Article:

Smith, ASJ (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). 503 - 515. ISSN 1366-5545

https://doi.org/10.1016/j.tre.2011.10.003

The Application of Stochastic Frontier Panel Models in Economic Regulation: Experience from the European Rail Sector

Andrew S.J. Smith[a]

[a] Institute for Transport Studies and Leeds University Business School, University of Leeds. Postal Address: Room 213, Institute for Transport Studies, 36-40 University Road, University of Leeds, Leeds, LS2 9JT UK. Tel (direct): + 44 (0) 113 34 36654. Tel (enquires): +44 (0)113 34 35325. Fax: + 44 (0) 113 34 35334. Email: a.s.j.smith@its.leeds.ac.uk

Abstract

This paper shows how stochastic frontier panel techniques can be used by economic regulators to benchmark regulated firms against international best practice. We utilise a unique, panel dataset of European rail infrastructure managers (1996-2006). A time-varying inefficiency model, with firm-specific time paths for inefficiency, is adopted. The results were used in the 2008 regulatory review of the British infrastructure manager, Network Rail, and showed that the company faced an efficiency gap of around 40% against European best practice – in line with engineering-based evidence. More widely, the paper highlights the advantages of the inefficiency specification adopted for use in economic regulation.

Keywords
Stochastic Frontier Model; Efficiency; Panel Data; International Benchmarking; Economic Regulation

JEL codes

C23 ,C81, L51, D24

1 Introduction

Starting with the 1991 European Commission Directive 91/440 (and later Directives 2001/12, 2001/13 and 2001/14), Europe has embarked on a process of rail deregulation, progressively opening up rail markets to competition. Via successive legislation, Europe's rail systems have been required to separate train operations and infrastructure (accounting and organisational), and many countries have gone further and implemented full institutional separation. The reforms have thus created a major regulatory challenge. Typically rail infrastructure managers (IMs) are natural monopolies and also lack domestic comparators that would support the use of yardstick competition to restrain costs to efficient levels. Thus, international benchmarking takes on an increased importance as compared with, for example, electricity distribution, where regional networks within a country may be compared against each other.

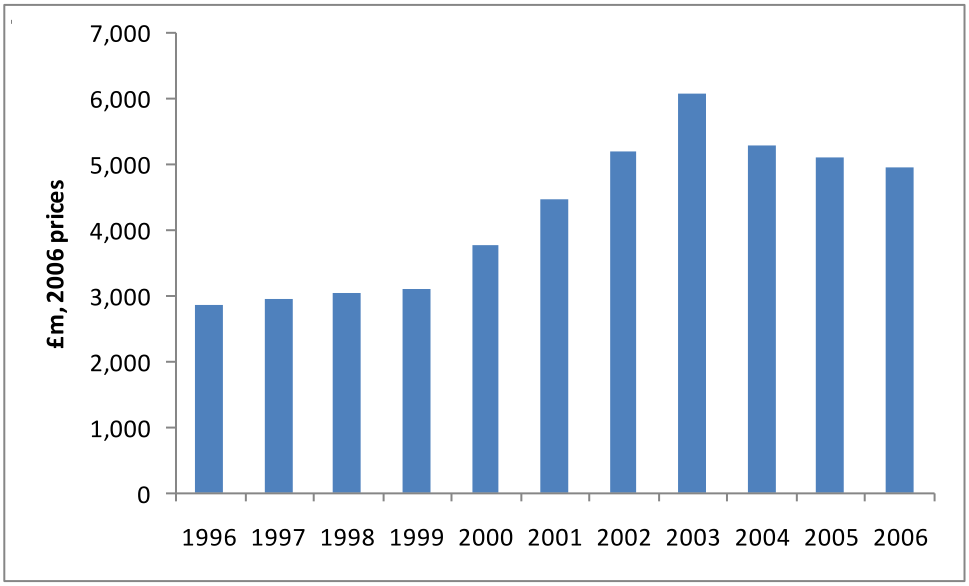
However, international efficiency comparisons raise serious challenges for economic regulators. Lack of comparable data across countries and over time is usually a major hurdle. Capital cost measurement is a particular problem in network industries. Controlling for cross-country network differences that have nothing to do with inefficiency is a further challenge. There is also the task of specifying an appropriate efficiency model specification for use in this context. Given that the number of comparators is often small, thus necessitating the use of panel data, the model needs to be sufficiently flexible to allow for inefficiency variation over time, whilst recognising that different countries may be experiencing very different trends in inefficiency. At the same time, the results need to be usable by regulators.

This paper demonstrates how stochastic frontier panel data techniques can be used by economic regulators to determine the efficiency of regulated firms against international best practice. The specific context is the situation which arose in Britain which took the most radical approach to rail reforms in Europe, opting for full institutional separation combined with privatisation of the separated infrastructure manager, Railtrack. As part of the British reforms, the Office of Rail Regulation (ORR)[1] was set up, principally to regulate the monopoly infrastructure element of the business. The traditional, RPI-X regulatory approach was adopted, in order to incentivise cost reduction (Beesley and Littlechild, 1988). One of ORR's main roles is to conduct periodic reviews of the costs that an efficient infrastructure company would incur in operating, maintaining, renewing and developing the rail infrastructure.

Following the reforms, rail infrastructure costs in Britain increased substantially (Figure 1), confounding theoretical expectations and empirical evidence concerning the impact of change in ownership on productive efficiency (see, for example, Viscusi et. al, 2005). The cost increases led to Railtrack being placed into administration, and replaced by Network Rail in 2002. Costs later started to fall, albeit from very high levels, following interim regulatory action by ORR in 2003 (Figure 1). However, despite falling costs, by the time of the 2008 Periodic review, it was clear that costs were still much higher than at privatisation. With no domestic comparators, the only way to gain an objective view of Network Rail's efficiency performance was to look to international comparisons.

[1] Originally named the Office of The Rail Regulator, but later re-named to reflect the move away from vesting all powers in a single person (the regulator), but instead adopting a regulatory board structure.

Figure 1: Rail Infrastructure Operation, Maintenance and Renewal Costs



We overcome the traditional data problems by utilising a unique, panel dataset of 13 European rail infrastructure managers (1996 to 2006). The International Union of Railways (UIC) has been collecting data on rail infrastructure managers over a number of years, and during that time work has been done to ensure the comparability of the data through, for example, the specification of common data definitions. The dataset contains a range of variables that capture variation in scale, usage and infrastructure characteristics. UIC agreed to provide this data for the purpose of econometric work in support of the 2008 Periodic Review process, though with strict confidentiality conditions meaning that the efficiency scores for countries other than Britain could not be disclosed. The results therefore focus on the results for Network Rail, and how they were applied in the context of the 2008 efficiency determination for the company.

We adopt a time varying inefficiency model (an augmented version of the model proposed by Cuesta, 2000) as our preferred model. This model has a number of desirable properties in the present context. It allows inefficiency to vary over time, whilst permitting firm-specific time paths of inefficiency. Further, non-monotonic change in inefficiency over time is permitted for Network Rail, which is important given the profile of cost changes for the company (Figure 1). Alvarez et. al. (2006) also note that a key advantage of this model is that it enables the unrealistic assumption of independence in inefficiency over time (a problem that plagues many comparator models) to be relaxed, which is important in a regulatory context. The robustness of our preferred model is ensured by testing against alternative specifications. Confidence intervals (or more precisely, prediction intervals) on the efficiency scores are also reported. The econometric results are then briefly compared against bottom-up, engineering evidence.

Whilst there is an extensive literature on railway system efficiency, this is the first paper to apply stochastic frontier techniques to study the relative efficiency performance of rail infrastructure managers. The results were used in the 2008 regulatory review of Network Rail, covering some £27 bn (c. €33 bn) of expenditure. The methodology also has wider applicability as new rail regulatory bodies are emerging elsewhere in Europe. The paper is structured as follows. Section 2 sets out the regulatory background. Section 3 explains the choice of efficiency model specification in the context of the academic and regulatory literatures. Section 4 describes the data and the issues to be overcome in conducting international comparisons. Section 5 presents the results and how they were used by the economic regulator. Section 6 offers some conclusions.

