 99/3, pp. 572-607.

LEK (2003), Regional Benchmarking: Report to Network Rail, ORR and SRA, London.

Levy, B. and P.T. Spiller (1994), "The institutional foundations of regulatory commitment: a comparative analysis of telecommunications regulation". *The Journal of Law, Economics, and Organization*, Vol. 10/2, pp. 201-246.

Littlechild, S.C. (1983), *Regulation of British Telecommunications' profitability: report to the Secretary of State, February 1983*. Department of Industry.

Mayat, T. (2017), Local Highways Maintenance Funding. Working note for PhD, University of Leeds.

UK Parliament (2016), *Local road maintenance, repairs and street works in England*, Briefing Paper Number SN739.

McCartney, S. and J. Stittle (2015), "Accounting for producer needs: The case of Britain's rail infrastructure". *Accounting Forum*, Vol. 39/2, pp. 109-120.

McNulty, R. (2011), Realising the Potential of GB Rail, Report of the Rail Value for Money Study. *Summary Report.*

Nash, C.A., B. Matthews and A.S.J. Smith (2017), Deliverable D1.7 Incentives Final Report Annex 1: Research on Incentives for Innovation.

National Audit Office (2011), Regulating Network Rail's efficiency, *Report by the Comptroller and Auditor General*.

Newbery, D.M. (1988), "Road user charges in Britain". The Economic Journal, Vol. 98/390, pp. 161-176.

Ofgem (2008), Electricity Distribution Price Control Review Policy Paper.

Ofgem (2009), Regulating Energy Networks for the Future: RPI-X@20 History of Energy Network Regulation. <u>https://www.ofgem.gov.uk/ofgem-publications/51984/supporting-paper-history-energy-network-regulation-finalpdf.</u>

Ofwat (1999), Future water and sewerage charges 2000 -05 Final determinations. <u>http://webarchive.nationalarchives.gov.uk/20150603222823/http://www.ofwat.gov.uk/pricereview/pr9</u> <u>9/det_pr_fd99.pdf</u>

Ofwat (2004), Future water and sewerage charges 2005 -10 Final determinations. <u>http://webarchive.nationalarchives.gov.uk/20150603192547/http://www.ofwat.gov.uk/pricereview/pr0</u> 4/det_pr_fd04.pdf

Ofwat (2007), Relative efficiency assessments 2006-07 – supporting information. <u>http://webarchive.nationalarchives.gov.uk/20140715093441/http://www.ofwat.gov.uk/regulating/repor</u> <u>ting/rpt_unc_2006-07supp.pdf</u>

Ofwat (2011). CAPEX bias in the water and sewerage sectors in England and Wales: Substance, perception or myth? Discussion paper.

Ofwat (2014), Basic cost threshold model Appendix C Enhancement modelling. <u>http://webarchive.nationalarchives.gov.uk/20150603202629/http://www.ofwat.gov.uk/pricereview/pr1</u> 4/pap_tec1402feederbasiccostappc.pdf

ORR (2013), PR13 Efficiency Benchmarking of Network Rail using LICB.

ORR (2015), Monitoring Highways England: The monitoring framework. October 2015. London.

ORR (2017), Improving Network Rail's renewals efficiency: A consultation. http://orr.gov.uk/__data/assets/pdf_file/0012/25221/pr18-improving-network-rail-renewals-efficiencyconsultation-july-2017.pdf

OXERA (April 2008), Network Rail's scope for efficiency savings in CP4: {repared for Office of Rail Regulation, London. <u>http://www.rail-reg.gov.uk/upload/pdf/pr08-oxeraeffic-160408.pdf</u>

Pollitt, M.G. (1999), The restructuring and privatization of the Electricity Supply Industry in Scotland. *Cambridge University, Department of Applied Economics, mimeo*.

Pollitt, C. (2005), Performance management in practice: A comparative study of executive agencies. *Journal of Public Administration Research and Theory*, Vol. 16/1, pp. 25-44.

Rees, R. (1984), The public enterprise game. *The Economic Journal*, Vol. 94, pp. 109-123.

