



**UNIVERSITY OF LEEDS**

This is a repository copy of *Regional Benchmarking of the British Rail Infrastructure Manager: A Long Panel Approach*.

White Rose Research Online URL for this paper:  
<https://eprints.whiterose.ac.uk/102603/>

Version: Accepted Version

---

**Proceedings Paper:**

Rivas Amiasorho, E, Wheat, PE [orcid.org/0000-0003-0659-5052](https://orcid.org/0000-0003-0659-5052) and Smith, A [orcid.org/0000-0003-3668-5593](https://orcid.org/0000-0003-3668-5593) (2016) *Regional Benchmarking of the British Rail Infrastructure Manager: A Long Panel Approach*. In: ITEA Annual Conference and School on Transportation Economics. International Transportation Economics Association (ITEA), 13-17 Jun 2016, Santiago, Chile.. ITEA .

---

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# REGIONAL BENCHMARKING OF THE BRITISH RAIL INFRASTRUCTURE MANAGER A LONG PANEL APPROACH

M<sup>a</sup> Eugenia Rivas Amiassorho<sup>1</sup> (Junior Scholar), Phill Wheat<sup>2</sup>, Andrew Smith<sup>2</sup>

<sup>1</sup>Universidad ORT Uruguay

<sup>2</sup>Institute for Transport Studies, University of Leeds

May 2016

## 1. INTRODUCTION

The performance of the rail industry in the last decades have included in some countries reductions in market shares, increases in subsidies or lack of adequate rates of return, which have been responsible for a wide range of regulation and institutional reforms, including vertical separation and privatisations (Nash, 2000). In this context, the assessment of the rail sector performance constitutes a key issue. Rail companies need to identify the cost and productivity drivers, and regulators should be able to assess if the costs of the regulated firms are reasonable and if promoting cost reductions through regulation is possible (Nash and Shires, 2000). Efficiency measurement is especially important when there is no strong competition in the market, as in the case of a rail infrastructure manager, and cost efficiency should be promoted through economic regulation.

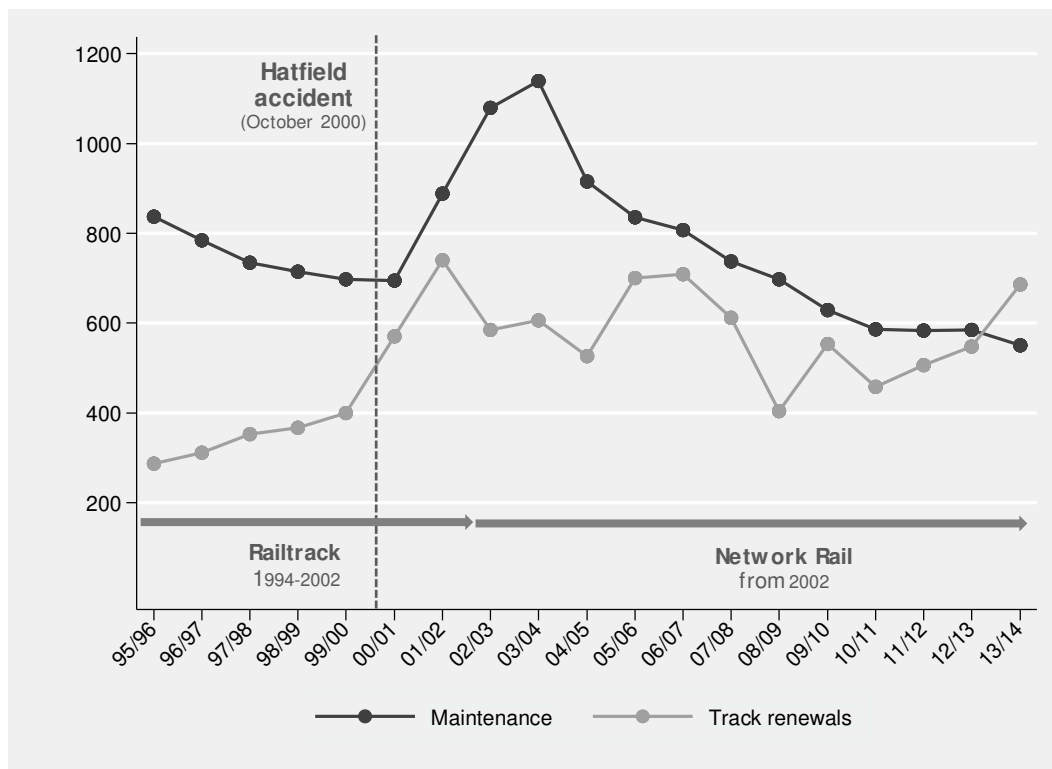
The efficiency analysis in the UK has been driven by the incentive regulatory system of price-cap regulation. The *RPI – X* regulation is extended in the utilities industries in the UK, where the benchmarking performance of utilities is used to determine *X* factors in *RPI – X* regulation (Jamasp and Pollitt, 2001). This incentive regulation is used by the Office of Rail Regulation<sup>1</sup> which controls Network Rail (Smith et al., 2010) and have commissioned several efficiency studies to assess the performance of the rail infrastructure manager in the context of periodic reviews.

The efficiency studies of the rail infrastructure manager in the UK have covered bottom-up (engineering) and top-down (statistic) benchmarking. The use of a top-down international benchmarking started with the initial studies of the *ORR* and the Institute for Transport

---

<sup>1</sup> Currently, Office of Rail and Road.

Studies (University of Leeds) in 2005, followed by its implementation in the 2008 Periodic Review (*PR08*) (Smith et al., 2010). In the 2013 Periodic Review (*PR13*), an efficiency gap between 13% and 24% was estimated for Network Rail with respect to its European peers. International comparisons are especially relevant in the case of Network Rail because of its monopolistic nature, and the necessity to keep costs under control, and promote the implementation of international best-practices. However, the comparison among units of the same company, or internal benchmarking, can also help to improve efficiency by understanding the intra-company differences and implementing internal best-practices. The last top-down benchmarking of the rail infrastructure benchmarking in the UK from an internal perspective is the work of Kennedy and Smith (2004) which evaluates the performance of Railtrack in the period 1995/96-2001/02. Unfortunately, this study only covers a short period after the Hatfield accident (October 2000), which led to a sharp increase in costs, the collapse of Railtrack and the establishment of Network Rail in 2002 (Figure 1).



**Figure 1: Maintenance and track renewal costs**  
Source: own work

The aim of this paper is to analyse the cost efficiency of the UK rail infrastructure manager in the period 1995/96 to 2013/14, covering the operation of Railtrack (1994-2002) and

Network Rail (from 2002), by examining the regional performance over time, with special focus after the Hatfield accident, and by comparing these results with the international benchmarking evidence. The contribution of this paper is to provide an updated internal econometric benchmarking identifying the top performing regions and the potential cost savings that can be achieved by Network Rail by understanding the intra-company differences.

The paper is divided in four main sections after this introductory section. The second section presents an overview of the context of the efficiency analysis and main techniques, and the key efficiency studies of the rail infrastructure manager in the UK. The third section describes the methodology with focus on the econometric techniques. The fourth section examines the main results of the internal benchmarking and its comparison with the international benchmarking. Finally, the last section presents the concluding remarks.

## **2. BENCHMARKING OF RAILWAY INFRASTRUCTURE MANAGERS**

### **2.1. Benchmarking approaches**

Different approaches have been used to assess efficiency in the rail industry, where it can be distinguished index numbers, total factor productivity, and models explaining the performance of individual railways (Oum et al., 1999). In particular, the traditional econometric methods allow the estimation of production and cost function, implicitly assuming that the firms are always on the frontier. On the other hand, frontier methods recognize that not all the firms are efficient, decomposing the error in the model into an inefficiency term and a random noise term. The method is known as a deterministic and stochastic frontier method depending on the inefficiency component being deterministic or stochastic respectively<sup>2</sup>.

In the case of top-down (statistical) approaches, it is possible to identify between internal or external benchmarking. The former is the comparison of different units within the same

---

<sup>2</sup> For a complete review see for example Parmeter and Kumbhakar (2014).

company, whereas the second one is the comparison with others outside the company<sup>3</sup>. The international comparisons can be a very useful tool, but the difficulty of comparing data of different countries over time is one of the main challenges for the regulators (Smith, 2012), especially when the data availability is compromised by organizational changes, and restructures in the ownership and integration of the railways (Makovsek et al., 2015). The risk of bias in the assessment of performance also has been noted as a drawback of benchmark analysis, where there is a potential risk of conducting ex-post benchmarking in order to support previous strategies of vertical separation (Hansen et al., 2013).

In this context, internal benchmarking presents the advantage of homogeneity regarding criteria and procedures used for collecting the data. Nevertheless, it is likely that results of internal benchmarking show less possibilities of efficiency gains than results of external benchmarking (Smith et al., 2010), because the best performing regions can be inefficient in some way.

## **2.2. Regulation and benchmarking**

The move towards a price incentive regulation such as  $RPI - X$  has required the conduction of performance analysis to establish the efficiency goals as well as ex-post assessments to evaluate the results of the regulations (Coelli and Lawrence, 2006). Efficiency assessment constitutes an important component of economic regulation, where regulators are interested in measuring the potential efficiency gains as part of periodic reviews or price controls (Smith et al., 2010).

In the  $RPI - X$  regulation, a regulated firm cannot increase its price by more than the increase in the retail price index ( $RPI$ ) minus a negotiated factor ( $X$ )<sup>4</sup> (Ricketts, 2006), introducing the incentives to retain the cost savings greater than  $X$ . However, because the costs of the firms are used to reset  $X$ , there is a risk of a non-efficient behaviour by the firms in order to affect the regulation in the future (Burns et al., 2006).

---

<sup>3</sup> Smith and Wheat (2012) analyse a dual-inefficiency model by dividing firm inefficiency into two components, one of them reflecting internal inefficiency (differences across units within a firm) and the other reflecting external inefficiency (persistent component across all units in the same firm).

<sup>4</sup> It can include changes in input prices which represent a significant part of costs and are outside firm's control.

In fact, as is pointed out by Smith et al. (2010), the assessment of  $X$  is far from simple. In simple terms,  $X$  can be defined as the expected total factor productivity ( $TFP$ ) growth in the productivity that the firm can achieve above the average performance of the economy reflected implicitly within the  $RPI$ . The factors which affect the potential growth of  $TFP$  can be divided into input price trends and expected productivity gains which includes scale effects, efficiency gains and technical progress (Smith, 2015). If  $X$  is equal to zero, it does not mean that there are no productivity gains. It implies that the regulated firm is expected to achieve the same efficiency gain than the economy as a whole, and it can increase the prices by exactly the same amount of the increase in the  $RPI$ . The determination of  $X$  has resulted into an extensive literature to benchmark the efficiency of the regulated firms through different efficiency and productivity techniques (Fried et al., 2008a), where the required accuracy of the benchmarking estimations can be different depending on the use of the benchmarking (Stern, 2003).

### **2.3. The United Kingdom case**

Several efficiency studies have been conducted to assess the performance of the infrastructure manager in the UK, including top-down (statistical) and bottom-up (engineering) approaches<sup>5</sup>. The studies have been mainly commissioned in the last years in the context of the Periodic Reviews. Even though *ORR* has commissioned different studies to evaluate the implementation of international benchmarking (NERA, 2000) and a bottom-up international benchmarking study during the 2003 access charges review (LEK et al., 2003), it was not until 2005 when a top-down international benchmarking was developed by *ORR* and the Institute for Transport Studies (University of Leeds) in order to use the results in the *PR08* and the following periodic reviews (Smith et al., 2010).

A summary of the main studies of efficiency of the railway industry based on the compilation of Kennedy and Smith (2004), Smith (2005) and own work, is presented in Table 1.

---

<sup>5</sup> The approaches have been used as complementary rather than competitive. For example, in the *PR13* the *ORR* commissioned engineering analysis in order to contribute to understand the efficiency gaps (Makovsek et al., 2015).