2 Regulatory background

A detailed review of the regulatory efficiency studies undertaken by ORR prior to the 2008 review can be found in Kennedy and Smith (2004) and Smith (2005). This section summarises the findings and issues relevant to this paper.

The first, post-privatisation periodic review carried out by ORR in 2000 resulted in the regulator setting an efficiency target of 17% over the five year period 2002 to 2006. This was based predominantly on comparison of trends in other UK privatised utilities (ORR, 2000); with the assumption being made that Railtrack should be capable of achieving similar savings following its change of ownership. More widely, both the theory and the evidence suggest that privatisation, combined with high powered RPI-X regulation should deliver productivity gains (see, for example, Viscusi et. al., 2005). At the time, there was insufficient data available in the public domain to do a comparison of productivity levels across countries, but some trend comparisons from US Class 1 railroads supported the evidence from other utilities (NERA, 2000).

In common with the approach taken by other UK economic regulators, ORR also commissioned bottom-up consultant reviews that corroborated the other evidence (Booz Allen Hamilton, 1999; 2000). These studies sought to identify specific efficiency improvement initiatives, quantify their impact, and thus determine an efficiency target from the bottom up. The bottom-up approach contrasts with top-down methods, such as those based on stochastic frontier modelling using actual data, which are the focus of this paper; but the two types of approaches are useful complements, as evidenced by their use amongst economic regulators.

However, just as the 2000 Periodic Review conclusions were being finalised, a train derailment on the East Coast Mainline at Hatfield[2], resulting from defective track, set off a chain of events which resulted in Railtrack being placed into administration roughly one year later (October 2001). The derailment heightened concerns over the condition of Britain's rail infrastructure and Railtrack management responded by imposing speed restrictions across the network. Maintenance and renewal activity was also stepped up, leading to a sharp increase in the costs of operating, maintaining and renewing Britain's rail network (Figure 1). The accident precipitated a major financial crisis at Railtrack, and following a year in administration, a new company, Network Rail, took ownership of Britain's rail infrastructure. Network Rail is a company limited by

[2] The East Coast Mainline is a high speed (125 mph (200 km/h)) line connecting London with Edinburgh in Scotland. Hatfield is a town just north of London.

guarantee, owned by members, not shareholders, and funded entirely by debt financing (see Pollitt and Smith, 2002).

Despite the change of ownership structure, the question of efficiency remains of central importance in the regulation of Network Rail. Given the very sharp cost rises that ensued, the ORR announced an Interim Review of the company's operations and efficiency (2003 Interim Review). In the more stable conditions prior to 2000, previous efficiency studies had focussed on trend based comparisons, based on the assumption that Railtrack could achieve savings in line with the efficiency gains achieved in other recently privatised industries. However, given the very substantial cost increases that had occurred, it was clear that such trend-based comparators would be insufficient to draw meaningful conclusions about the scope for Network Rail to get costs down going forward.

ORR commissioned a wide range of studies during the 2003 Interim Review. However, it did not prove possible to conduct top-down international comparisons in the time available, and ultimately ORR's efficiency determination was largely based on two bottom-up consultant reviews of Network Rail's business plan (LEK, TTCI and Halcrow, 2003 and Accenture, 2003). These results were supplemented by internal benchmarking, which indicated the kind of savings that could be achieved if Network Rail implemented its own best practice consistently across the network.

Though bottom-up evidence is an important part of an economic regulator's toolkit, the chief weakness of such studies is that they rely heavily on the judgement of the consultants involved. There is also a danger that their findings will depend on the context in which the review is being conducted, meaning that the studies may to some extent "find what they are expected to find" (see Smith, 2005). The internal benchmarking work, though based on top-down techniques, of course did not offer an external perspective.

Ultimately then, the 2003 Interim Review was unable to provide a clear, empirically based assessment of Network Rail's relative efficiency position based on hard data from external sources. ORR nevertheless set a tough efficiency target of 31% over 5 years (2004-2009). This was based on the evidence available, which though not totally conclusive, suggested that efficiency had deteriorated during Railtrack's period of administration and immediately afterwards. However, costs were starting from a very high base. Thus, although costs then started to fall as Network Rail set about delivering its efficiency targets, by the time of the next periodic review in 2008, the scene was set to take the benchmarking approach a step forward by attempting international comparisons.

3 The model

Given the many different factors influencing railway costs, previous international benchmarking studies have noted the problem of making comparisons based on simple unit cost measures – that is, measures that relate costs to a single measure of scale or output. For example, very different results are obtained in railway comparisons, depending on whether unit costs are measured relative to track kilometres or train kilometres. It is thus important to develop an econometric model that simultaneously take account of variation in track km and train km, as well as other relevant cost drivers (see for example, Nash, 1985 and NERA, 2000). Statistical methods ultimately provide a single, and more definitive measure of relative efficiency, than can be obtained from comparing the results of several different unit cost measures.

A range of possible econometric techniques exist. The stochastic cost frontier model estimated in this paper can be represented as:

$$C_{it} = f(Y_{it}, P_{it}, N_{it}, \tau_t; \beta) + v_{it} + u_{it} \tag{1}$$

where the first term ($f(Y_{it}, P_{it}, Q_{it}, \tau_t; \beta)$) is the deterministic component, and Y_{it} is a vector of output measures, P_{it} is a vector of input prices, N_{it} is a vector of exogenous network characteristic variables, τ_t is a vector of time variables which represent technical change and β is a vector of parameters to be estimated. As is standard in the literature we start with a translog model, and test the Cobb-Douglas restriction (see section 5). C_{it} represents the infrastructure cost variable to be explained. The precise specification of the deterministic component and the definition of costs used in the analysis are discussed in section 4. Here we focus on the error specification.

The v_{it} term is a random component representing unobservable factors that affect the firm's operating environment. This term is distributed symmetrically around zero (more specifically assumed to be normally distributed with zero mean and constant variance). A further one sided random component is then added to capture inefficiency (u_{it}).

Since we have panel data there are a range of possible assumptions concerning the path of the inefficiency (u_{it}) over time available from the literature (see for example, Greene, 2008). The simplest assumption is the pooled stochastic frontier model which treats each data point as a separate firm (see Pitt and Lee 1981; Model II). This model assumes that the u_{it} are independent over time.

Several alternatives, which recognise the panel structure of the data, whilst allowing inefficiency to vary over time have been proposed. Building on the work of Schmidt and Sickles (1984), which used traditional panel data techniques to obtain estimates of inefficiency, Lee and Schmidt (1990) specify the following model for the firm-specific inefficiency terms:

$$\alpha_{it} = \theta_t \delta_i \tag{2}$$

where α_{it} is the intercept term for firm i at time t. This model assumes that the temporal pattern of technical inefficiency (represented by θ_t) is completely unrestricted[3], but that this pattern is the same for all firms. As a result, the ordinal rankings of the firms remain the same in each year and firms cannot therefore overtake one another. This model was extended by Lee (2006) to permit firms in different groups to have different time paths of inefficiency.

[3] That is, inefficiency may increase one year, and fall back the next.

In the same vein, Cornwell, Schmidt and Sickles (1990) specify a quadratic function for the time-varying efficiency terms:

$$\alpha_{it} = \theta_{i1} + \theta_{i2}t + \theta_{i3}t^2 \tag{3}$$

where the parameters of the model can be estimated using the fixed or random effects panel data methods. This model provides additional flexibility compared to the Lee and Schmidt (1990) model, since the ranking of firms is permitted to change over time. One of the weaknesses of the formulation in equation (3) is that it requires the estimation of three parameters ((θ_1, θ_2, and θ_3) for each cross-section.

