7TH WCTR
SYDNEY AUSTRALIA
1995

TOPIC 14
PERFORMANCE
MEASUREMENT

ASSESSING PUBLIC SECTOR PERFORMANCE: THE CASE OF WESTERN EUROPEAN RAILWAYS

JOHN M. PRESTON
Institute for Transport Studies
University of Leeds
Leeds LS2 9JT UK

Abstract

The need for studies of public sector performance is outlined with specific reference to European railways and three analytical techniques are reviewed. Using the results from 12 different measures, it is found that there is considerable variation in the performance of Western Europe's railways over time, at any particular time and across methodologies.

INTRODUCTION

The problems of assessing the performance of publicly owned, non-profit making organisations are well known. Of the three key indicators of a business's success—productivity, profitability and rate of returns on assets—only productivity is usually relevant. Moreover, in a paper to the last World Conference (Nash and Preston 1992) we highlighted the specific problems that arise in comparing the productivity of the publicly owned, non-profit making railways of western Europe. It is not the aim of this paper to restate these problems. It is sufficient to say that these problems remain and must be borne in mind when analysing performance.

It is the aim of this paper to compare a number of different approaches to productivity performance with respect to western Europe's railways. In the next section, we will outline three broad approaches to productivity assessment, two of which we will examine in more detail in this paper. Then, in the following 2 sections, we develop non-parametric index numbers of partial factor productivity and total factor productivity respectively. We then develop a parametric factor productivity model based on the translog cost model. In the final section, we compare our different measures in terms of the ranking of the 14 railway companies in our data set.

APPROACHES TO MEASURING PERFORMANCE

A useful, recent review is provided by Hensher and Waters (1993). They identify three broad approaches: non-parametric, index number approaches; parametric factor productivity model estimation; and non-parametric Data Envelopment Analysis. We shall discuss these approaches in turn.

Non-parametric, index number approach

This is the simplest approach and one that is in vogue given the current popularity of benchmarking (see for example Boxwell 1994). Partial Factor Productivity (PFP) indices are simply ratios of aggregate output per unit of partial input. To the extent that increases in productivity with respect to one input (eg labour) can be achieved at the expense of reducing the productivity of other inputs (eg capital, fuel), PFPs will inaccurately portray the gain/loss in productivity (Talvitie and Obeng 1991). On the other hand, they have the advantage of being transparent and readily understandable by policy makers.

Total factor productivity (TFP) indices are the amount of aggregate output produced per unit of aggregate input. They are comprehensive measures that have been used extensively in recent work in Australia (Hensher, Daniels and DeMellow 1992). Dodgson (1985) highlights a number of problems. These include the definition of the input and output weights (and in particular whether they should be based on the before or after situation), the determination of whether productivity gains are due to reductions in technical inefficiency or due to external technical progress or scale effects and whether the combination of inputs is cost minimising.

Parametric factor productivity models

TFP and PFP measures can be estimated from neo-classical cost functions which can isolate the impacts of economies of scale and external factors from improvements in technical efficiency (see, for example, Obeng 1985). Early American econometric work on the rail industry was based on the Cobb-Douglas production function (Keeler 1974) or a linear cost function (Harris 1977), although subsequently the translog cost model has come to dominate the literature (eg Caves et al. 1985, Friedlander et al. 1993). The main drawback with these neoclassical approaches is that they assume cost-minimising behaviour which may not be appropriate for some railway firms. One approach has been to develop stochastic cost and production frontier functions (eg Cornwell, Schmidt and Sickles 1990). In this approach, the error term to a cost or production function has two parts, a non-negative random variable capturing the impact of inefficiency and a two-sided variate accounting for exogenous shocks.

Non parametric data envelopment analysis

Another approach, Data Envelopment Analysis (DEA), estimates a production function using linear programming techniques. The main practical advantage of DEA techniques is that they do not require multilateral price indices for inputs and outputs, whilst their main theoretical advantage is their greater flexibility than approximations such as the translog function. The main disadvantage is that DEA is non-parametric. There is no error theory and hence no tests of statistical significance. Results can be sensitive to outliers. Tobit regression is often used with DEA in order to decompose the productivity measure in order to account for systematic sources of productivity differences. A similar ad hoc procedure, using ordinary least squares regression, is often used to decompose the variation in TFP determined from a non parametric index number approach. DEA and tobit regression have recently been used in a study of the economic efficiency of OECD railways (Oum and Yu 1994).