**Table 1: Infrastructure manager efficiency studies - UK**

Periodic Review	Study	Type of study	Details
<i>PR00</i>	Booz--Allen and Hamilton Ltd. (1999); (2000) <sup>1</sup>	Bottom-up	<ul style="list-style-type: none"> <li>· Assessment of potential efficiency gains in Railtrack's asset areas.</li> <li>· Efficiency targets around 4%.</li> </ul>
	NERA (2000) <sup>1</sup>	Top-down / International comparison	<ul style="list-style-type: none"> <li>· Analysis of international evidence on rail infrastructure costs through the comparison of productivity levels among different countries and productivity trends.</li> </ul>
	Europe Economics (2000) <sup>1</sup>	Top-down / Comparison with UK privatised monopolies	<ul style="list-style-type: none"> <li>· The experience of other UK privatised network monopolies is recognized as a guide to establish the potential efficiency gains of Railtrack (for the closest comparators the real unit cost reduction is in the range of 3%-7% per annum).</li> <li>· This results were supported by Horton 4 Consulting (2000)</li> </ul>
<i>PR03</i>	LEK et al. (2003) <sup>2</sup>	Bottom-up Review	<ul style="list-style-type: none"> <li>· Review of the planned volume of activity and its necessity and sensibleness regarding scope and timing.</li> </ul>
	Halcrow et al. (2003) <sup>2</sup>	Bottom-up Review	<ul style="list-style-type: none"> <li>· Review of contracting strategy of Network Rail considering international best-practices.</li> </ul>
	LEK (2003) <sup>2</sup>	Internal benchmarking	<ul style="list-style-type: none"> <li>· Comparison of operating, maintenance and renewal unit costs for regions and contract areas, although it was not possible to estimate efficiency scores.</li> </ul>
Other	Kennedy and Smith (2004) <sup>3</sup>	Top-down / Internal benchmarking	<ul style="list-style-type: none"> <li>· Internal benchmarking of the infrastructure manager (1995/96-2001/02).</li> <li>· <i>COLS</i> and <i>SFA</i> techniques are applied. Potential improvements are possible by implementing best-practices across the network</li> </ul>
<i>PR08</i>	Smith (2008) <sup>3</sup> ORR (2010) <sup>3</sup>	Top-down/ International benchmarking and updating	<ul style="list-style-type: none"> <li>· The <i>PR08</i> found a gap of 40% against the frontier for the preferred model (1996-2006).</li> <li>· The updating of the <i>PR08</i> found a gap between 34% and 40% compared to the European infrastructure managers of the peer group (1996-2008).</li> </ul>
<i>PR13</i>	ORR (2013b) <sup>3</sup>	Top-down/ International benchmarking	<ul style="list-style-type: none"> <li>· Based on the approach used in <i>PR08</i>, incorporates recent developments in the efficiency benchmarking field and employs a range of models (1996-2010).</li> <li>· Four models were selected and the efficiency gap in 2010 was estimated in a range of 13% to 24% depending on the model.</li> </ul>

Source: based on <sup>1</sup>Kennedy and Smith (2004), <sup>2</sup>Smith (2005), and <sup>3</sup>own work.

In particular, this paper is motivated by the internal benchmarking conducted by Kennedy and Smith (2004) which analyses the technical inefficiency in European railways. The analysis is divided into the period before the Hatfield accident and after it, where the most efficient zones are identified, using *COLS* and *SFA* to assess the performance. Finally, the authors point out that the internal benchmarking contributes to assess the efficiency of Network Rail and highlight that there are potential efficiency gains by eliminating the cost differences within the company.

In addition, the updating of the *PR08* and the *PR13* constitutes the framework to analyse the results of the internal benchmarking conducted in the current study. The results of the *PR08*, which covers the period 1996-2008, show that Network Rail was 40% less efficient than the top European infrastructure managers, whereas the 2010 updating of the *PR08* find a gap ranging 34% to 40% (ORR, 2010). The *PR13* was developed based on the approach used in the *PR08* and taking into account the recommendations from its reviews. The efficiency gap estimated in 2010 is in the range of 13% to 24% depending on the estimated model.

### **3. METHODOLOGY**

#### **3.1. Data base**

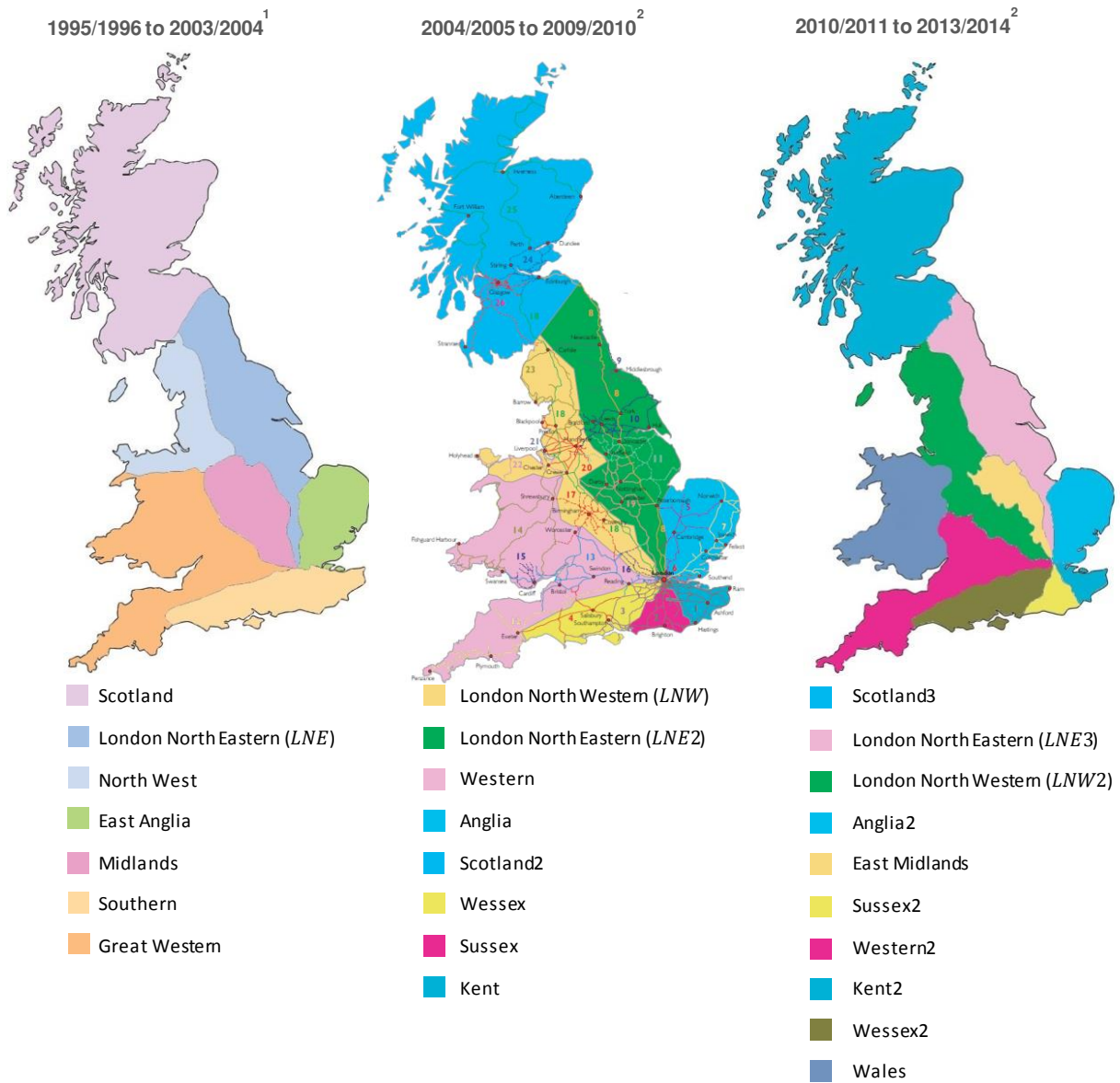
##### **3.1.1. Zone configuration and variables**

The current paper performs a panel data analysis for different organization zones of the UK rail infrastructure manager over a period of almost twenty years. The long period under analysis (1995/96-2013/14) comprehends four different zone configurations, reduced to three zones by combining Midland and Continental with London North Eastern during the period 2008/09-2010/11 (Figure 2). Because some variables are generated based on assumptions, and it is not possible to disaggregate the zones directly, this zone combination serves to simplify the analysis.

Even though, the unbalanced structure of panel data is formed by 25 regions across three different zone configurations, it is possible to aggregate the zones over time since some



regions remain almost constant. Based on the network size and the map configuration, the 25 zones are reduced to 16 aggregated zones (Table 2).



**Figure 2: Configuration of zones considered in the analysis and nomenclature**

<sup>1</sup>Source: Kennedy and Smith (2004) and Annual Return to the Rail Regulator

<sup>2</sup>Source: Annual Return – Network Rail

**Table 2: Aggregated zones**

Number of aggregated region	Regions
1	East Anglia, Anglia and Anglia2
2	Great Western and Western
3	Western2
4	Wales
5	London North Eastern and London North Eastern3
6	London North Eastern2
7	Midlands
8	East Midlands
9	North Western
10	London North Western
11	London North Western2
12	Scotland
13	Southern
14	Wessex and Wessex2
15	Sussex and Sussex2
16	Kent and Kent2

Source: own work

The dataset includes cost variables, quality data and final outputs. In addition, a set of network characteristics assumed time-invariant are considered, which contribute to control for heterogeneity. A summary of the variables is presented in Table 3 and average values by configuration zone in Table 4. It is important to highlight that a wide range of assumptions were taken into consideration due to the lack of public data in some periods, especially in the case of passenger and freight miles between 2002/03 and 2009/10 except for 2005/06.

**Table 3: Dataset variables**

Variables	Sources
<b>Costs</b>	
· Maintenance costs ( <i>MAIN</i> )	· 1995/96-2001/02: Dataset Kennedy and Smith (2004) · 2002/03-2009/10: Annual Returns · 2010/11-2013/14: Regulatory Financial Statements
<b>Quality</b>	
· Infrastructure-manager attributed delays ( <i>DELAY</i> )	· 1995/96-2001/02: Data Base Kennedy and Smith (2004) · 2002/03-2009/10: Annual Returns · 2010/11-2013/14: National Rail Trends Portal, ORR
· Broken rails ( <i>BRAIL</i> )	· 1995/96-2001/02: Dataset Kennedy and Smith (2004) · 2002/03-2013/14: Annual Returns

Outputs	
· Passenger train miles (PTRN)	<ul style="list-style-type: none"> <li>· 1995/96-2001/02: Dataset Kennedy and Smith (2004)</li> <li>· 2002/03-2004/05: Annual Returns</li> <li>· 2005/06: Dataset Wheat (2006)</li> <li>· 2010/11-2013/14: Regulatory Financial Statements</li> </ul>
· Freight train miles (FTRN)	
· Track miles (TRACK)	<ul style="list-style-type: none"> <li>· 1995/96-2001/02: Dataset Kennedy and Smith (2004)</li> <li>· 2002/03-2009/10: Annual Returns</li> <li>· 2010/11-2011/12: National Rail Trends Portal, ORR</li> <li>· 2012/2013-2013/14: Annual Returns</li> </ul>
· Freight gross tonne miles (FGTM)	<ul style="list-style-type: none"> <li>· 1995/96-2001/02: Dataset Kennedy and Smith (2004)</li> <li>· 2003/04-2004/05: Annual Returns</li> <li>· 2005/06-2013/14: National Rail Trends Portal, ORR</li> </ul>
Network characteristics	
<ul style="list-style-type: none"> <li>· Proportion of track miles with ages greater than 30 years (AGE_G30)</li> <li>· Proportion of track length which is Continuously Welded Rail (CWR)</li> <li>· Proportion of track length with maximum linespeed greater than 100mph (LP_G100)</li> <li>· Proportion of track length with maximum axle load greater than 25 tonnes (AX_G25)</li> <li>· Proportion of Track Electrified (ELECTR)</li> </ul>	<ul style="list-style-type: none"> <li>· Dataset Wheat (2006)</li> </ul>