Within the sphere of maximum likelihood stochastic frontier methods a number of other methods have been developed in parallel. Battese and Coelli (1992) developed an approach which allowed for variations in efficiency over time, but imposed structure on those variations. Their model is specified as:

$$u_{it} = u_i \cdot \exp(-\eta \cdot (t - T)) \tag{4}$$

where the u_{it} term denotes the inefficiency effect for the ith firm in time period t and the random variable u_i can be thought of as the inefficiency effect for the ith firm in the last period of the panel (period T). In turn, the inefficiency effects in earlier years are either an increasing ($\eta>0$) or decreasing ($\eta<0$) function of the level of inefficiency in the last year, where η is a parameter to be estimated (a trend is therefore imposed on the efficiency scores over time). Kumbhakar (1990) specifies a similar, alternative specification:

$$u_{it} = u_i / [1 + \exp(bt + ct^2)] \tag{5}$$

In both of the above models, all firms are forced (by assumption) to have the same direction of efficiency change over time. The change in efficiency over time is also monotonic[4].

The Battese and Coelli (1992) model was extended by Cuesta (2000) to permit firm-specific time paths of inefficiency.

$$u_{it} = u_i \cdot \exp(-\eta_i \cdot (t - T)) \tag{6}$$

where

$$u_i \sim N^+\left(0, \sigma_u^2\right)$$

As noted in Greene (2008), the Battese and Coelli (1992) and Cuesta (2000) models can also be extended to permit non- monotonic changes in efficiency over time:

[4] Although Battese and Coelli (1992) proposed a non-monotonic version of their model, the form in equation (4) is the one that is generally implemented.

$$u_{it} = u_i \cdot \exp(\eta_{i1} \cdot (t - T) + \eta_{i2} \cdot (t - T)^2) \tag{7}$$

Kumbhakar and Heshmati (1995) proposed a model that allows inefficiency to be decomposed into a persistent and time varying component. This model has more recently been re-interpreted, with the time invariant inefficiency term assumed to be unobserved heterogeneity, and the (random) time varying component to be inefficiency ("true" models; see Greene, 2005 (though it is unclear how to distinguish persistent inefficiency from unobserved heterogeneity; see Farsi, Filippini and Greene, 2006). A further drawback for both types of models is that the time varying component is assumed to vary randomly over time, which is a problem for testing the temporal variation in efficiency. The class of models in which time-related variables influence of the means of the inefficiency distribution (for example, Battese and Coelli, 1995) suffer from the same problem, as noted by Alvarez et. al. (2006).

In this paper we adopt the specification set out in equation (7). This model has a number of important and desirable features for efficiency estimation in a regulatory context. First the model permits inefficiency to vary over time, which is important given the length of the panel. Second, it allows firm-specific time paths of inefficiency. Given the international nature of the dataset, and the different regulatory arrangements in place across countries, it is important allow inefficiency to move in different directions for different firms. The latter is particularly important given the sharp changes in costs (and potentially efficiency) experienced by Network Rail, which it might not be expected to be replicated across other rail networks in Europe.

Third, it permits non-monotonic change in efficiency over time, which given the number of parameters to be estimated, and the focus in this paper on the efficiency performance of the British operator, we operationalise only for Network Rail. This additional flexibility is considered important given the particular profile of cost changes seen over the period of the analysis for the company (see Figure 1), which are much more pronounced than those seen in the other countries. We permit further flexibility for Britain by splitting the data into two companies: Railtrack (first four years of the sample) and Network Rail (last seven years of the sample). However, the results are not sensitive to this data split.

Finally, as noted by Alvarez et. al. (2006), a key advantage of such models is that there is no need to make the unrealistic assumption of independence in inefficiency over time; a problem that plagues a number of the alternative models discussed above. We consider this feature of the proposed model to be particularly important in a regulatory context, where economic regulators set efficiency targets and expect to see firm inefficiency to change in a structured and not random way over time. In the empirical analysis that follows we compare the results of the preferred model against a wide range of alternative model specifications.

The model adopted therefore has precedents in the academic literature. Turning to regulatory best practice, in general, economic regulators that have applied econometric methods have used relatively simple approaches. UK economic regulators have made extensive use of the corrected ordinary least squares (COLS) technique (see Greene, 1980), for example, by the Office of the Gas and Electricity Markets (OFGEM), the Office of Water Services (OFWAT) and the Office of Communications (OFCOM).

The Postal Services Commission (POSTCOMM) and OFCOM have also applied stochastic frontier analysis as part of the regulatory review process. In all of the aforementioned cases the approaches have been applied to cross-sectional data, although some are starting to utilise panel data, though using standard regression techniques (e.g. OFGEM, 2009). More widely, the highly regarded study commissioned by the German Network Agency in respect of gas and electricity distribution benchmarking in Germany applied stochastic frontier techniques, though again to cross-sectional data (Sumicsid (2007))[5].

Our proposed model is therefore well suited to application in a regulatory context as compared to alternative specifications. Further, we are not aware of any previous regulatory studies that have that have applied techniques of the kind adopted here[6]. From a methodological perspective the analysis therefore represents an advance on previous methods used in a regulatory context (at least in cases where panel data has been available and utilised). Its application to international comparison of rail infrastructure managers is also unique in the literature, and was made possible by the existence of a pre-existing dataset collected over a number of years, as described in the next section.

4 Data

This section first summarises the data. The choice of dependent variable, and in particular the question of how to measure capital costs, is then discussed.

4.1 Summary of the data

As noted in the introduction, international benchmarking is often hampered by the lack of comparable data across countries and over time. Here we have been able utilise a unique, panel dataset of 13 European rail infrastructure managers (1996 to 2006) that had already been collected by the UIC. The dataset has been developed over a number of years (starting in 1995), and forms the basis for UIC's own benchmarking methodology (see UIC, 2007). Each participating company submits data every year to UIC, according to a common set of data definitions. The UIC agreed to provide this data for the purpose of econometric work in support of the 2008 Periodic Review process (the data had not previously been subject to econometric analysis). Having access to a pre-existing dataset from UIC clearly impacted on the ability of ORR to place international benchmarking, based on the models shown in this paper, at the forefront of its efficiency analysis in the 2008 Periodic review. It should be noted that European legislation requires member states to separate train operations and infrastructure (at least into separate divisions with their own accounts), and it is therefore possible to readily separate rail infrastructure costs from those of train operations.

The dataset covers the following countries: Britain, The Netherlands, Norway, Portugal, Finland, Sweden, Ireland, Belgium, Germany, Austria, Italy, Denmark and Switzerland. Owing to the

[5] For example, pages 35-38.
[6] This point was also made in a keynote speech on international experience of regulatory efficiency analysis, by Dr Michael Pollitt, University of Cambridge, at the North American Productivity Workshop (NAPW), NYU Stern, New York 2008.

large changes in cost occurring in Britain over the period, in the analysis that follows, Britain is split into two companies: Railtrack (first four years of the sample) and Network Rail (last seven years of the sample). This offers even more flexibility concerning the time path of inefficiency for Network Rail, although the results are not sensitive to this data split. The dataset is summarised in Table 1 below. The data was provided by UIC on a confidential basis. This commitment was a formal requirement prior to obtaining the data and without which the data would not have been released for analysis. Thus, we are able to report all aspects of the model results, but are only permitted to reveal the efficiency scores of Network Rail.