PARTIAL FACTOR PRODUCTIVITY ANALYSIS

The background to our work in this area is as follows. The Institute for Transport Studies (ITS), University of Leeds and the British Railways Board (BRB) carried out a comparative study of western European railways in the late 1970s (BRB and University of Leeds 1979). Follow-up work was carried out by ITS in work financed by the Social Science Research Council and reported by Nash (1985). In 1992, ITS was commissioned by BRB to reactivate this work for 14 western European railways for the base year of 1990. The results are reported in detail by Preston et al. (1994). The approach adopted by these studies was essentially to develop a series of indices as follows:

$$\frac{\text{Receipts}}{\text{Traffic unit}} \times \frac{\text{Traffic unit}}{\text{Train kms}} \times \frac{\text{Train kms}}{\text{Staff nos}} \times \frac{\text{Staff nos}}{\text{Staff costs}} \times \frac{\text{Staff costs}}{\text{Total costs}} = \frac{\text{Receipts}}{\text{Total costs}}$$

Of these indices, Train Kms/Staff Nos was classified as the key measure of operating performance, Receipts/Traffic Unit and Traffic Unit/Train Km were measures of commercial performance and Receipts/Total Costs was the key measure of financial performance. Staff No/ Staff Costs and Staff Costs/Total Costs were regarded as being largely determined by factor prices. The railways studied in our latest work (and their abbreviations) are listed in Table 1 and the results for our key indicators given in Table 2.

In terms of commercial performance as measured by receipts per traffic unit, BR has the highest rates at 5.8 pence per traffic unit km and CP and SJ have the lowest rates at 2.1 and 2.2 pence per traffic respectively unit adjusted using purchasing power parity rates). This in turn explains the relatively low market share of railways in Britain and the high market share (at least for freight) in Sweden (see Table 3). Moreover, the use of a homogenous traffic unit is misleading. European railways vary greatly in their mix of output. For example, at one extreme 78% of NS's traffic units are passenger kms, whilst, at the other extreme, the corresponding figure for SJ is 24%. SJ's low receipts per traffic unit are due to very low freight rates (around one pence per tonne km) which in turn are due to product mix (low value products such as timber and iron ore are important) and lengths of haul. In fact, at over 5 pence per passenger km, SJ has relatively high receipts per passenger carried.

Table 1 Railways Included in the study

Acronym	Name	Country
BR	British Rail	Great Britain
CFF	Chemins de Fer Federaux Suisses	Switzerland
CIE	Coras Iompair Eireann	Eire
CP	Caminhos de Ferro Portugueses	Portugal
DB	Deutche Bundesbahn	West Germany
DSB	Danske Statsbaner	Denmark
FS	Ente Ferrovie dello Stato	Italy
NS	Nederlandse Spoorwegen	Netherlands
NSB	Norges Statsbaner	Norway
OBB	Osterreichische Bundesbahn	Austria
RENFE	Red nacional de los Ferrocarriles Espanoles	Spain
SNCB	Societe Nationale des Chemins de fer Belges	Belgium
SNCF	Societe Nationale des Chemins de fer Francais	France
SJ/BV	Statens Jarnvager/Banverket	Sweden

Table 2 Key indicators for 14 European railways (1990)

	Receipts/ Traffic Unit (pence per km)	Traffic Unit/ Train km	Train km/ Staff Nos	Staff Costs/ Staff Nos (£)	Staff Cost/ Total Cost	Receipts/ Total Costs
BR	5.8	113.97	3193	15054	0.59	0.82
CFF	3.9	158.11	3033	21197	0.57	0.51
CIE	3.7	127.48	2693	12804	0.48	0.45
CP	2.1	175.21	1857	10362	0.52	0.34
DB	4.3	173.76	2559	26296	0.61	0.44
DSB	4.3	119.27	2709	13360	0.43	0.45
FS	2.4	212.34	1568	21332	0.44	0.16
NS	3.2	120.45	4484	18711	0.50	0.46
NSB	3.5	127.29	2504	13596	0.60	0.49
OBB	3.4	181.48	1750	14935	0.49	0.35
RENFE	2.6	170.64	3459	19473	0.54	0.42
SJ/BV	2.2	249.23	3501	14844	0.45	0.59
SNCB	3.0	96.94	3402	24591	0.68	0.27
SNCF	3.4	234.66	2413	18729	0.49	0.50

Note:
Traffic Unit = Passenger Km and Freight Tonne Km

Table 3 Rail market shares (1990%)

	Passenger	Freight	Average	Ranking
Great Britain	5.4	9.9	7.7	11
Switzerland	10.8	41.6	26.2	2
Republic of Ireland	3.6	10.3	7.0	13
Portugal	7.0	12.7	9.9	8
Germany	6.3	20.6	13.5	6
Denmark	7.1	16.0	11.6	7
Italy	7.1	10.1	8.6	10
Netherlands	6.9	4.6	5.8	14
Norway	5.1	14.3	9.7	9
Austria	11.1	46.1	28.6	1
Spain	7.6	7.0	7.3	12
Sweden	6.1	42.5	24.3	3
Belgium	10.1	17.8	14.0	5
France	9.2	26.7	18.0	4

High receipts per traffic unit are associated with low loadings (traffic units per train km) and vice versa. BR has the second lowest loading of the 14 railways studied, with only SNCB having lower. SJ has the highest loading, followed by SNCF and FS. However, aggregating passenger and freight traffics again masks important differences. BR and SNCB's low loadings are due to low passenger loadings which in turn are related to these railways operating relatively short, frequent passenger trains. SJ's high overall loading is due to its freight business; its passenger loadings are relatively low (around 100 passenger km per train km) and are surpassed by SNCF and FS by a considerable margin (at around 200 passenger km per train km).