Source: own work

**Table 4: Dataset – Total average by configuration of zone periods**

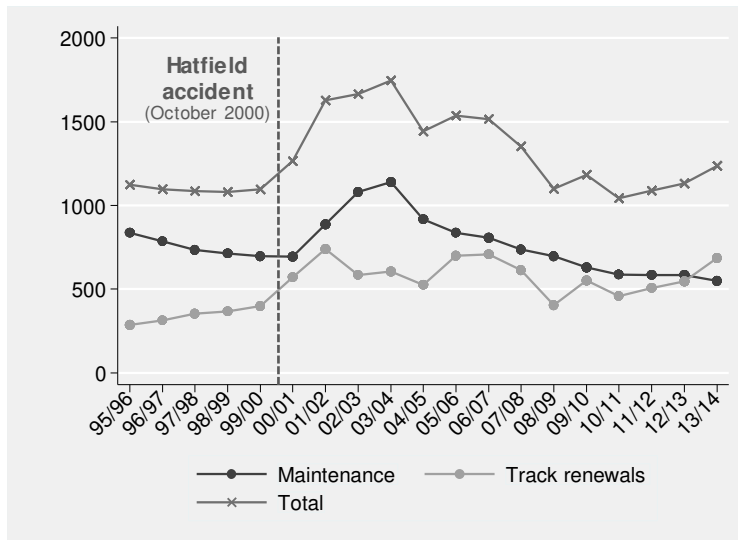
PERIOD	MAIN	REN	TOTC	DELAY	BRAIL	ALLTM	TRACK
1995/96 to 2003/04	841.0	469.1	1,310.1	12,048	680	284	19,290
2004/05 to 2009/10	770.6	584.1	1,354.6	9,774	221	305	19,319
2009/10 to 2013/14	576.3	549.6	1,126.0	8,861	151	334	19,315

Source: own work

### 3.1.2. Rail trends

Following Makovsek et al. (2015), it is possible to identify three main periods in the evolution of the infrastructure spending in the UK (Figure 3). The first period, after privatisation and before Hatfield accident, was characterised by a decrease in *MAIN* which can be attributed to efficiency gains but also to less maintenance by subcontractors with fixed-costs contracts. The second period, after Hatfield accident until 2004/05, was the period of the collapse of Railtrack in October 2000 and the creation of Network Rail in 2002, characterised by increasing costs as a result of Network Rail efforts to manage the poor levels of maintenance and investment of Railtrack; in addition, Network Rail's cost inefficiency was affected by the temporary suspension of regulatory controls. Finally, the last period, starting in 2004/05, is characterised by Network Rail improvements in efficiency

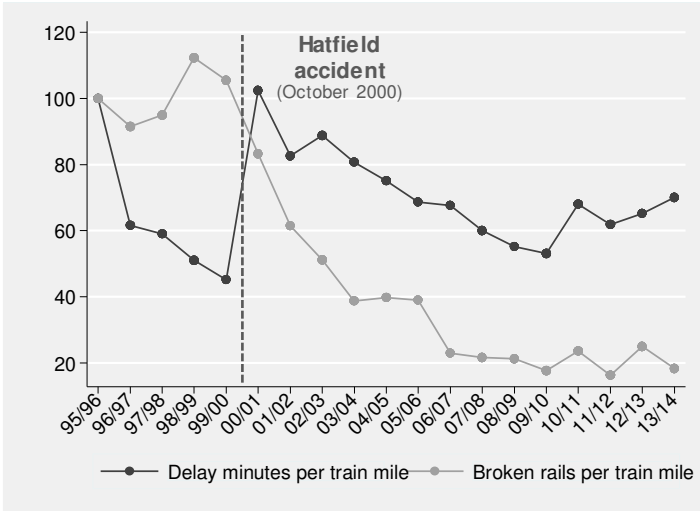
under regulatory pressures. However, the international comparisons conducted in the last periodic review found an efficiency gap of 13%-24%. McNulty (2011) observed different reasons that explain the higher costs, mainly focused on fragmentation of structures, misalignment of incentives, lack of encouragement for cost reduction in the franchising system, management procedures which fail in the implementation of best practices, and lack of partnership capability.



**Figure 3: Maintenance, track renewal and total costs**

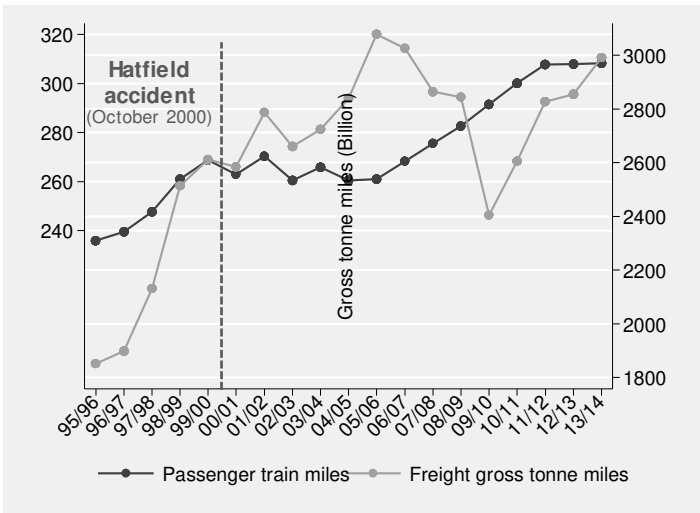
Source: own work

The structural change after Hatfield accident identified in the infrastructure spending can be slightly appreciated in the evolution of the quality indicators of delay minutes, which increased sharply in the year 1999/00-2000/01, and number of broken rails per train mile which also decreased sharply in the same period (Figure 4). Regarding *PTM* and *FGTM*, there is a decrease in both measures during the period 1999/00-2000/01 which is followed by an increase in the next year (Figure 5). In the case of *PTM*, a smooth trend can be noticed in the period 2006/07-2009/10 as a result of the assumptions made to resolve the lack of data. Finally, the unit costs by train mile also evidence the structural break after the Hatfield accident, returning the decreasing path after 2003/04 (Figure 6).



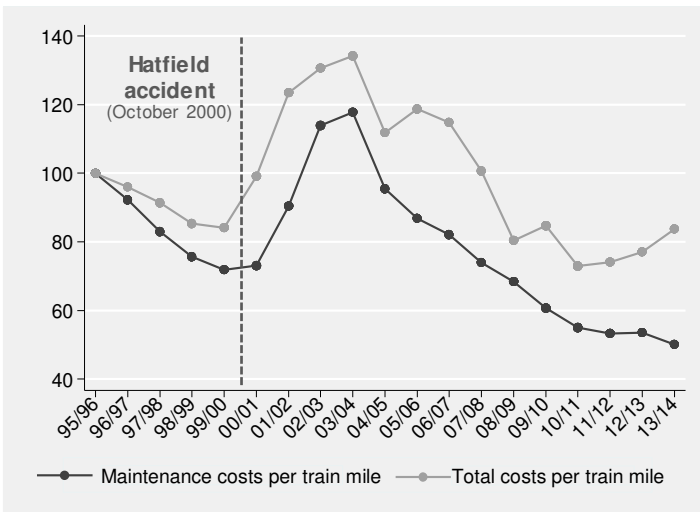
**Figure 4: Delay and broken rails per train mile**

Source: own work



**Figure 5: Passenger train miles and Freight gross tonne miles**

Source: own work



**Figure 6: Unit costs by train mile**

Source: own work

### 3.2. Internal benchmarking

The regions' performance analysis is based on frontier analysis, building an "efficiency frontier" in which the regions on it are efficient and the inefficiency of the rest of the zones is measured by the distance from the frontier (Smith et al., 2008), which represents the potential efficiency "catch-up" that inefficient regions can achieve.

Two main approaches are considered to analyse the performance of the different zones, a deterministic frontier and a stochastic frontier approach. The deterministic approach does not differentiate between noise and inefficiency when the distance to the frontier is analysed, while the stochastic frontier does. Although the data does not need to be balanced, in order to simplify the notation a balanced panel is considered.

#### 3.2.1. Cost functional form and specification details

A cost function, based on cost minimization assumptions, explain costs ( $C_{it}$ ) as a function of input prices ( $w_{it}$ ), outputs ( $y_{it}$ ), exogenous network characteristics ( $N_{it}$ ), and technical progress ( $t$ ), which indicates the cost reduction as a consequence of technical change (Smith, 2012):

$$C_{it} = f(y_{it}, w_{it}, N_{it}, t; \beta) \quad (1)$$

where  $\beta$  represents a vector of parameters. The cost function describes the minimum expenditure required to produce a given output  $y$ . The errors in optimization, technical or allocative, are reflected in a higher cost and the deviations from the frontier can be interpreted as technical and allocative inefficiency (Greene, 2008).

Different functional forms can be specified for the cost function; however, Cobb-Douglas and Translog forms are the most common in *SFA* (Fried et al., 2008b). Following the usual practices in the literature (Smith, 2012), a Translog form is estimated and tested if collapses into a Cobb-Douglas model. The statistical comparison among the Cobb-Douglas and Translog models is done considering the likelihood ratio (*LR*) test, due to the former model being nested in the latter model. In addition, a simplified version of Translog form is considered by including only the squared terms for output variables. The variables are expressed in logarithms which ensure that the estimated coefficients of the regression are cost elasticities.

Regarding the specification of the variables, in previous efficiency studies, *TOTC* has been considered as the main dependent variable (such as in the case of Kennedy and Smith (2004)), whereas, in the context of an international comparison, considering both costs together solves the problem of risk of different cost allocation among countries (Smith et al., 2008) and it was the approach used in *PR08* and *PR13*. In addition, a steady-state adjustment was made to consider the significant increase in renewal expenditure faced by Network Rail as a consequence of the underinvestment levels of Railtrack in the period before the Hatfield accident. However, the information of track renewed by zone -necessary to get the final steady state adjustment-, is not available before 2000/01<sup>6</sup> which is the crucial period regarding underinvestment made by Railtrack. Therefore, in view of the limited information available and the lumpy behaviour of renewal costs over the period, the analysis of this paper is focused on *MAIN*.

On the other hand, volume variables were expressed as densities dividing them by track miles (*PTMD*, *FGTMD*), and *BRAIL* and *DELAY* were divided by train miles to normalise by the size of the zones ( $BRAIL_{Mile}$ ,  $DELAY_{Mile}$ ). Finally, in order to model a technical change and a potential different trajectory after the Hatfield accident, a time trend, a dummy variable for the Hatfield accident and an interaction of them are included in the cost function:

$$\gamma_1 D_{Hatfield} + \gamma_2 (1 - D_{Hatfield}) Time_t + \gamma_3 D_{Hatfield} Time_t \quad (2)$$

Where, following Kennedy and Smith (2004),  $D_{Hatfield}$  is a dummy variable to reflect the Hatfield accident, which takes the value of 0.5 in 2000/01 in the case of *TOTC* (the Hatfield accident happened in the middle of the financial year) and unity onwards, and takes the value of 1 in 2001/02 in the case of *MAIN* and unity onwards, because the sharp increase did not start immediately after the accident (Figure 1). The objective of  $\gamma_2$  and  $\gamma_3$  is to capture a different technical change before and after the Hatfield accident, being the difference between both statistically tested.

### 3.2.2. Deterministic frontier approach

*COLS*, proposed by Winsten (1957), is the simplest deterministic frontier model. It corrects the estimation of Ordinary Least Squares (*OLS*) by shifting the line down by the amount of

---

<sup>6</sup> The information for the rest of the years is available on Annual Returns.

the minimum *OLS* residual ( $\min_i \{\hat{u}_i\}$ ) generating a cost frontier which is in or under the data (Parmeter and Kumbhakar, 2014):

$$\hat{\beta}_0^{COLS} = \hat{\beta}_0 + \min_i \{\hat{u}_i\} \quad (3)$$

$$\hat{u}_i^{COLS} = \hat{u}_i - \min_i \{\hat{u}_i\} \geq 0 \quad (4)$$

and the cost efficiency of each unit ( $CE_i$ ), which is the ratio of the minimum cost to actual cost, is calculated as follows:

$$CE_i = \{-\hat{u}_i^{COLS}\} \quad (5)$$

The problem of *COLS* is that all the deviations from the frontier are attributed to inefficiency and does not allow random noise, making the estimation highly sensitive to extreme values. Even considering this strong assumption, this approach is employed in the UK for its simplicity (Nash and Smith, 2014).

Finally, as pointed out earlier, *COLS* does not take into account noise existence. Therefore, a noise-adjustment is applied to the efficiency scores. This approach implies that 25% of the inefficiency is attributed to random events, and it is recognising that the split between inefficiency and noise is unknown:

$$x_{i,t}^* = x_{i,t} + 0.25 \times (1 - x_{i,t}) \quad (6)$$

where  $x_{i,t}^*$  is the adjusted efficiency score and  $x_{i,t}$  is the original efficiency score.