Shleifer, A. (1985), A theory of yardstick competition. *The RAND Journal of Economics*, pp. 319-327.

Smith, A.S. (2006), "Are Britain's railways costing too much? Perspectives based on TFP comparisons with British Rail 1963–2002". *Journal of Transport Economics and Policy (JTEP)*, Vol. 40/1, pp. 1-44.

Smith, A.S.J. (2012), "The application of stochastic frontier panel models in economic regulation: Experience from the European rail sector". *Transportation Research Part E: Logistics and Transportation Review*, Vol. 48/2, pp. 503-515.

Smith, A.S.J. (2008). International Benchmarking of Network Rail's Maintenance and Renewal Costs: An Econometric Study Based on the LICB Dataset (1996-2006). *Report for the Office of Rail Regulation*.

Smith, A.S.J. and P.E. Wheat (2012), "Estimation of cost inefficiency in panel data models with firm specific and sub-company specific effects", *Journal of Productivity Analysis*, Vol. 37, pp.27-40. <u>http://eprints.whiterose.ac.uk/82983/</u>

Smith A.S.J. Wheat P.E. and Smith G. (2010). "The role of international benchmarking in developing rail infrastructure efficiency estimates", *Utilities Policy*, Vol. 18, pp.86-93. http://eprints.whiterose.ac.uk/82981/

Stern, J. (1997), "What makes an independent regulator independent?" *Business Strategy Review*, Vol. 8/2, pp. 67-74.

Viscusi, W.K., J.E, Harrington and J.M. Vernon (2005), *Economics of regulation and antitrust*. MIT press.

Appendix 1. Research questions and outputs of the Working Group on Private Investment in Infrastructure

Introduction: Getting the basics right

What are the economic characteristics of infrastructure? What is infrastructure and what are operations? What are the models of private participation in infrastructure and through which significant private investment actually takes place?

Can private investment improve productive efficiency? Improve project selection? Close the infrastructure funding gap? Have other positive effects when it is private?

What have the private investment trends in transport infrastructure been over the last 20 years? How much of that was foreign private investment?

Makovšek, D. (2019), "What is Private Investment in Transport Infrastructure and Why is it Difficult?", Working Group Paper, International Transport Forum, Paris.

Makovšek, D. (2019), "The Role of Private Investment in Transport Infrastructure", Working Group Paper, International Transport Forum, Paris.

Mistura, F. (2019), "Quantifying Private and Foreign Investment in Transport Infrastructure", Working Group Paper, International Transport Forum, Paris.

Defining the challenge: How uncertainty in contracts matters

How does uncertainty affect risk pricing? Beyond investors, do suppliers in PPPs also have issues with risk pricing? How does its transfer to the private sector affect competition? What does uncertainty mean for the public vs. private cost of financing?

Makovšek, D. and Moszoro, M. (2018), "Risk pricing inefficiency in public-private partnerships", Transport Reviews, 38(3), 298-321.

<i>Is uncertainty also an issue in long-term services/operations contracts?</i>	Beck et al. (forthcoming), "Uncertainty in Long-term Service Contracts: Franchising Rail Transport Operations", Working Group Paper, International Transport Forum, Paris.
What is the competition for large transport infrastructure projects in the EU Market? Is there a difference between traditional procurement and PPPs?	Roumboutsos, A. (forthcoming),"Competition for Infrastructure Projects: Traditional Procurement and PPPs in Europe", Working Group Paper, International Transport Forum, Paris.

Addressing uncertainty for suppliers: the construction phase as example

Adversarial vs. collaborative procurement – is collaborative contracting the future?	Eriksson et al. (forthcoming), "Collaborative Infrastructure Procurement in Sweden and the Netherlands", Working Group Paper, International Transport Forum, Paris.
What lessons in dealing with risk and uncertainty were learnt in Danish mega projects from Storebaelt to Femernbaelt?	Vincentsen, L. and K. S. Andersson (2018), "Risk Allocation in Mega-Projects in Denmark", Working Group Paper, International Transport Forum, Paris.
What can governments do in the short run to reduce inefficient pricing of risk by construction contractors?	Kennedy et al. (2018), "Risk Pricing in Infrastructure Delivery: Making Procurement Less Costly", Working Group Paper, International Transport Forum, Paris.