Table 1: Summary of the Data

	Mean	Standard Deviation	Variable name[3]
Maintenance costs (PPP Euros[1], 2006 prices)	Cost data confidential		MAIN
Renewal costs (PPP Euros[1], 2006 prices)	Cost data confidential		RENEW
Total costs (PPP Euros[1], 2006 prices)[2]	Cost data confidential		TOTAL
Route-km	8,566	9,467	ROUTE
Passenger train-km per route km	18,477	11,175	PASSDR
Freight train-km per route km	4,180	2,294	FRDR
Single-track km divided by route-km	0.61	0.23	SING
Electrified track km divided by track-km	0.61	0.26	ELEC
Average salary	Cost data confidential		WAGE

[1] German Euros as the numeraire

[2] Total costs here is the sum of maintenance and renewal costs, but excludes operational costs

[3] Expressed as logs of the variables

As is standard in the rail cost function literature, given the highly regulated nature of the railways, the main output vector comprises passenger and freight train-km (expressed as densities[7]), together with route-km in order to distinguish between scale and density effects. The model also includes variables that capture exogenous differences in the network characteristics, in this case the proportion of track electrified and the ratio of single track to route-km (see, for example, Nash, 1985; Cantos et. al., 2010; Wheat and Smith, 2008).

Input price variation between countries and over time is addressed in two ways. First, The cost data was converted into a common currency and price level using Purchasing Power Parity (PPP) exchange rate data from the OECD to convert the data to a common currency and price level (with German Euros in 2006 as the numeraire). Thus general (economy wide) differences in price (and wage) levels between countries are taken into account. In addition, we also test the impact of rail specific wage rates (the WAGE variable in Table 1). It should be noted that the wage rate variable was not complete for all firms in all years, and therefore had to be in-filled in some

[7] That is, per route-km.

cases. Inspection of the data also suggested that this variable was not updated every year, indicating some measurement error in the data. We thus have less confidence in the accuracy of this variable, but we test its inclusion in the model (see section 5).

Ideally, we would also want to include measures of asset age/condition and performance, but such data was not systematically available for all countries in all years. We return to this point in the results section.

4.2 The dependent variable

The dependent variable for our preferred model is annual total cost, which is the sum of annual maintenance and annual renewal cost. Maintenance expenditure refers to expenditure required to maintain the condition and performance of the network to the required standards but without any significant replacement of assets (this item includes inspection of the network). Renewals expenditure relates to replacement of the existing infrastructure on a like-for-like basis (such that it does not result in any change in performance as compared to the original asset). The cost of day-to-day operations (for example, the cost of signalling and timetabling operations are not included in the dataset) and the cost of new investment – that is enhancements to the network - are not included due to lack of data. Thus, in our analysis, firms are assumed to minimise total maintenance and renewal costs subject to exogenously fixed traffic volume on a network of fixed size, with exogenously determined network characteristics.

There are two reasons for preferring a measure of total costs as compared with analysing maintenance and renewal costs separately. Firstly, there may be capital substitution effects that distort comparisons across countries. Secondly, despite efforts to harmonise definitions, there may be some residual differences in definitions at the boundary between maintenance and renewals (though our experience suggests that these are likely to be small)[8].

However, whilst maintenance cost is an ongoing cost that can easily be compared across rail infrastructure managers, the measurement of capital is problematic. The longevity of assets employed in the rail industry (and other network industries) means that historic-cost based accounting based measures of capital costs are unlikely to be a good measure in general. More specifically, different railways apply different depreciation policies and some may have undertaken revaluations; others not. Furthermore, hidden subsidies within some countries mean that the book value of net assets in the accounts may exclude subsidies provided by government subsidies in some cases.

Thus there are strong reasons for preferring a cash-based measure of capital which, in the case of the UIC database, is annual expenditure on renewal activity. This approach is also widely used in comparative efficiency analysis of other regulated network industries (see Frontier Economics, 2010). However, capital expenditure-based measures can fluctuate from year-to-year, for example as part of the natural cycles that may be present in the profile of required renewal activity. There is thus a danger that an observation of low (or high) costs in a particular year is mistaken for

[8] It is possible that there are also some remaining differences in definitions at the boundary between renewals and enhancement costs, although the UIC dataset was put together based on an agreed set of definitions, and there was no specific evidence of this problem in the data.

efficient (or inefficient) operation, when in fact it simply reflects the potential lumpy nature of renewal activity.

One alternative is to use maintenance costs plus renewal costs averaged over a number of years. However, this approach has the strong disadvantage of reducing the number of observations (reduces T), thus in part negating the benefits of panel data which are so important in regulatory studies where the number of firms (N) is small.

In this paper we adopt the following approach. Given the particular profile of cost changes for Network Rail in recent years (see Figure 1), which in part is driven by increased renewal activity, the consensus within the industry is that in the early years of the sample the equity-funded Railtrack was renewing below the level required to maintain the network in a steady-state condition, whilst in the latter years renewal activity has been above this steady-state level. Thus an adjustment was made to Network Rail's track and signalling renewals expenditure based on engineering analysis conducted by ORR (see ORR, 2008)[9]. This analysis suggested that renewal of 2.5% track and signalling assets represents a reasonable approximation to a steady-state level of renewal activity. Thus, where renewal activity fell below (or above) this level, the track and signalling renewal cost element of the dependent variable for Network Rail was adjusted up (or down) accordingly.

There was insufficient hard data to make similar adjustments for other railways. However, we consider that our approach is robust for the following reasons. First, inspection of the data and anecdotal evidence did not give any reason to believe that the frontier firms - which are of particular importance in determining the position and shape of the frontier - were substantially away from steady-state (see, ORR, 2008[10]). Second, the stochastic frontier approach itself (which seeks to separate noise and inefficiency), and the use of panel data over an 11-year period, provide further safeguards against the risk of mis-interpreting low costs in one particular year as evidence of efficient operation, and thus creating an unrealistic benchmark. Third, as final checks on the modelling work, the results of the preferred model are compared against a model that does not include any adjustment to Network Rail's raw cost data, and also against a maintenance-only model. Finally, the results are compared against the results of engineering-based, bottom-up efficiency studies.

5 Results

This section is divided into three sub-sections. We first present the frontier parameter estimates, before going on to report and discuss the efficiency results from the preferred model and comparator inefficiency specifications. Confidence intervals (or more precisely, prediction intervals) on the firm-specific efficiency estimates for Network Rail are also reported. Finally, we briefly compare the results of the econometric modelling work compares against bottom-up engineering based evidence and comment on the regulator's use of the econometric results in reaching its final efficiency determination for Network Rail in 2008.

[9] See pages 123 and 133.
[10] Paragraph 7.4.2.

5.1 Frontier parameter estimates

The frontier and efficiency parameter estimates for the preferred model (and two comparator models) are shown in Table 2 below. The variable names are as specified in Table 1, with two additional variables, TIME and TIME2 (a time trend and squared time trend respectively). As described in section 4, the dependent variable for the preferred model is total costs (maintenance plus renewals), with the British data adjusted where renewal activity is considered to be away from steady-state. The dependent variables for the two comparator models are total costs (with no adjustment for the British data) and maintenance costs only.