### 3.2.3. Stochastic frontier approach with panel data

The Stochastic Frontier Analysis (*SFA*), developed simultaneously by Aigner, et al. (1977) and Meeusen and van den Broeck (1977), identifies a deterministic component of the frontier and a stochastic component allowing the effect of random external factors.

$$C_{it} = \underbrace{f(y_{it}, w_{it}, N_{it}, t_t; \beta)}_{\text{Deterministic Frontier}} + \underbrace{v_{it}}_{\text{Noise}} + \underbrace{u_{it}}_{\text{Inefficiency}} \quad (7)$$

Stochastic Frontier

Panel data structure allows capturing some heterogeneity, impossible to capture with cross sectional data. However, the heterogeneity can arise from inefficiency or specific heterogeneity of the firms, which is a-priori unknown (Parmeter and Kumbhakar, 2014).

Following Parmeter and Kumbhakar (2014) and Cuesta (2000), the *SFA* specifications with panel data can be classified according to the behaviour of the inefficiency over time and the



methods of estimation (Table 5). Firstly, time-invariant inefficiency models imply that the technical inefficiency is the same over time, whereas time-variant inefficiency models imply that the inefficiency can vary over time. Secondly, the models can be classified according to if they are specified following traditional panel data techniques or if they are estimated using Maximum Likelihood (ML) techniques. However, not all models are suitable for the case under analysis. Firstly, the long period of the panel data make it implausible to consider time-invariant inefficiency, especially in light of the impact of the Hatfield accident and the background behind it, and for the regions covering the whole period. Secondly, the assumption of the same efficiency pattern for all the regions does not seem to be appropriate, in particular considering the changes in zone configuration. Therefore, a Cuesta (2000) model and its variations are considered.

**Table 5: SFA specifications with panel data<sup>1</sup>**

	<b>Time-invariant technical inefficiency models</b>	<b>Time-varying technical inefficiency models</b>
<b>Traditional Panel data methods</b>	$y_{it} = \alpha_i + X'_{it}\beta + v_{it}$ $\alpha_i = \beta_0 + u_i$ <ul style="list-style-type: none"> <li>· Schmidt and Sickles (1984) <math display="block">\hat{u}_i = \hat{\alpha}_i - \min_i\{\hat{\alpha}_i\} \geq 0</math> </li> <li>· Other approaches: Seale (1990), Atkinson and Cornwell (1994) <math display="block">CE_i = \{-\hat{u}_i\}</math> </li> </ul>	$y_{it} = \alpha_{it} + X'_{it}\beta + v_{it}$ <ul style="list-style-type: none"> <li>· Cornwell, Schmidt, and Sickles (1990): <math display="block">\alpha_{it} = \alpha_{0i} + \alpha_{1i}t + \alpha_{2i}t^2</math> <math display="block">\hat{u}_{it} = \hat{\alpha}_{it} - \min_j\{\hat{\alpha}_{jt}\}</math> </li> <li>· Lee and Schmidt (1993): <math display="block">u_{it} = u_i\lambda_t</math> <math display="block">\hat{u}_{it} = \{\hat{u}_i\lambda_t\} - \min_i\{\hat{u}_i\lambda_t\}</math> <math display="block">CE_i = \{-\hat{u}_i\}</math> </li> </ul>
<b>Maximum Likelihood methods</b>	$y_{it} = \alpha_i + X'_{it}\beta + v_{it}$ $\alpha_i = \beta_0 + u_i$ <ul style="list-style-type: none"> <li>· Pit and Lee (1981) <p><i>Firm specific inefficiency → JLMS type mean estimators estimated from the mean <math>E(u_i/\varepsilon_i)</math> or the mode <math>M(u_i/\varepsilon_i)</math> using ML, or alternative predictor of Battese and Coelli (1988).</i></p> </li> </ul>	$y_{it} = X'_{it}\beta + v_{it} - u_{it}$ $u_{it} = G(t)u_i \text{ and } G(t) > 0$ <p>Different specifications of <math>G(t)</math>:</p> <ul style="list-style-type: none"> <li>· Kumbhakar (1990): <math display="block">G(t) = [1 + \exp(\gamma_1 t + \gamma_2 t^2)]^{-1}</math> </li> <li>· Battese and Coelli (1992): <math display="block">G(t) = \exp[\gamma(t - T)]</math> </li> <li>· Cuesta (2000): <math display="block">G(t) = \exp[\eta_i(T - t)]</math> </li> <li>· Cuesta (2000) extension: <math display="block">G(t) = \exp[\eta_{i1}(T - t) + \eta_{i2}(T - t)^2]</math> <p><i>Firm specific inefficiency → estimated from the mean <math>E(u_i/\varepsilon_i)</math> or the mode <math>M(u_i/\varepsilon_i)</math> using ML, or alternative predictor of Battese and Coelli (1988).</i></p> </li> </ul>

Source: based on Parmeter and Kumbhakar (2014), Cuesta (2000) and own work.

<sup>1</sup> Recently developments have included methodologies to capture unobserved heterogeneity, disentangling inefficiency from heterogeneity, such as for example Greene (2005) and Farsi et al. (2005). Other extensions of the stochastic approach include dynamic panel data models, allowing, for example, for serial correlation between technical inefficiency levels and firm-specific patterns, such as in Ahn and Sickles (2000).

### 3.3. International context

The results of the internal benchmarking are analysed in the context of the international benchmarking conducted by *ORR*. It is expected that results of internal benchmarking show less possibilities of efficiency gains than results of external benchmarking (Smith et al., 2010). As indicated previously, *ORR* controls Network Rail through Periodic Reviews. The 2008 Periodic Review (*PR08*) considers the Control Period 4 (*CP4*) which covers the period April 2009 - March 2014; whereas the 2013 Periodic Review (*PR13*) considers the Control Period 5 (*CP5*) which covers the period April 2014 - March 2019. The results of these Periodic Reviews are compared with those obtained in the internal benchmarking.

## 4. RESULTS

### 4.1. Internal benchmarking

#### 4.1.1. Cost functional form

The models were estimated using *MAIN* as a dependent variable and focusing the analysis on it, as explained before. Firstly, both a translog function and a reduced version of it, by considering only squared terms for output variables, were tested. However, the obtained results were not plausible, and a Cobb-Douglas function was employed although the former models were statistically preferred to the latter. Secondly, three set of output variables were tested, *PTMD* and *FGTMD* (set 1), *PTMD* and *FTRND* (set 2) and *ALLTMD* (set 3). Even though the first set was preferred because it allows identifying the effects of passenger and freight services, its results were against the empirical evidence of economies of density. Therefore, *ALLTMD* (set 3) was considered. Finally, the network characteristic variables were excluded from the model because less strong economies of density were obtained. The cost functional form was defined as follow:

$$\ln(Cost_{it}) = \beta_0 + \beta_1 \ln(ALLTMD_{it}) + \beta_2 \ln(TRACK_{it}) + \beta_3 \ln(DELAY_{Mile_{it}}) + \beta_4 \ln(BRAIL_{Mile_{it}}) + \gamma_1 D_{Hatfield} + \gamma_2 (1 - D_{Hatfield}) Time_t + \gamma_3 D_{Hatfield} Time_t \quad (8)$$

where  $Cost_{it}$  corresponds to  $MAIN$ . The function allows to capture economies of density (the cost implication of increasing train density while keeping the network size constant), and economies of scale (the cost implication of increasing network size while holding constant train density), through the coefficients of  $ALLTMD$  and  $TRACK$  respectively.

The exclusion of network characteristics results in slightly more plausible estimations in line with the literature in the field. However, the cost elasticity with respect to traffic remained higher. This model represents a big issue from an economic perspective because the inclusion of network characteristics constitutes a source of heterogeneity across firms that should be taken into account in order to assess efficiency. However, the weaker economies of densities conflict with the inclusion of the network variables, and the decision of excluding them is taken.

In order to analyse if the efficiency results are significantly sensitive to the network variables, Spearman's rank was used. It is appropriate to compare the differences among rankings (not the actual values)<sup>7</sup>. The ranking of the reduced model (without network characteristics) estimated using  $COLS$  is compared with the ranking obtained by estimating the complete model (including network characteristics). The results show a  $\rho_s = 0.9550$  which is statistically significant at the 1% level of significance ( $\alpha = 0.01$ ) implying a highly strong positive relationship among rankings.

#### **4.1.2. Deterministic frontier approach**

The results of the estimated  $COLS$  for  $MAIN$  are presented in Table 6. Because the estimated model is log-log, the coefficients represent the elasticity of cost with respect to the cost driver under analysis. The output variable presents the expected sign and magnitude, and it is statistically significant. An increase in volume densities ( $ALLTMD$ ), leads to an increase in costs<sup>8</sup>. The result shows economies of density in line with the empirical evidence. However, this result is significantly less strong than other empirical results, such as the  $PR13$  which found a coefficient around 0.6 for train density for total cost models (ORR, 2013b). In fact, the analysis focused on railway infrastructure finds cost elasticities with respect to traffic around 0.49 in the case of  $TOTC$  and in the range of 0.20-0.35 if only

---

<sup>7</sup> It has been widely used in the railway literature (for instance Oum and Yu (1994) and Cantos et al. (2012)).

<sup>8</sup> A 1% increase in  $ALLTMD$  leads to a 0.777% increase in  $MAIN$ .

*MAIN* are considered (Smith, 2012), such as the case under analysis. Regarding track miles (*TRACK*), the  $H_0$  of returns to scale cannot be rejected, also in line with the evidence in the field which finds constant returns to scale in vertically integrated railways (Smith, 2012)<sup>9</sup>.

On the other hand, the quality variables under analysis present different sign. An increase in  $DELAY_{Mile}$  might increase *MAIN*<sup>10</sup>. Conversely, an increase in  $BRAIL_{Mile}$  leads to a decrease of -0.046% in *MAIN*. This impact is statistically significant at  $\alpha = 0.10$ .

Finally, the estimated time trend coefficients allow identifying different patterns of technical change before and after Hatfield Accident. Annual cost reductions of -5.6% and -6.7%<sup>11</sup> are achieved as a consequence of technical change during the period before and after Hatfield accident respectively. Where the positive coefficients of  $D_{Hatfield}$  evidence a sharp increase in costs of 72%<sup>12</sup> for *MAIN* from the accident onwards<sup>13</sup>. Finally, the *Adjusted R*<sup>2</sup> of 0.939 evidences a very good fit of the model.

**Table 6: COLS Regression**

	Maintenance	
	Coef.	Std. Err.
$\ln(ALLTMD)$	0.777***	(0.055)
$\ln(TRACK)$	0.948***	(0.027)
$\ln(DELAY_{Mile})$	0.195***	(0.042)
$\ln(BRAIL_{Mile})$	-0.046*	(0.028)
$D_{Hatfield}$	0.542***	(0.073)
$Time_{BeforeHatfield}$	-0.058***	(0.013)
$Time_{AfterHatfield}$	-0.069***	(0.006)
Constant	3.798***	(0.741)
<i>Adjusted R</i> <sup>2</sup>	0.939	

Source: own work

The efficiency scores by zone in the case of *COLS* were calculated by considering a noise-adjustment, where 25% of the inefficiency was attributed to noise, as described previously.

<sup>9</sup> A 1% increase in *TRACK* might increase *MAIN* by 0.948%.

<sup>10</sup> A 1% increase in  $DELAY_{Mile}$  might increase *MAIN* by 0.195%.