Addressing uncertainty in long-term contracts in the absence of continuous pressure for efficiency

What is the public sector organisational counterfactual on which private investment should seek to improve?	Holm, K.V. and T.H. Nielsen (2018), "The Danish State Guarantee Model for Infrastructure Investment", Working Group Paper, International Transport Forum, Paris.
Partial fixes to the Private-Public Partnership approach	
How would an organisational structure consisting of PPPs come close to a network-wide management approach? What benefits would it yield?	Vasallo, J. (forthcoming), "Public-Private Partnerships in Transport: Unbundling Prices from User Charges", Working Group Paper, International Transport Forum, Paris.
Should the public or the private side bear the cost of long- term uncertainty? How could we design a PPP contract to avoid hold-up due to incomplete contracts?	Engel et al., (forthcoming), "Dealing with the Obsolescence of Transport Infrastructure in Public-Private Partnerships", Working Group Paper, International Transport Forum, Paris.
Long-term strategic approach	
<i>How do the PPP and regulated utility model (RAB) compare in terms of efficiency incentives?</i>	Makovšek, D. and D. Veryard (2016), "The Regulatory Asset Base and Project Finance Models", International Transport Forum Discussion Papers, No. 2016/01, Paris.

What basic considerations underlie the choice between a PPP and RAB approach?	Hasselgren, B. (forthcoming), "Risk allocation in Public-Private Partnerships and the Regulatory Asset Base Model", Working Group Paper, International Transport Forum, Paris.
Which are the preconditions a country would need to take to establish a RAB model on a motorway network? Is user- charging a must?	Alchin, S. (forthcoming), "A Corporatised Delivery Model for the Australian Road Network", Working Group Paper, International Transport Forum, Paris.
From the investors' point of view, does a RAB need to be fully reliant on user-charging?	Francis, R. and Elliot, D. (2019), "Infrastructure Funding: Does it Matter Where the Money Comes From?", Working Group Paper, International Transport Forum, Paris.
Incentive regulation can also yield perverse incentives. Can the capex bias be managed?	Smith, A. et al. (2019), "Capex Bias and Adverse Incentives in Incentive Regulation: Issues and Solutions", Working Group Paper, International Transport Forum, Paris.
Does it make sense to pursue hybrid solutions between PPP and RAB?	Zhivov, N. (2018), "The Thames Tideway Tunnel: A Hybrid Approach to Infrastructure Delivery", Working Group Paper, International Transport Forum, Paris.

Uncertainty and private investment mobilisation in transport infrastructure

What lessons can we draw from recent attempts to mobilise private investment in infrastructure in the aftermath of the global financial crisis?	Makovšek, D. (2018), "Mobilising Private Investment in Infrastructure: Investment De-Risking and Uncertainty", Working Group Paper, International Transport Forum, Paris.
Synthesis	ITF (2018), Private Investment in Transport Infrastructure: Dealing with Uncertainty in Contracts, Research Report, International Transport Forum, Paris

CAPEX Bias and Adverse Incentives in Incentive Regulation

Issues and Solutions

Most privately financed transport infrastructure is delivered through Public-Private Partnerships. An alternative approach is the Regulatory Asset Base model, but this can be biased towards capital expenditure to the detriment of operating expenditure. This report investigates whether economic regulators can hold "CAPEX bias" in check. The paper is part of a series of 19 papers and a synthesis report produced by the International **Transport Forum's Working Group** on Private Investment in Transport Infrastructure.

International Transport Forum 2 rue André Pascal F-75775 Paris Cedex 16 +33 (0)1 73 31 25 00 contact@itf-oecd.org www.itf-oecd.org





Transport Forum



Transport Forum





Transport Forum



Transport Forum













ubllished: 08/2019 | Photo credit: Aarón Borrás