Table 2: Frontier and Efficiency Parameter Estimates

Preferred model Dependent variable: Total costs (steady-state adjusted)			**Comparator model** Dependent variable: Total costs (unadjusted)			**Comparator model** Dependent variable: Maintenance costs		
	Coeff.			Coeff.			Coeff.	
Frontier parameters								
CONSTANT	6.2453	***	CONSTANT	6.2382	***	CONSTANT	5.4770	***
ROUTE	1.0743	***	ROUTE	1.0913	***	ROUTE	0.8430	***
PASSDR	0.3345	***	PASSDR	0.3115	***	PASSDR	0.1362	**
FRDR	0.1792	***	FRDR	0.1472	***	FRDR	0.1567	***
SING	-0.9181	***	SING	-0.9681	***	SING	-0.7146	***
ELEC	-0.0370		ELEC	-0.0690		ELEC	0.0733	
TIME	0.0556	***	TIME	0.0561	***	TIME	0.0469	***
TIME2	-0.0048	***	TIME2	-0.0048	***	TIME2	-0.0027	**
Efficiency parameters[1]								
λ	4.0541	***	λ	4.1810	***	λ	3.6678	***
σ_u	0.4560	***	σ_u	0.4694	***	σ_u	0.3374	***
η_{R1}	0.0585		η_{R1}	-4.5467		η_{R1}	0.1634	**
η_{N1}	0.2252		η_{N1}	0.2031	**	η_{N1}	0.2689	**
η_{N2}	-0.0570	**	η_{N2}	-0.0513	**	η_{N2}	-0.0520	***

*** (**, *) indicates parameter significance at the 1% (5%, 10%) level

[1] Other firm specific η parameters are included in the model but not shown for confidentiality reasons. $\lambda = \sigma_u/\sigma_v$

The preferred model was arrived by starting with a translog model (also including a labour input price variable) and testing down. Based on a likelihood ratio (LR), the Cobb-Douglas restriction could not be rejected at the normal levels of statistical significance, though a second order time trend was clearly significant and this variable was therefore retained[11]. The wage variable was statistically insignificant. This could mean that the PPP adjustment has adequately controlled for sector-specific variation in input prices, or be a result of measurement error in the WAGE variable (see section 4). As explained in section 4, for confidentiality reasons, only the η

[11] The translog model also produced negative estimates on some of the output variables at the sample mean.

parameters for Britain are shown (subscript R for Railtrack, and subscript N for the successor company, Network Rail; and the N1 and N2 subscripts refer to the coefficients on the first order and second order terms in equation 7).

In general, the parameter estimates are well behaved in that their signs make sense from an engineering perspective and are line with previous econometric work. In respect of the signs of the coefficients, the estimates for ROUTE, PASSDR and FRDR are positive as expected. The coefficient on the SING variable is negative which reflects the fact that, for a given level of route-km, more single track implies less track-km to maintain / renew. There may be an additional negative impact of single track on costs, in that track possessions may be handled more efficiently on single track routes if the line can be completely closed during work[12]. The sign of the coefficient on the electrification variable is potentially ambiguous; on the one hand, more electrification assets should require higher M&R activity. On the other, it is possible that the electrification variable is acting as a proxy for other network characteristics.

The parameter estimates are also statistically significant, with the exception of the electrification coefficient. However, we prefer to retain it in the model, based on theoretical considerations - that is, we expect this variable to impact on costs. It is possible that its effect in the M&R model is being obscured by its correlation with some of the other explanatory variables. It may also be possible that the third rail electrification system in Britain – which will have a different impact on costs – could be impacting on the coefficient in respect of this variable, although the British third rail network will be a comparatively small share of total European electrified track.

The size of the parameter estimates are also in-line with previous evidence. Since the volume measures are expressed as densities (per route-km), the elasticity of costs with respect to scale is given by the coefficient on the ROUTE variable. The point estimate is marginally greater than one, but we cannot reject the hypothesis of constant returns to scale at the 5% level. There is a wide range of evidence in the literature from increasing through constant to decreasing returns. For vertically integrated railways the evidence suggests constant returns to scale, whilst the evidence is more mixed in respect of European railways (see, for example, Caves et. al., 1985; Gathon and Perelman, 1992; Andrikopolous and Loizides, 1998). More recent evidence based on infrastructure-only data suggests increasing returns to scale, though this second body of evidence utilises disaggregate data (e.g. at track section level), so the scale estimates are not directly comparable (see, for example, Wheat et. al. , 2009).

With respect to the parameter estimates on PASSDR and FRDR, the preferred model implies a total elasticity of costs with respect to traffic of 0.51. This finding implies strong economies of density, which again is in line with the general rail cost literature. The specific infrastructure-only cost modelling literature referred to above, suggests a range of cost elasticities with respect to traffic of 0.20-0.35 for maintenance-only, and up to 0.49 for maintenance and renewal (renewals cost are in general more responsive to traffic levels); see Wheat et. al. (2009). The comparator models also produce reasonable parameters estimates in line with the preferred model; and we

[12] On the other hand there may be additional advantages to multiple track (e.g. materials can be delivered more cheaply if it is possible to use the adjoining track); but overall, the negative sign on the SING variable is in line with expectations.

note in particular that the overall traffic cost elasticity for the maintenance-only model is lower, as expected, and within the range of previous work (0.29).

The above discussion suggests then that the frontier parameter estimates are reasonable in the context of the previous literature. At this point it is worth noting an important benefit of parametric methods, as compared for example with Data Envelopment Analysis (DEA), namely that we can check whether the frontier parameter estimates are reasonable prior to looking at the inefficiency estimates resulting from the model. We consider that it is therefore possible to have more confidence in the efficiency scores resulting from a parametric model for this reason.

5.2 Efficiency parameters and scores

We now turn to discuss the efficiency results from the models shown in Table 2. We note that for all three models the null hypothesis of no inefficiency effects can clearly be rejected based on a likelihood ratio (LR) test (at any reasonable levels of significance). Further, one of our motivations for adopting the flexible inefficiency specification was that we considered it important to permit firm-specific time trends in inefficiency, thus allowing the direction (and size) of inefficiency change to differ by firm.

We therefore test the preferred model against two restricted (nested) versions of the preferred model. First, the Battese and Coelli (1992) restriction, which estimates a single inefficiency trend parameter for the whole industry, is rejected at any reasonable levels of significance (LR test statistic =47.02; 1% critical value = 29.14). The same applies for the Pitt and Lee (1981) model, which restricts inefficiency to be time invariant for all firms (LR test statistic =48.98; 1% critical value = 30.58). We therefore convincingly prefer the flexible time varying inefficiency model over the more restrictive alternatives. This preference also carries over into the two comparator models shown in Table 2.

Table 3 shows the results of the Battese and Coelli (1992) and Pitt and Lee (1981) models, together with two other alternative models (corrected ordinary least squares (COLS) and a time invariant model estimated by generalised least squares (GLS)). All models indicate that the British infrastructure manager faces a substantial cost efficiency gap as compared to its European peers. We also estimated the "true" fixed and random effects model and the Battese and Coelli (1995) models. We experienced convergence problems with the true random effects[13] and Battese and Coelli (1995) models[14], a problem that has been reported previously in the literature, particularly for complex specifications (see, for example, Farsi and Filippini, 2008; Oum, Yan and Yu, 2006; and Shehu et. al., 2007).

Returning to the preferred model, the parameters that capture the time varying path of inefficiency for Network Rail are statistically significant, so that the null hypothesis of no variation in efficiency for Network Rail over time can be rejected at the 1% level in the preferred

[13] The true fixed effects model results in the "wrong skew" error message, indicating model mis-specification (this being unsurprising, given the a prior expectation that inefficiency does not vary randomly from year to year).

[14] As compared to our preferred model, the Battese and Coelli (1995) requires double the number of parameters to be estimated (a firm-specific dummy and a dummy interacted with the time trend needs to be included in the list of variables influencing the mean inefficiency), so the problem of non-convergence is unsurprising.

model (based on an LR test). During the early years after privatisation, Figure 2 shows Railtrack's efficiency improving modestly (even after the steady-state adjustment). This accords with the expectation that privatisation would be expected to produce some efficiency gains, although the   parameter for Railtrack is not statistically significant.