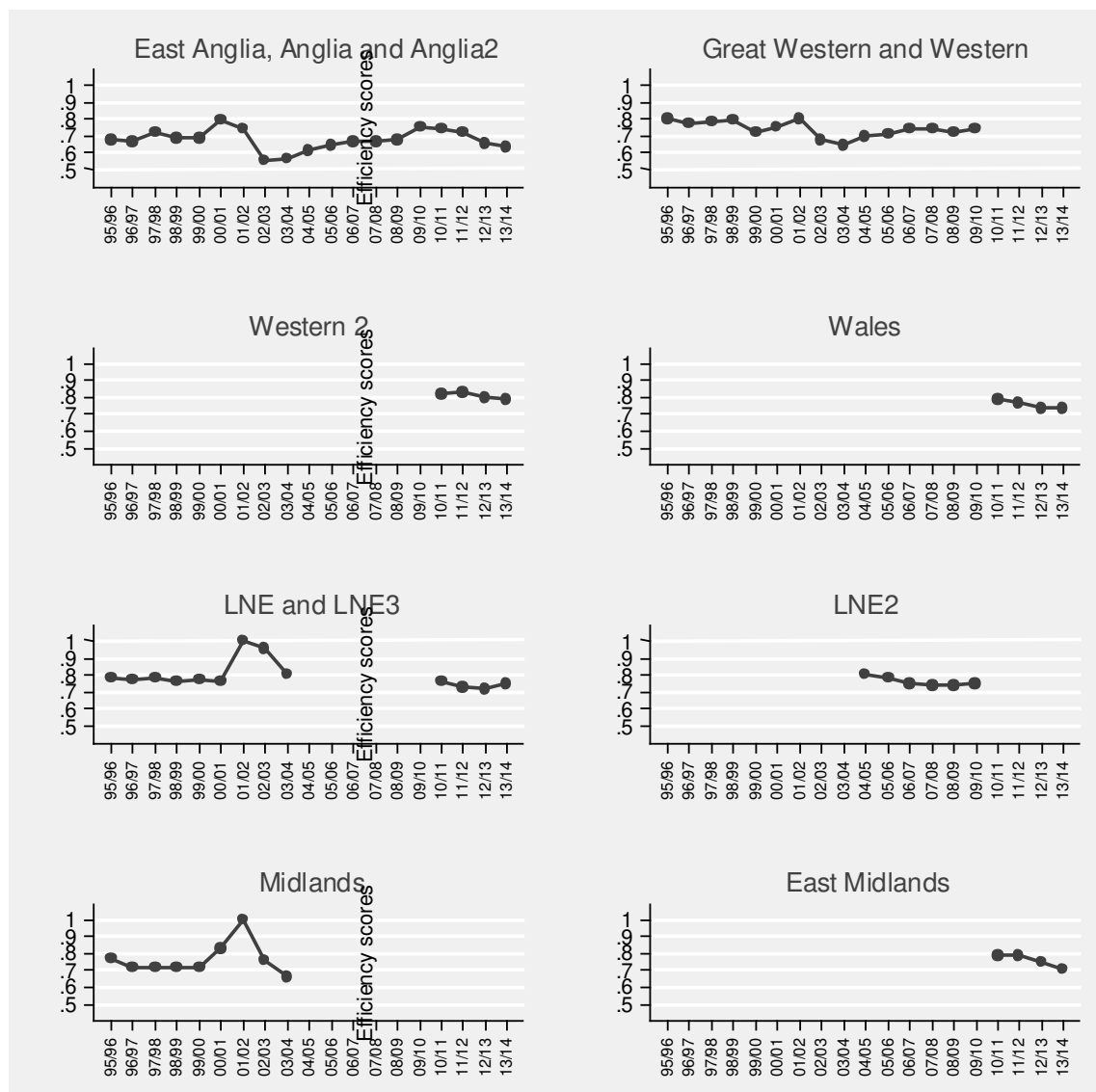
<sup>11</sup>  $100\% \times [\exp(-0.058) - 1] \times 100 = -5.6\%$ ;  $100\% \times [\exp(-0.069) - 1] = -6.7\%$ .

<sup>12</sup>  $100\% \times [\exp(0.542) - 1] = 72\%$ .

<sup>13</sup> The wide range of models estimated, including different sets of output variables as well as network characteristics, report similar results.

The results of the analysis show that the most efficient region is *LNE* in the 2001/02 financial year where the efficiency score is equal to 1. The rest of the efficiency scores were calculated considering this region and year as the best performer.

Figure 7 and Figure 8 present the evolution of the efficiency scores over time by aggregated zone. Hatfield accident represents a significant decrease in the efficiency performance across zones, generally starting in 2001/02 and finalising in 2003/04. In the regions with more years of information, generally, it is possible to identify a positive time trend in the efficiency until 2010/11 and a decreasing trend after that. Even though the preferred cost functional form does not include network characteristic variables, *COLS* was tested with the complete model with network characteristics and the same pattern was reproduced.



**Figure 7: COLS - Efficiency scores by zones over time (1)**

Source: own work

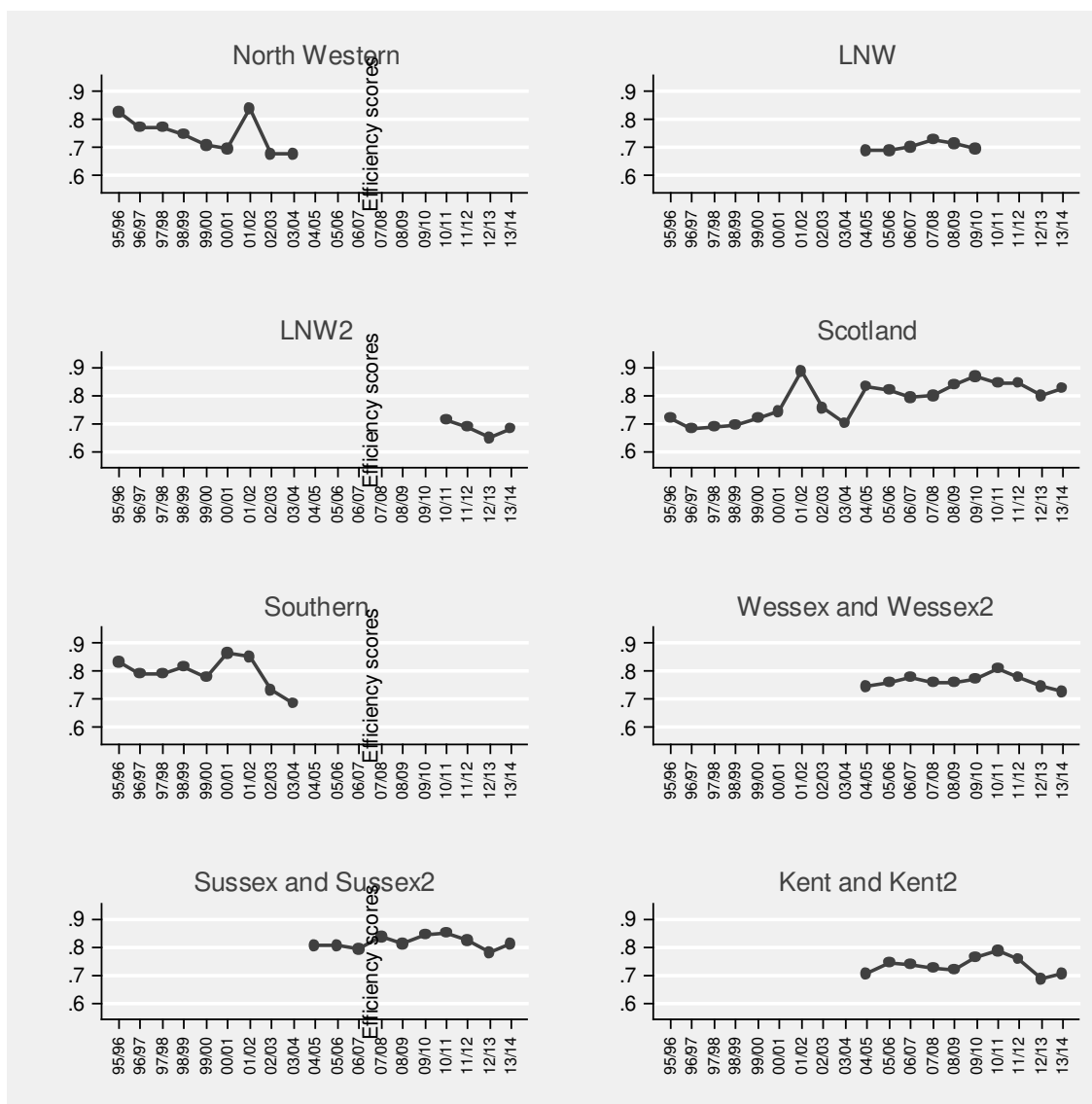


Figure 8: COLS - Efficiency scores by zones over time (2)

Source: own work

#### 4.1.3. Stochastic frontier approach with panel data

The structure of panel data was used to estimate the stochastic frontier models. Different models were estimated<sup>14</sup> assuming different patterns for the inefficiency regarding variation over time and estimation methods, as previously discussed. A half normal distribution was assumed for all the estimated models. As indicated above, not all models are suitable for the case under analysis, and *CUES00* is the preferred one to assess efficiency because allows for firm-specific patterns of temporal change of  $u_{it}$ .

<sup>14</sup> The models were estimated using *sfp* command created by Belotti et al. (2012), with the exception of Cuesta (2000) model. This model was estimated based on an adjustment of *sfp* command created by Kumbhakar et al. (2014).

A wide range of versions of *CUES00* were tested considering different specifications such as including time-squared terms, excluding time terms extremely implausible ( $|\eta_i| > 1$ ,  $|\eta_i^2| > 1$ ) and excluding time terms not statistically significant. The simple version of Cuesta (2000) includes all time terms, *CUES00(1)* model. All the models were compared using *LR* test, and the preferred one was the model which includes all time terms and only the significant time-squared terms. Based on this model, additional modifications were tested, and the *CUES00(2)* model considering only one time trend<sup>15</sup> and excluding *BRAIL<sub>Mile</sub>*, was preferred.

In addition, a sensitivity analysis was conducted. The objective of this analysis was to evaluate which were the impacts on the efficiency path of different assumptions about the specification of the model and the data base, especially in the last years of the period (2010/11 to 2013/14). The analysis was conducted based on the Cuesta (2000) model which includes all time terms and only the significant time-squared terms. The specification of the model sensitivities covered the inclusion of network characteristics, a squared time trend for all the period and a squared time trend for the period after the Hatfield accident. On the other hand, the database sensitivities included changes in the main database assumptions, and exclusions of the zones from the database one by one. None of the sensitivities generated a significant change in the pattern of the average efficiency regarding the original models, with the exception of excluding Scotland, *CUES00(3)* model. Additionally, this model excludes *BRAIL* because it is not significant, and it keeps both time trends for the period before and after the Hatfield accident, because the difference between both of them is statistically significant.

The results of these frontier models are presented in Table 7. The economies of density are stronger in the simple version *CUES00(1)*. The elasticity of cost with respect to traffic is 0.693 in *CUES00(2)* model. Even though this result is in line with economies of density, the cost elasticity is significantly higher than that resulting from the empirical evidence, as highlighted before. On the other hand, the regions evidence a positive technical change of 6.3%<sup>16</sup> per year according to *CUES00(2)*.

---

<sup>15</sup> There was no statistically difference between the time trends before and after the Hatfield accident.

<sup>16</sup>  $100\% \times [\exp(-0.065) - 1] = -6.3\%$ .