Table 3: Efficiency Scores from Comparator Models

Dependent Variable - Total Costs (Steady-state Adjusted)	Time varying[1] COLS	Time invariant[2] GLS	Time invariant[2] Pitt and Lee 1981	Time varying Battese and Coelli 1992	Time varying Preferred model
Network Rail score 2006	0.406	0.584	0.646	0.705	0.597
Network Rail rank in 2006 (out of 14)	12	11	7	7	10
Rank correlation (all firms in 2006) against preferred model	0.929	0.830	0.709	0.396	1.000
LR test (preferred model against restricted model)			48.98***	47.02***	

1 The COLS score against upper quartile is 0.630

2 For these models, the 2006 score is the same as for all other years (time invariant efficiency model).

The model shows that efficiency then deteriorated sharply following the Hatfield accident, Railtrack's period in administration, and the company's subsequent replacement by Network Rail; before starting to improve following the 31% efficiency target set for the company at the 2003 interim review (starting from 2004 onwards). As noted in the previous paragraph, this variation in efficiency over time is found to be statistically significant. The profile of efficiency change over time for Railtrack / Network Rail therefore appears plausible, given the changes in underlying costs over the period as well as other evidence (see section 2 and 5.2). It should be noted that in the preferred model, Network Rail's renewal costs have been adjusted to take account of fluctuations around steady-state renewal volumes, so the inefficiency gap at the end is additional to cost increases resulting from rising renewal volumes above their steady-state levels.

In respect of the pattern of efficiency variation over time for Network Rail, of the two other time varying efficiency models in Table 3, the COLS model produces a similar profile to that of the preferred model (see Figure 3); though with lower absolute scores as expected, since this model does not distinguish noise from inefficiency. By contrast, the Battese and Coelli (1992) model shows efficiency always improving for most of the period, which is not a credible result given the doubling of the company's costs over the period and previous evidence suggesting a deterioration in efficiency over this period (see section 2). This finding gives further support for our preferred model, since the restricted model, which forces the same direction of efficiency change for all firms, produces counterintuitive results.

Figure 2: Profile of Network Rail Efficiency Scores: Preferred Model

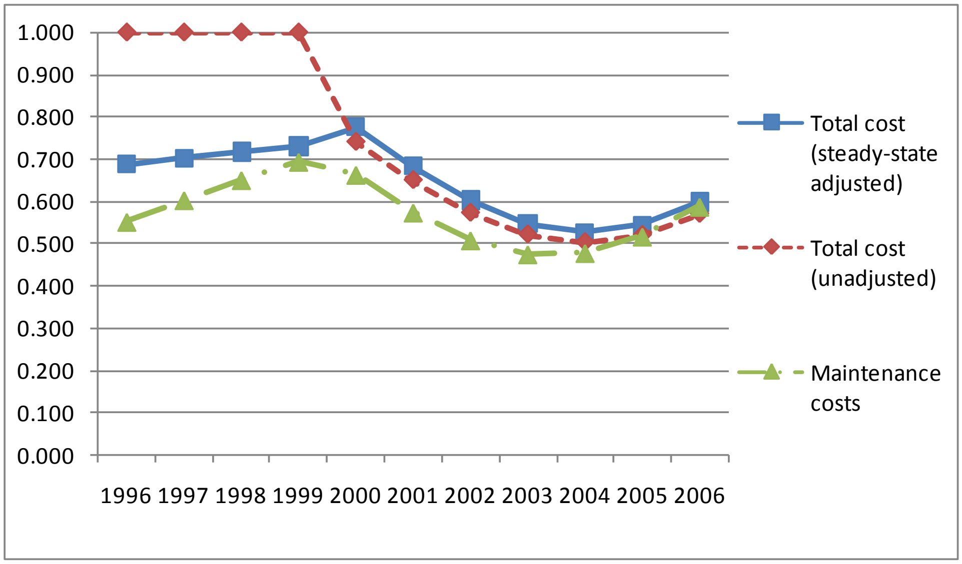
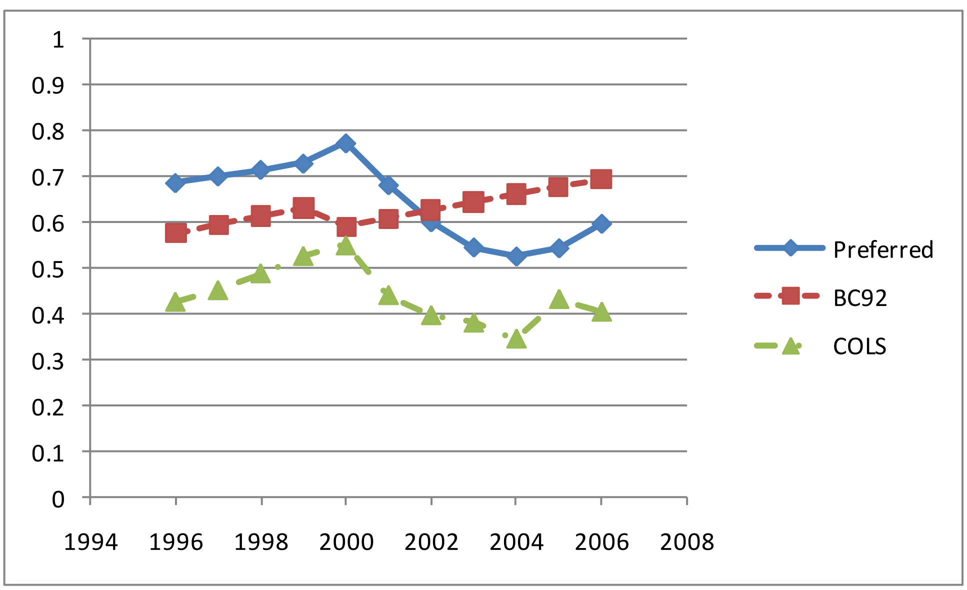


Figure 3: Network Rail Efficiency Score Profiles: Preferred and Comparator Models



A slightly different picture emerges for the total cost model without the steady-state adjustment, with efficiency being above that of the preferred model during the early years (when renewal volumes were argued to be below steady-state levels), and marginally below that of the preferred model in the later years (when renewal volumes were argued to be above steady-state levels). The deterioration in efficiency is also more pronounced and occurs earlier in the adjusted model. These results are intuitive, since the steady-state adjustment in the preferred model increases

Network Rail's costs compared to the raw data in the early years, and reduces them in the later years when renewal volumes were stepped up.

The maintenance-only model again shows a similar temporal pattern of efficiency variation and a very similar efficiency gap for Network Rail in 2006; though with a more substantial improvement in efficiency performance in the early years as maintenance costs were brought down after privatisation. Maintenance, which includes the cost of inspections, tamping and minor repairs, is naturally less prone to annual fluctuations as a result of changes in the volume of activity, since inadequate maintenance activity would be likely to feed through into reduced operational performance relatively quickly. The finding that the maintenance-only model produces similar results to the total cost model (with steady-state adjustment) thus suggests that the steady-state adjustment carried out to renewal costs prior to estimation of the preferred (total cost) model has produced a sensible model in terms of the efficiency results for Network Rail.

Of course, it has been argued within the European rail industry that Railtrack was not carrying out sufficient maintenance during the early years after privatisation, and that maintenance had to be stepped up accordingly. Nevertheless, even during the Railtrack years, the company was still substantially away from the frontier in terms of the efficiency of its maintenance expenditure. Further, based on comparison with its peers, our model indicates that in 2006 Network Rail was spending considerably more than suggested by best practice elsewhere in Europe.

To complete this section, we consider the precision of the firm-specific efficiency estimates for Network Rail. Horrace and Schmidt (1996) derived confidence intervals for the firm-specific efficiency estimates resulting from stochastic frontier models (or more precisely, prediction intervals; see Coelli et. al., 2005). These intervals take account of the uncertainty surrounding the decomposition of the composite error into its noise and inefficiency components[15]. In any efficiency study it is important to consider the precision of the efficiency estimates derived from the model, and perhaps even more so in a regulatory context, where the results will be used to set efficiency targets for the regulated firm.

In the 2008 Periodic Review, the efficiency score for Network Rail resulting from the preferred model was used as a central piece of evidence in the regulator's efficiency determination for the company. The efficiency score in 2006, the final year of the study, is therefore the one of most interest. Table 4 below shows the prediction intervals for Network Rail's efficiency score in 2006. These are computed using the general formula for panel models in the (econometric software) LIMDEP 9.0 Manual[16]. Table 4 shows that the prediction interval is fairly tight around the standard Jondrow et. al. (1982) efficiency estimate resulting from the preferred model, giving further confidence in the use of this model as a key piece of evidence in setting efficiency targets for Network Rail as part of the regulatory review process.

[15] As Horrace and Schmidt (1996) note, these intervals do not take into account uncertainty concerning the estimated parameters of the model

[16] See page E33-58. The code is available from the author on request.

Table 4: Prediction Intervals for Network Rail's Efficiency Score in 2006

	Lower Confidence Bound 95%	Efficiency Estimate (Preferred Model)	Upper Confidence Bound 95%
Network Rail (2006)	0.550	0.597	0.648

5.3 Other evidence and regulatory interpretation

It is standard regulatory practice in the UK to adopt a wide range of methods to assess the efficiency performance of regulated companies. These may be top-down econometric methods, as described above, broader top-down methods based on analysing trends in unit operating costs in comparator industries, or bottom-up approaches. The latter utilise engineering and management expertise to determine the specific initiatives that could be implemented to improve efficiency. Thus, top-down and bottom-up benchmarking complement each other, with the former identifying the size of the efficiency gap, and the latter seeking to explain the reasons for the gap and how it could be closed.

ORR used the results of the preferred econometric model presented in this paper as a central piece of evidence in its final efficiency determination for Network Rail in 2008. However, in line with best practice, ORR commissioned a wide range of other studies and concluded that the finding of the international econometric study, which showed a substantial gap against the frontier for Network Rail, was supported by other evidence (both top-down and bottom-up; see ORR, 2008 and Table 5).

Following the practice of other regulators, ORR also applied the results conservatively, with the final efficiency target for maintenance and renewal costs being set at 22%, as compared to the 40% efficiency gap indicated by the preferred econometric model. Network Rail challenged the findings throughout the process, but ultimately accepted the final determination and did not exercise their right to appeal to the Competition Commission. The experience of this review, and efficiency reviews in other regulated industries, demonstrates the importance of combining top-down studies with bottom-up evidence, in order to ensure the credibility of the results, and its acceptance by regulators and regulated firms.

Table 5: 2008 Periodic Review Efficiency Studies

Study	Key finding
International econometric study (this paper)	40% gap against the frontier 37% gap against upper quartile
ORR final determination based on the above	37% gap against upper quartile 22% to be closed over the 5 year control period
Network Rail's own estimates	17.6% over 5 years
Ernst and Young review of Network Rail estimates	"Does not believe it is unreasonable to expect that the 17.6% total efficiency target could be increased"
BSL gap analysis	Adjusted gap of 27% to 44% against the average elsewhere in Europe
RailKonsult study of technology and working practices	Technology / working practices from overseas could reduce costs substantially (e.g. inspection costs by 75%; re-railing costs by 40%)
LEK/TTCI comparisons against North America	US Railroads average efficiency gains since 1980 de-regulation: 4-5% per year (implies 18-20% over 5 years) US average costs between 3.3 and 5.1 times lower than Network Rails (post adjustments)
Lloyds Register Rail review of track renewals efficiency	Savings of 33% attainable on track renewal unit costs through new methods Savings of around 40% on switches and crossing renewals possible
OXERA study on opex efficiency (trends from other sectors)	Savings of 4-6.5% per year, or 18-28% over five years
Network Rail track productivity benchmarking	Network Rail more productive for both conventional and "high output" track renewals
ORR international visits	"Evidence of potentially more effective and/or efficient practices in other countries"
AMCL asset management benchmarking	"..could deliver significant savings in both capital and operational expenditure"
Abbot review (Canadian railway engineer)	"..there is tremendous scope for improvement in productivity"
Lloyds Register Rail study of possessions practices	"…many areas where overseas practice is more efficient than Britain"

Source: ORR (2008)

6 Conclusions

The contributions of this paper are as follows.

1. It is the first paper in the literature to apply stochastic frontier techniques to study the relative efficiency performance of rail infrastructure managers in Europe. The results in respect of the British rail infrastructure manager, Network Rail, indicate that in 2006 the company faced a cost efficiency gap of around 40% against European best practice. To our knowledge the paper is also the first to produce prediction intervals for firm-specific

efficiency scores for the Cuesta (2000) model adopted. The 95% prediction interval is shown to be fairly tight around the standard Jondrow et. al. (1982) efficiency estimate, giving a range for Network Rail's efficiency gap of between 35% and 45%.

2. The paper demonstrates that stochastic frontier panel data techniques can be used by economic regulators to determine the efficiency of regulated firms against international best practice. The results formed the basis of ORR's 2008 efficiency determination for Network Rail, covering some £27 bn (c. €33 bn) of expenditure, and the findings were supported by a wide range of other, engineering-based evidence. In general, economic regulators have adopted relatively simple efficiency modelling approaches, and we are not aware of any that have applied techniques of the kind adopted here. The study utilised a unique, panel dataset that had already been used by the International Union of Railways for benchmarking purposes, thus enabling ORR to place international benchmarking – using econometric techniques - at the forefront of its efficiency determination.

3. The paper highlights the desirable properties of the preferred, time varying stochastic frontier model adopted, which are important in a regulatory context. The model allows inefficiency to vary over time, whilst permitting firm-specific time paths of inefficiency. In an international benchmarking context, this flexibility is particularly important, since firms are operating under different regulatory regimes. The flexibility allowed by the model is shown to be important in this case, with the restrictions implied by less flexible alternatives, in particular the Battese and Coelli (1992) model, being clearly rejected based on statistical testing. Further, the model enables the unrealistic assumption of independence in inefficiency over time (a problem that plagues many comparator models; see Alvarez et. al., 2006) to be relaxed, and for the temporal variation over time to be statistically tested.

In recent years rail regulators have emerged across Europe, as countries have implemented the requirements of European legislation. To date, however, outside Britain, these regulatory bodies have been chiefly concerned with the issue of ensuring non-discriminatory access for train operators to the rail infrastructure. Looking forward, there are signs that other European regulators are taking more of an interest in infrastructure manager efficiency, although they do not yet have the powers[17]. Further, transport ministries in any case have a strong interest in efficiency, as the high cost of rail is coming under closer scrutiny; and European legislation requires pressure to be applied on costs, whether by a regulator, or the ministries.

Given these trends, and the lack of domestic comparators for rail infrastructure, we consider that international benchmarking will take on an increased importance in the (European) sector in the future. This paper has demonstrated the power of stochastic frontier panel data techniques as a key measurement tool to be adopted by regulatory or policy bodies.

What challenges remain? First, it is important to continue to collect good quality, comparable data, including extending existing data sources over time. In 2010 the British Infrastructure

[17] See the proceedings of the Florence School of Regulation Workshop on the Aims, Models and Powers of Rail Regulators (November 2010). http://www.florence-school.eu/portal/page/portal/FSR_HOME/Transport/Policy_events/Workhops/2010/Rail_Regulation

manager and regulatory body started the ongoing process of updating the dataset and carrying out further checks to verify data comparability. Second, obtaining more variables that capture heterogeneity between countries is important. In particular, the collection of measures of the condition and age of the infrastructure, and its performance, would enhance existing models. Extending the steady-state adjustment more formally to other countries would be a useful development if new variables cannot be collected. Finally, future research could focus on the impact of different institutional arrangements (e.g. different regulatory regimes; vertical integration versus vertical separation) on the efficiency of rail infrastructure managers in Europe.

Acknowledgements

This work was funded partly by the British Office of Rail Regulation (ORR) and partly by Rail Research UK funding provided by the UK Engineering and Physical Sciences Research Council. I would like to acknowledge the valuable contributions of numerous people at ORR, including Paul McMahon, Hannah Nixon, Greg Smith and Gian Carlo Scarsi who contributed to the specification of the study and commented on the results; as well as many others who contributed, but are too numerous to mention individually. I would also like to thank Gerard Dalton of the International Union of Railways (UIC), Network Rail and the other infrastructure managers who provided data during the course of this study, as well as Phill Wheat for his technical comments on this research. Finally I acknowledge the comments of three anonymous referees. All remaining errors are those of the author.