**Table 7: CUES00 and its variations**

Variable	<i>CUES00</i> (1)	<i>CUES00</i> (2)	<i>CUES00</i> (3) (Excluding Scotland)
$\ln(ALLTMD)$	0.604***	0.693***	0.703***
$\ln(TRACK)$	0.841***	0.861***	1.022***
$\ln(DELAY_{Mile})$	0.086**	0.101***	0.102***
$\ln(BRAIL_{Mile})$	-0.031		
$D_{Hatfield}$	0.602***	0.411***	0.204**
$Time_{pre-Hatfield}$	-0.061***		-0.074***
$Time_{post-Hatfield}$	-0.081***		-0.045***
$Time$		-0.065***	
$Constant$	6.086***	5.538***	4.160***
$\eta_1$	-0.048**	0.184**	0.293***
$\eta_2$	-0.107**	0.357**	0.600***
$\eta_3$	-0.624	-0.423	-3.347
$\eta_4$	-1.231	-0.596	-0.979
$\eta_5$	-0.289**	-0.232*	-1.621
$\eta_6$	-0.127	-0.076	-0.062
$\eta_7$	-0.1	-0.851**	-0.654**
$\eta_8$	-1.122	-1.377	-1.023
$\eta_9$	-0.514	-0.380**	-0.272***
$\eta_{10}$	-0.002	0.03	0.119
$\eta_{11}$	-0.084	-0.085	-0.116
$\eta_{12}$	-1.553	0.108*	
$\eta_{13}$	-0.796*	-0.679**	-0.527**
$\eta_{14}$	-0.251	-0.072	0.128
$\eta_{15}$	-0.803	-1.558	0.157*
$\eta_{16}$	-0.138	-0.041	0.069
$\eta_1^2$		-0.013***	-0.016***
$\eta_2^2$		-0.040***	-0.055***
$\eta_7^2$		0.093**	0.067**
	* p<0.1; ** p<0.05; *** p<0.01		

Source: own work



The estimated parameters reflecting the temporal variation of technical inefficiency ( $\eta_i$ ) show that, in the case of *CUES00(2)* model, 9 parameters are not statistically different from zero and the  $H_0$  of time-invariant inefficiency cannot be rejected for them; whereas from the remaining 7 regions, 2 of them experience a concave-up inefficiency pattern ( $\eta_{1,2} > 0$ ,  $\eta_{1,2}^2 < 0$ ), one of them experiences a concave-down inefficiency pattern ( $\eta_7 < 0$ ,  $\eta_7^2 > 0$ ), and the rest a decreasing inefficiency pattern over the period ( $\eta_i < 0$ ). All the squared-terms included in the model, which allows for turning points in the inefficiency path, are statistically different from zero at  $\alpha = 0.05$ . The same results are appreciated in the case of *CUES00(3)*, except that regions 14, 15 and 16 present an increasing inefficiency pattern over the period ( $\eta_{14,15,16} > 0$ ) when Scotland is excluded from the database. The ranking of regions changes every year due to the flexibility of the model. The following graphs (Figure 9 and Figure 10) show the performance evolution of each region over time for *CUES00(1)*, *CUES00(2)* and *CUES00(3)* models, where the described firm-specific patterns mentioned before can be appreciated. Finally, the ranking of the regions according to the preferred *CUES00(2)* model for 2013 shows that Scotland is the best performing region, Sussex2 the second one, and Wales the third one.

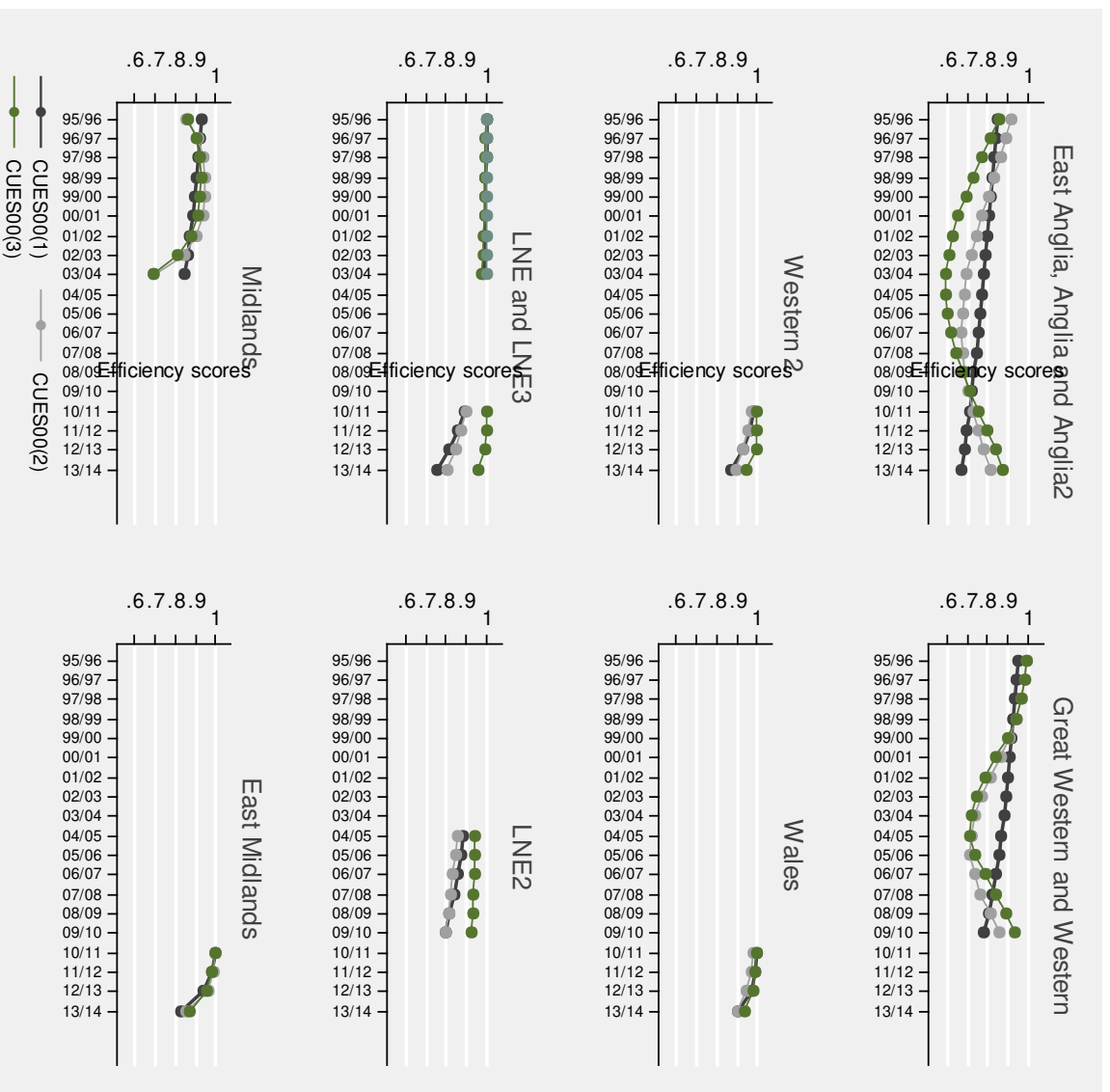


Figure 9: Preferred stochastic frontier models - Efficiency scores by zones over time (1)  
Source: own work

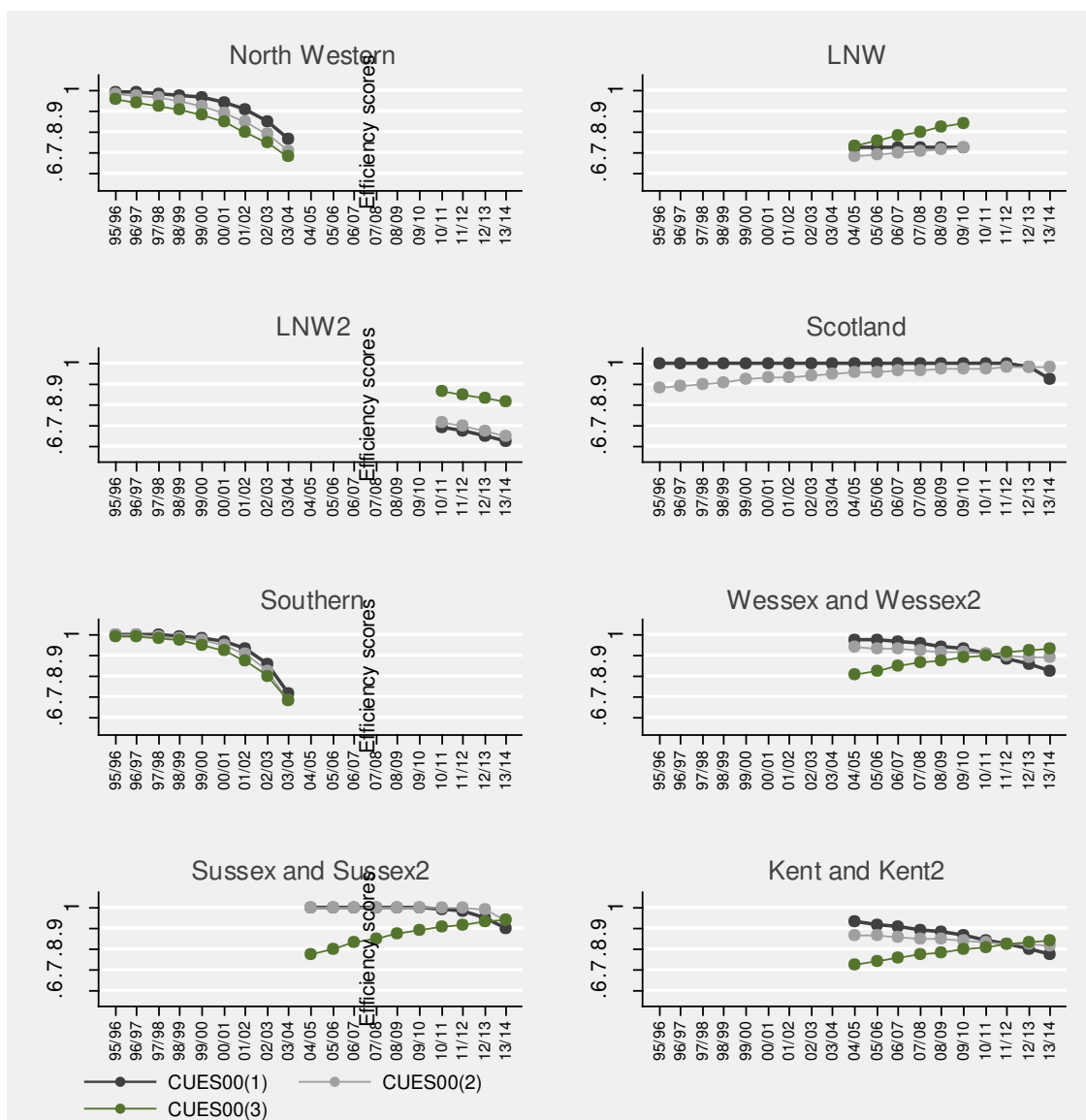


Figure 10: Preferred stochastic frontier models - Efficiency scores by zones over time (2)

Source: own work

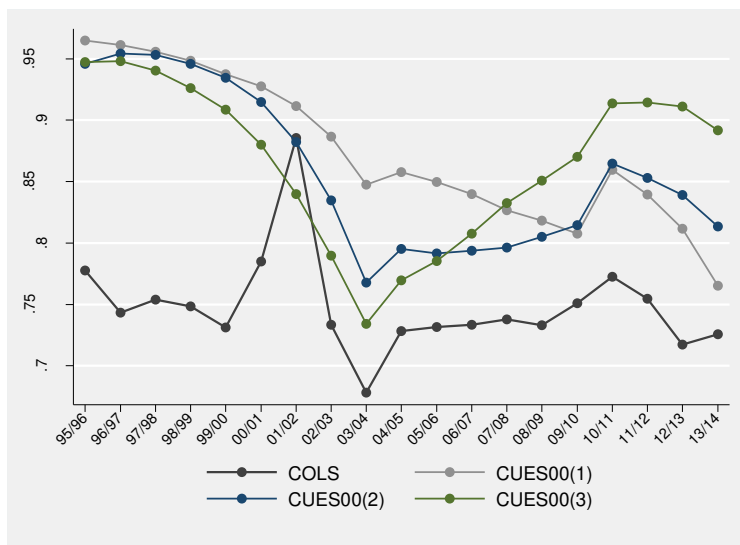
#### 4.1.4. Comparison of results

A cost-weighted average<sup>17</sup> of the efficiency was estimated to analyse the overall performance and to control by a relationship of the inefficiency with the region size<sup>18</sup> (Figure 11). It is possible to identify three breaks in the data set corresponding with the three zone configuration changes in 2003/04 and 2009/10. In the case of *COLS* model, a significant decrease in efficiency starts in 2001/02, whereas in the stochastic models the decrease

<sup>17</sup> The efficiency of each region was weighted by its cost over the total in each year.

<sup>18</sup> The average performance over time (without weights) follows almost exactly the same evolution than the cost-weighted efficiency.

starts from almost the beginning of the period under analysis until 2003/04. The overall performance follows a slightly positive trend until 2010/11, to decrease again, recovering a positive trend in the last year in the case of *COLS*. The results of the cost-weighted average efficiency scores in 2013/14 evidence potential inefficiency<sup>19</sup> of 26.3% (efficiency score equal to 0.737) in the case of *COLS* model, and 19.4%, 14.6% and 9.8% in the cases of *CUES00(1)*, *CUES00(2)* and *CUES00(3)* respectively (Table 8) which are analysed in the international comparison framework in the next section.