References

Accenture, 2003. Review of Network Rail's Supply Chain. London.

Alvarez, A., Amsler, C., Orea, L. and Schmidt, P., 2006. Interpreting and testing the scaling property in models where inefficiency depends on firm characteristics. Journal of Productivity Analysis 25, 201-212.

Andrikopoulos, A. and Loizides, J, 1998. Cost Structure and Productivity Growth in European Railway Systems. Applied Economics 30, 1625-1639.

Battese, G.E. and Coelli, T.J., 1992. Frontier Production Functions, Technical Efficiency and Panel data: With Application to Paddy Farmers in India. Journal of Productivity Analysis, 3, 153-169.

Battese, G.E. and Coelli, T.J., 1995. A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data. Empirical Economics 20, 325-332.

Beesley, M. and Littlechild, S., 1988. The Regulation of Privatized Monopolies in the United Kingdom. RAND Journal of Economics 20, 454-72.

Booz-Allen and Hamilton, 1999. Railtrack's Expenditure Needs 2001-2006: A Report to the Office of the Rail Regulator. London.

Booz-Allen and Hamilton, 2000. Response to Railtrack's May 2000 Cost Submission: A Report to the Office of the Rail Regulator. London.

Caves, D.W., Christensen, L.R., Tretheway, M.W. and Windle, R.J., 1985. Network Effects and the Measurement of Returns to Scale and Density for U.S. Railroads. In: Daughety, A.F. (Eds.), Analytical Studies in Transport Economics, Cambridge, Cambridge University Press, pp. 97-120.

Cantos, P., Pastor, J.M. and Serrano, L., 2010. Vertical and Horizontal Separation in the European Railway Sector and its Effects on Productivity. Journal of Transport Economics and Policy, 44 (2), 139-160.

Coelli, T.J., Rao, D.S.P, O'Donnell, C.J. and Battese, G.E., 2005. An Introduction to Efficiency and Productivity Analysis, 2nd edition. New York, Springer.

Cornwell, C., Schmidt, P. and Sickles, R.C., 1990. Production Frontiers With Cross-Sectional And Time-Series Variation in Efficiency Levels. Journal of Econometrics 46, 185-200.

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, 139-158.

Farsi, M. and Filippini, M, 2008. An Analysis of Cost-Efficiency in Swiss Multi-utilities. CEPE Working Paper, No. 60.

Farsi, M., Filippini, M. and Greene, W., 2006. Application of Panel Data Models in Benchmarking Analysis of the Electricity Distribution Sector. Annals of Public and Cooperative Economics 77 (3), 271-290.

Frontier Economics, 2010. RP-X@20: The future role of benchmarking in regulatory reviews. A final report prepared for OFGEM. London.

Gathon, H.J. and Perelman, S., 1992. Measuring Technical Efficiency in European Railways: a Panel Data Approach. The Journal of Productivity Analysis 3, 135-151.

Greene, W.H., 1980. Maximum Likelihood Estimation of Econometric Frontier Functions. Journal of Econometrics 13, 27-56.

Greene, W.H., 2005.Reconsidering Heterogeneity in Panel data Estimators of the Stochastic Frontier Model. Journal of Econometrics 126, 269-303.

Greene, W. H., 2008. The Econometric Approach to Efficiency Analysis. In: Fried, H. O., Lovell, C. A. K., Schmidt, S. S. (Eds.), The Measurement of Productive Efficiency and Productivity Growth, Oxford University Press, New York.

Horrace, W.C. and Schmidt, P., 1996. Confidence Statements for Efficiency Estimates from Stochastic Frontier Models. Journal of Productivity Analysis 7 (2/3), 257-282.

Kennedy, J. and Smith, A.S.J , 2004. Assessing the Efficient Cost of Sustaining Britain's Rail Network: Perspectives Based on Zonal Comparisons. Journal of Transport Economics and Policy, vol. 38 (2), 157-190.

Kumbhakar, S.C., 1990. Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency. Journal of Econometrics 46, 201-211.

Kumbhakar, S., and A. Heshmati, 1995. Efficiency Measurement in Swedish Dairy Farms: An Application of Rotating Panel Data,1976-88. American Journal of Agricultural Economics 77 (3), 660-674.

Jondrow, J., Lovell, C.A.K., Materov, I.S. and Schmidt, P., 1982. On Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model. Journal of Econometrics 19, 233-238.

Lee, Y.H., 2006. A stochastic production frontier model with group-specific temporal variation in technical efficiency. European Journal of Operational Research 174, 1616-1630.

Lee, Y.H. and Schmidt, P., 1990. A Production Frontier Model With Flexible Temporal Variation in Technical Efficiency. In: Fried, H.O., Lovell, C.A.K. and Schmidt, S.S., (Eds.), The Measurement of Productive Efficiency, New York, Oxford University Press.

LEK, TTCI, Halcrow, 2003. Bottom-up review of Network Rail's business plan: 2003/04-2005/06. Report to ORR. London.

Nash, C.A., 1985. European Railway Comparisons – What Can we Learn? In: K.J Button and D.E. Pitfield (Eds.), International Railway Economics, Aldershot, Gower, pp. 237-269.

NERA, 2000. Review of Overseas Railway Efficiency: A Draft Final Report for the Office of the Rail Regulator. London.

OFGEM, 2009. Electricity Distribution Price Control Review Methodology and Initial Results Paper. London.

Office of the Rail Regulator, 2000. The Periodic Review of Railtrack's Access Charges: Final Conclusions. London.

Office of Rail Regulation, 2008. Periodic review of Network Rail's outputs and funding for 2009-2014. London.

Oum, T.H., Yan.J. and Yu, C., 2006. Ownership Forms Matter for Airport Efficiency: Results from Bayesian Estimation of Stochastic Cost Frontiers for Worldwide Airports. Mimeo.

Pitt, M.M. and Lee, L.F., 1981. Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry. Journal of Development Economics 9,43-64.

Pollitt, M.G. and Smith, A.S.J, 2002. The Restructuring and Privatisation of British Rail: Was it Really that Bad? Fiscal Studies, vol. 23 (4), 463-502.

Schmidt, P. and Sickles, R.C., 1984. Production Frontiers and Panel Data. Journal of Business & Economic Statistics 2 (4), 367-374.

Shehu, J.F., Mshelia, S.I. and Tashikalma, A.K., 2007. Analysis of Technical Efficiency of Small-scale Rain-fed Upland Rice Farmers in North-west Agricultural Zone of Adamawa State, Nigeria. Journal of Agriculture and Social Sciences 3 (4), 133-136.

Smith A.S.J., 2005. The Role of Efficiency Estimates in UK Regulatory Price Reviews: The Case of Rail. Utilities Policy 13 (4), 294-301.

Sumicsid, 2007. Project Gerner / AS6: Development of Benchmarking Models for German Electricity and Gas Distribution: Final Report.

UIC, 2007. Lasting Infrastructure Cost Benchmarking (LICB): 10 years of Benchmarking 1996-2005.

Wheat, P. and Smith, A., 2008. Assessing the Marginal Infrastructure Maintenance Wear and Tear Costs for Britain's Railway Network. Journal of Transport Economics and Policy, 42(2), 189-224.

Wheat, P., Smith, A. and C. Nash, 2009. CATRIN (Cost Allocation of TRansport INfrastructure cost), Deliverable 8 - Rail Cost Allocation for Europe. Funded by Sixth Framework Programme. VTI, Stockholm.

Viscusi, W.K., Harrington, J.E. and Vernon, H.M., 2005. Economics of Regulation and Antitrust, MIT Press, Cambridge, Massachusetts.