**Figure 11: Efficiency scores over time – Cost-weighted average**  
Source: own work

**Table 8: Cost-weighted average inefficiency in 2013/14**

Model	Efficiency gain
<i>COLS</i>	26.3%
<i>CUES00 (1)</i>	19.4%
<i>CUES00 (2)</i>	14.6%
<i>CUES00 (3)</i>	9.8%

Source: own work

It is important to highlight that the exclusion of Scotland from the data base generates a change in the pattern of the average efficiency in the period 2010/11-2012/13. The average efficiency remains almost constant rather than decrease as in the original case. However, the decrease in the last year under analysis (2013/14) is the same than in the other versions of *CUES00* models. Among all the changes tested in the sensitivity analysis, only the exclusion of Scotland generates a modification of the average efficiency pattern in the last years of the

<sup>19</sup>  $(1 - \text{Efficiency score})$

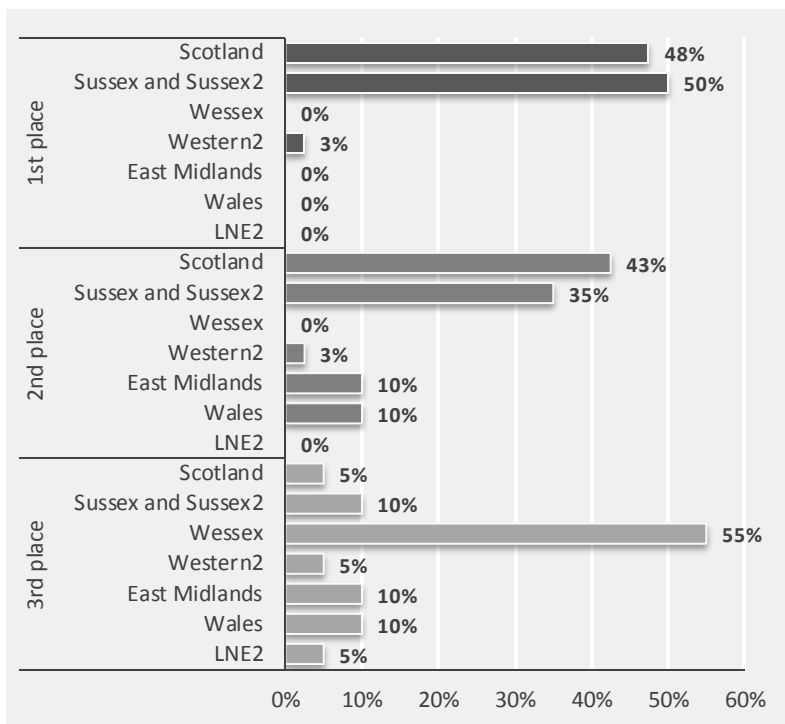
period. However, it seems that it is the behaviour of Scotland for itself that produces the pattern because the results are not significantly sensitive to the database assumptions considered about Scotland.

The results of the Spearman correlation matrix suggest a high positive correlation between *COLS* and *CUES00(1)* and *CUES00(2)*. Even though the estimated inefficiency levels are different among models, the ranking of the regions are highly correlated supporting the results of the top performing regions. In order to analyse the sensibility of the regions' performance, the rankings of the regions for *COLS* and the stochastic models based on Cuesta (2000)<sup>20</sup> for the period 2004/05-2013/14 are considered, being Scotland and Sussex and Sussex2 the best performing regions, followed by Wessex. On the other hand, Western2, LNE2, East Midlands and Wales, get sporadically the second and third places in the rankings (Figure 12).

The results of this analysis show that Scotland and Sussex and Sussex2 have been implementing the best practices among all the regions in the last periods, achieving the top positions in terms of cost efficiency. As was pointed out by Smith (2006), internal benchmarking can be an useful tool for the infrastructure manager to reduce costs by eliminating the intra-company cost differences. In this sense, Network Rail could take advantage of the information about the best performing regions, analyse their practices in order to learn about them, and develop strategies to apply in the rest of the regions.

---

<sup>20</sup> *CUES00(1)* and *CUES00(2)*.



**Figure 12: Best performing regions 2004/05 - 2013/14 – COLS and stochastic models based on Cuesta (2000), CUES00(1) and CUES00(2)**

Source: own work

## 4.2. International context

The international comparisons of Network Rail using top-down benchmarking started its implementation during the *PR08*. The dataset for the international comparison was collected as part of the Lasting Infrastructure Cost Benchmarking (*LICB*) exercise of the *UIC*, including a slightly different set of countries between *PR08* and *PR13*.

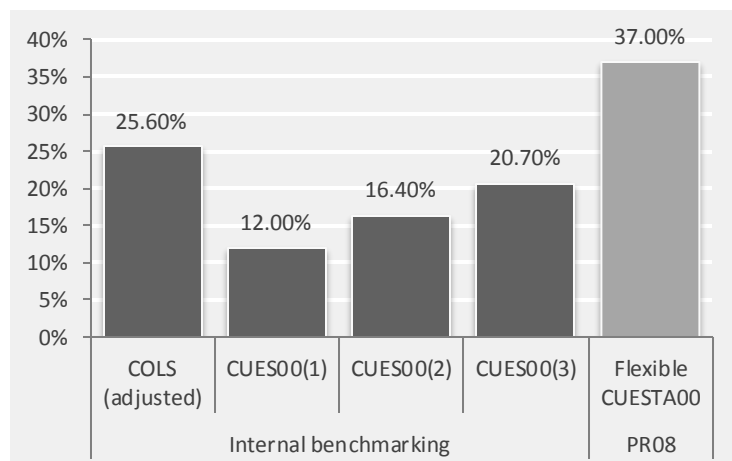
In the *PR08* a preferred model was selected, which is a flexible version of Cuesta (2000), including a squared time trend for Network Rail to capture the effect of the Hatfield accident, while in the *PR13* was decided to take the approach of selecting a set of models.

According to the preferred econometric model in the *PR08*, including steady state adjustment, an efficiency gap between Network Rail and the top quartile of the European comparators of 37% was estimated in 2006. This gap was translated to an efficiency target of 23% for maintenance and renewal costs for the end of the *CP4* (ORR, 2008). During the last periodic review, *PR13*, an efficiency gap ranging 13% to 24% was estimated depending on the model, with an estimated efficiency gap of 23% if only the inner range from these models is examined. Network Rail's *PR13* strategic business plan set out a 24% reduction on its support costs over *CP5* (ORR, 2013a). The *ORR* points out that the efficiency gap can be

explained by “differences in contracting and possessions strategy, system renewals, asset condition monitoring, renewal backlogs, workforce protection and effective network size” (Beck et al., 2013, p.18).

The results of the current internal benchmarking are placed in the international context based on the results described previously. As stated by Smith et al. (2010), the potential efficiency gains arising from internal benchmarking are lower than those arising from international comparisons. Effectively, the expected efficiency gains identified by the internal benchmarking are lower in both 2008 and 2010 estimations (Figure 13 and Figure 14 respectively). In particular, the results of the *COLS* model are closer to those obtained from the international analysis. The internal benchmarking only identifies the efficiency gains that can be achieved by moving to the best practice presented by Network Rail. If the best performing regions are inefficient in some way, the identified potential cost efficiency savings will be less than the savings that could be achieved.

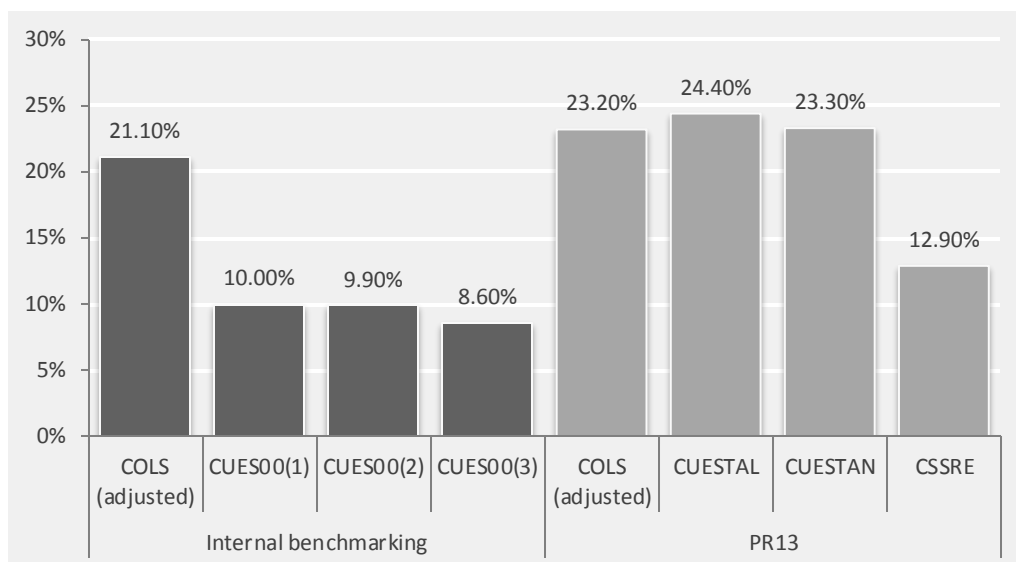
The observed differences are not only arising from the approaches (internal vs. international benchmarking). There are differences regarding the variables included, the adjustments and specifications of the models, and mostly due to the independent variables considered in both approaches (maintenance costs vs. total costs). However, the framework of the international benchmarking is helpful to verify the consistency of the estimations obtained from the internal benchmarking.



**Figure 13: Estimates of Network Rail's efficiency gap in 2008 (preferred models)**

*Internal benchmarking: cost-weighted average efficiency*

Source: Smith (2008) and own work



**Figure 14: Estimates of Network Rail's efficiency gap in 2010 (preferred models)**

*Internal benchmarking: cost-weighted average efficiency*

Source: ORR (2013b) and own work

## 5. CONCLUSIONS

The aim of this paper has been to analyse the efficiency performance of the UK rail infrastructure manager from an internal perspective. Several efficiency analysis have been commissioned by the *ORR* in the last years, including internal econometric benchmarking, but focused mainly on international econometric benchmarking by comparing Network Rail with other European comparators. However, the last internal econometric benchmarking was done by Kennedy and Smith (2004), covering only two years after the Hatfield accident. In this context, the current analysis has focused on providing an updated internal perspective to the efficiency analysis of Network Rail, focusing on the period after the Hatfield accident by identifying the top performing regions and putting the results obtained from the internal benchmarking in the international context.

A frontier analysis was conducted by building a cost efficiency frontier, where the efficient regions lie on the frontier and the inefficiency of the remaining regions is measured as the distance from it. The analysis has included a deterministic frontier approach which attributes the entirely distance from the frontier to inefficiency (*COLS*), and a stochastic frontier approach which distinguishes noise from inefficiency. Even though the final results have found evidence of economies of density with a cost elasticity with respect to traffic of



around 0.7, they are significantly weaker than those found by the literature in the case of maintenance costs which are in the range of 0.2-0.35 (Smith, 2012).

The deterministic and stochastic frontier approaches are consistent in the cost-weighted average path followed by the infrastructure manager after the Hatfield Accident, with a significant decrease until 2003/04, followed by a positive trend until 2010/11 to decrease again, returning to a positive trend in 2013/14 only in the case of *COLS*. In the case of the stochastic frontier models, Cuesta (2000) model and different versions of it were considered because it allows for region-specific time varying inefficiency. The results have allowed to identify cost efficiency gains from an internal perspective and to identify the best performing regions. The findings are consistent across *COLS* and the preferred stochastic models, identifying Sussex and Sussex2 and Scotland as the top performing regions in the period 2004/05-2013/14 occupying the first and second places, followed by Wessex. A sensitivity analysis was conducted to evaluate which were the impacts on the average efficiency path of specification of the model and data base assumptions. Only the exclusion of Scotland generates a modification of the average efficiency pattern in the last years of the period, but it seems that it is the behaviour of Scotland for itself that produces the pattern, rather than the assumptions considered about Scotland. The potential cost-weighted average efficiency gains were estimated in a range of 10% and 26% in 2013/14 depending on the estimated models. This provides Network Rail the opportunity of achieving important cost savings by understanding the differences intra-company.

These internal benchmarking findings were compared with the results obtained during the *PR08* and the *PR13*. The potential average efficiency gains obtained from the internal benchmarking for 2008 and 2013, were lower than those resulting from the international comparison for Network Rail as predicted by the theory (Smith et al., 2010). In particular, the results of the adjusted *COLS* from the internal benchmarking are closer to those arising from the international benchmarking, but still below than the international results. Even though, these results are not directly comparable because both approaches differ in the independent variable considered in the models (maintenance costs vs. total costs), as well as in the specifications and adjustments of the models, they have helped to verify the consistency of the internal benchmarking results.

Nonetheless, the results should be analysed in the light of the limitations of the analysis and further analysis should be conducted to overcome these limitations. Firstly, the consideration of the original data is recommended, avoiding assumptions in the cases that the information is available. In addition, a steady-state adjustment of the renewal costs will lead to a far-reaching efficiency assessment, by considering total costs rather than only maintenance costs, and providing more sensitive results.

The results of the analysis have shown that Network Rail has opportunities to learn about its internal best practices, achieving cost savings based on the experience of the top performing regions. However, further research is required to improve the quality of the dataset and its analysis by taking into account the limitations of the current study. The proposed considerations will allow building a more robust analysis by improving its quality, ensuring that the robustness of the analysis is maximized. This study will provide Network Rail with specific tools to develop strategies to reduce the cost inefficiency by sharing the best practices of the best performing regions across the network. Network Rail has the opportunity to achieve efficiency gains through the implementation of their own best practices without investments in technology, but further research is required to improve the quality of the analysis and identify the magnitudes of the potential efficiency gains in a more precise way.

## REFERENCES

- Anderson, R., Hirsch, R., Trompet, M. and Adeney, W. 2003. Developing benchmarking methodologies for railway infrastructure management companies. In: *Proceedings of the European Transport Conference 2003*.
- Battese, G.E. and Coelli, T.J. 1988. Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of econometrics*. 38(3), pp.387-399.
- Beck, A., Bente, H. and Schilling, M. 2013. Railway efficiency: An overview and a look at opportunities for improvement. In: International Transport Forum Discussion Paper.
- Ben-Akiva, M.E. and Lerman, S.R. 1985. *Discrete choice analysis: theory and application to travel demand*. MIT press.
- Booz-Allen and Hamilton Ltd. 1999. *Railtrack's Expenditure Needs 2001–2006: A Report to the Office of the Rail Regulator*. London.
- Booz-Allen and Hamilton Ltd. 2000. *Response to Railtrack's May 2000 Cost Submission: A Report to the Office of the Rail Regulator*. London.
- Bottasso, A. and Conti, M. 2011. Quantitative techniques for regulatory benchmarking.
- Burns, P., Jenkins, C., Reichmann, C. and Mikkers, M. 2006. The Role of the Policy Framework for the Effectiveness of Benchmarking in Regulatory Proceedings,. In: Coelli, T. and Lawrence, D.A. eds. *Performance Measurement and Regulation of Network Utilities*.
- Butcher, L. 2012. *Railways: regulation*. [Online]. [Accessed 26 July 2015]. Available from: <http://researchbriefings.files.parliament.uk/documents/SN02071/SN02071.pdf>
- Campos, J. and Cantos, P. 2001. Railways. In: Estache, A. and de Rus, G. eds. *Privatization and Regulation of Transport Infrastructure. Guidelines for Policymakers and Regulators*. The World Bank Washington, D.C.: WBI Development Studies. , pp.171-234.
- Coelli, T. and Lawrence, D.A. 2006. *Performance measurement and regulation of network utilities*. Edward Elgar Publishing.
- Coelli, T.J., Rao, D.S.P., O'Donnell, C.J. and Battese, G.E. 2005. *An introduction to efficiency and productivity analysis*. Springer Science & Business Media.
- Cuesta, R.A. 2000. A production model with firm-specific temporal variation in technical inefficiency: with application to Spanish dairy farms. *Journal of Productivity Analysis*. 13(2), pp.139-158.
- De Rus, G. 2006. The prospects for competition. In: Gomez-Ibanez, J. and De Rus, G. eds. *Competition in the Railway Industry: An International Comparative Analysis*. . Transport Economics, Management and Policy. Edward Elgar Publishing Limited.
- Europe Economics. 2000. *Review of Railtrack Efficiency: A Report for the Office of the Rail Regulator*. London.

- Farsi, M. and Filippini, M. 2009. An analysis of cost efficiency in Swiss multi-utilities. *Energy Economics*. 31(2), pp.306-315.
- Fried, H.O., Lovell, C.K. and Schmidt, S.S. 2008a. Efficiency and Productivity. In: Fried, H.O., et al. eds. *The measurement of productive efficiency and productivity growth*. Oxford University Press.
- Fried, H.O., Lovell, C.K. and Schmidt, S.S. 2008b. *The measurement of productive efficiency and productivity growth*. Oxford University Press.
- Greene, W.H. 2008. The econometric approach to efficiency analysis. *The measurement of productive efficiency and productivity growth*. pp.92-250.
- Halcrow, TCI and LEK. 2003. *International Benchmarking: Report to ORR, Network Rail and SRA*. London.
- Hansen, I.A., Wiggensraad, P.B. and Wolff, J.W. 2013. Benchmark analysis of railway networks and undertakings. In: *RailCopenhagen2013: 5th International Conference on Railway Operations Modelling and Analysis, Copenhagen, Denmark, 13-15 May 2013*: International Association of Railway Operations Research (IAROR).
- Hilmola, O.-P. 2007. European railway freight transportation and adaptation to demand decline: Efficiency and partial productivity analysis from period of 1980-2003. *International Journal of Productivity and Performance Management*. 56(3), pp.205-225.
- House of Commons - Committee of Public Accounts. 2008. *Delivering a sustainable railway: a 30-year strategy for the railways? Tenth Report of Session 2007-08*. TSO (The Stationery Office).
- Jamasb, T. and Pollitt, M. 2001. Benchmarking and regulation: international electricity experience. *Utilities Policy*. 9(3), pp.107-130.
- Kennedy, J. and Smith, A.S. 2004. Assessing the efficient cost of sustaining Britain's rail network: Perspectives based on zonal comparisons. *Journal of Transport Economics and Policy*. pp.157-190.
- Kumbhakar, S.C., Lien, G. and Hardaker, J.B. 2014. Technical efficiency in competing panel data models: a study of Norwegian grain farming. *Journal of Productivity Analysis*. 41(2), pp.321-337.
- Kumbhakar, S.C. and Lovell, C.K. 2003. *Stochastic frontier analysis*. Cambridge University Press.
- Kumbhakar, S.C., Wang, H.-J. and Horncastle, A. 2015. *A Practitioner's Guide to Stochastic Frontier Analysis Using Stata*. Cambridge University Press.
- LEK. 2003. *Regional Benchmarking: Report to Network Rail, ORR and SRA*. London.
- LEK, TCI and Halcrow. 2003. *Bottom-Up Review of Network Rail's Business Plan: 2003/04-2005/06: Report to ORR*. London.
- Makovsek, D., Benezech, V. and Perkins, S. 2015. Efficiency in Railway Operations and Infrastructure Management.

- McNulty, R. 2011. Realising the Potential of GB Rail, Report of the Rail Value for Money Study, Summary Report. *Department for Transport and ORR*.
- Mirrlees-Black, J. 2014. Reflections on RPI-X regulation in OECD countries. *Utilities Policy*. 31, pp.197-202.
- Nash, C. 2000. Modelling performance: rail. *Handbook of transport modelling, Pergamon*. pp.565-575.
- Nash, C. 2005. Rail infrastructure charges in Europe. *Journal of Transport Economics and Policy*. pp.259-278.
- Nash, C. and Shires, J. 2000. Benchmarking European Railways. An assessment of current data and recommended indicators. Institute for Transport Studies, University of Leeds, United Kingdom. *Transport benchmarking - Methodologies, Applications & Data Needs*. ECMT-OECD, pp.119-136.
- Nash, C. and Smith, A. 2014. Rail efficiency: cost research and its implications for policy. *International Transport Forum*.
- Network Rail. 2014. *Final CP5 volume incentive baselines*. Network Rail.
- Network Rail. 2015. *Track and route mileage*. [Online]. [Accessed 19 July 2015]. Available from: <http://www.networkrail.co.uk/asp/10538.aspx>
- ORR. 2008. *Periodic review 2008: Draft determinations*
- ORR. 2010. *International cost efficiency benchmarking of Network Rail*. ORR.
- ORR. 2013a. Periodic Review 2013: Final determination of Network Rail's outputs and funding for 2014-19.
- ORR. 2013b. *PR13 Efficiency Benchmarking of Network Rail using LICB*. Office of Rail Regulation.
- ORR. 2015a. *Methodology*. [Online]. [Accessed 19 July 2015]. Available from: <https://dataportal.orr.gov.uk/displayreport/report/html/ea0ad5a2-daca-47ec-a72e-b6e50a36a3f1>
- ORR. 2015b. *Million gross tonne miles by freight train operators - Table 13.9*. [Online]. [Accessed 19 July 2015]. Available from: <https://dataportal.orr.gov.uk/displayreport/html/html/36ac9a10-b60e-48c5-875b-a35ed58ff772>
- ORR. 2015c. *Number of broken rails by route classification - Table 2.56*. [Online]. [Accessed 19 July 2015]. Available from: <https://dataportal.orr.gov.uk/displayreport/html/print/fc88f120-61e3-49e3-9e5d-f50c2b3d5190>
- ORR. 2015d. *What we do*. [Online]. [Accessed 26 July 2015]. Available from: [orr.gov.uk/about-orr/what-we-do](http://orr.gov.uk/about-orr/what-we-do)
- Oum, T.H., Waters, W. and Yu, C. 1999. A survey of productivity and efficiency measurement in rail transport. *Journal of Transport economics and Policy*. pp.9-42.
- Oxera. 2009. *Recommendations on how to model efficiency for future price reviews* Prepared for Office of Rail Regulation.

- Oxera. 2014. *What is the contribution of rail to the UK economy?* : Prepared for the Rail Delivery Group.
- Parker, D., Dassler, T. and Saal, D.S. 2006. 6 Performance benchmarking in utility regulation: principles and the uk's experience. *International Handbook on Economic Regulation*. p117.
- Parmeter, C.F. and Kumbhakar, S.C. 2014. Efficiency Analysis: A Primer on Recent Advances. *Foundations and Trends (R) in Econometrics*. 7(3-4), pp.191-385.
- Pollitt, M.G. and Smith, A.S. 2002. The restructuring and privatisation of British Rail: was it really that bad? *Fiscal Studies*. 23(4), pp.463-502.
- Ricketts, M. 2006. Economic regulation: principles, history and methods. In: Crew, M.A. and Parker, D. eds. *International handbook on economic regulation*. Edward Elgar Publishing.
- Sheskin, D.J. 2003. *Handbook of parametric and nonparametric statistical procedures*. crc Press.
- Smith, A. 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. 2015. *Efficiency Analysis in UK Price Reviews*. Economics of Transport Regulation.: Institute for Transport Studies, University of Leeds.
- Smith, A., Wheat, P. and Nixon, H. 2008. International Benchmarking of Network Rail's Maintenance and Renewal Costs. *Report written as part of PR2008*.
- Smith, A., Wheat, P. and Smith, G. 2010. The role of international benchmarking in developing rail infrastructure efficiency estimates. *Utilities policy*. 18(2), pp.86-93.
- Smith, A.S. 2005. The role of efficiency estimates in UK regulatory price reviews: The case of rail. *Utilities Policy*. 13(4), pp.294-301.
- 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*. pp.1-44.
- Smith, A.S. 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*. 48(2), pp.503-515.
- Stern, J. 2003. *Econometric Benchmarking and its Uses by ORR: A Review*. CCRP Working Paper No 21.
- Van Keilegom, I. and Wilson, P.W. 2011. *Exploring Research Frontiers in Contemporary Statistics and Econometrics: A Festschrift for Léopold Simar*. Springer Science & Business Media.
- Wheat, P. 2006. Dissertation MA Transport Economics. University of Leeds. [Online].
- Wheat, P. and Smith, A. 2008a. Assessing the Marginal Infrastructure Maintenance Wear and Tear Costs for Britain's Railway Network. *Journal of Transport Economics and Policy*. 42, Part 2, pp.189-136.

Wheat, P. and Smith, A. 2008b. *A Response to the LECG and Horton 4 Consulting Reports on the ITS/ORR International Benchmarking Study (Based on the LICB Dataset)*. Report written as part of PR2008, October 2008. .

Wooldridge, J.M. 2010. *Econometric analysis of cross section and panel data*. MIT press.