In terms of operating performance, as measured by train km per staff, there is huge variation. Staff in the most productive railway (NS) produce over 2.5 times the output of staff in the least productive railway (OBB). However, this masks a number of important differences. Firstly, product mix again influences this index as freight railways, particularly those carrying large volumes of general merchandise, may be expected to be more labour intensive than passenger railways due to additional labour requirements for loading/unloading, marshalling etc. Secondly, the number of hours worked per member of staff varies, particularly between Great Britain and Ireland (48 hours per week on average) and continental Europe (where 38 hours per week is the norm). There is also great variation in the level of skill of railway labour. Thirdly, there is much variation in the mean length of passenger and freight trains operated. Fourthly, low labour productivity may be offset to some extent by highly productive use of other inputs such as fuel and equipment. Fifthly, the results for some railways may have been distorted by the use of contract labour.

In terms of annual staff costs per member of staff, DB workers have the highest annual salary with a mean of £26,362, whilst CP workers have the lowest mean annual salary at £10,362 (adjusted using purchasing power parity rates). The average hourly salary cost ranges from £17.46 an hour (DB) to £4.44 (CP). In terms of staff costs as a proportion of total costs, the highest figure is recorded by SNCB (68%) and the lowest by DSB (43%). In part, this reflects the treatment of capital costs. Depreciation (based on historic costs) and interest account for only 6% of BR's costs but account for 29% of DSB's costs.

In terms of financial performance, as measured by receipts divided by total costs, BR had by far the highest cost recovery ratio at 82% and FS the lowest at 16%. The measure in Table 2 is based on data for the rail business only. If data from the non rail business are included, the mean cost recovery ratio for our sample of 14 firms increases from 46% to 63%, indicating that non rail business are generally much more profitable than rail.

Table 4 compares our indicators for operating (train km per staff member), commercial (market share) and financial (receipts divided by total costs) performance in 1990 with those for 1977 (1976 for market share data) for nine railways. Staff productivity has increased by 27% on average or 1.8% per annum, with the highest growth being for DB and SNCB and the lowest growth for FS and NSB. Rail's share of the passenger market had decreased by 12% (1.0% per year) and of the freight market by 20% (1.7% per year), with the rail freight market also declining in absolute terms (Preston 1994a). Rail's financial performance has worsened, with the average cost recovery ratio decreasing from 59% to 46%, with only BR showing an improvement in its finances. The picture portrayed by Table 4—sluggish productivity growth and declining market shares and cost recovery ratios—helps explain the widespread belief amongst policy makers that European railways require drastic reform. What Table 4 and the other information presented in this section is less useful in doing is prescribing appropriate remedies, we shall return to this issue later.

Table 4 Trends in key indicators between 1977 and 1990 for selected western European railways

		Train-km per member of staff		Market share freight (%)		Market share passenger (%)		Receipts/ costs %	
		1977	1990	1976	1990	1976	1990	1977	1990
Netherlands	NS	3909	4484	4.9	4.6	6.4	6.9	56	46
Sweden	SJ/BV	2830	3501	44.6	42.5	5.4	6.1	83	59
Belgium	SNCB	1800	3402	22.1	17.8	11.4	-	50	27
Great Britain	BR	2417	3193	16.8	9.9	6.5	5.4	71	82
Denmark	DSB	2242	2709	15.0	16.0	7.3	7.1	61	45
W Germany	DB	1750	2559	26.1	20.6	6.4	6.3	61	44
Norway	NSB	2267	2504	23.2	14.3	5.6	5.1	60	49
France	SNCF	2096	2413	34.1	26.7	11.0	9.2	55	50
Italy	FS	1411	1568	18.2	10.1	12.1	7.1	32	16
Mean		2302	2926	22.7	18.1	7.6	6.7	59	46

TOTAL FACTOR PRODUCTIVITY ANALYSIS

Previously we briefly highlighted some of the problems associated with non parametric, total factor productivity measurement. A theoretically attractive index which has been widely adopted is the translog multilateral productivity index proposed by Caves, Christensen and Diewert (1982) which allows comparisons based on cross sectional, time-series or pooled data. The index is defined as:

$$\ln\left[\frac{TFP_k}{TFP_b}\right] = \frac{1}{2}\sum_i (R_{ki}+\overline{R_i})(\ln Y_{ki}+\overline{\ln Y_i}) - \frac{1}{2}\sum_i (R_{bi}+\overline{R_i})(\ln Y_{bi}+\overline{\ln Y_i})$$
$$- \frac{1}{2}\sum_n (W_{kn}+\overline{W_n})(\ln X_{kn}+\overline{\ln X_n}) + \frac{1}{2}\sum_n (W_{bn}+\overline{W_n})(\ln X_{bn}+\overline{\ln X_n})$$

where

k = each individual observation, $k=1, \ldots, K$

b = base observation

i = outputs, $i=1, \ldots, I$

n = inputs, $n=1, \ldots, N$

R_i = weights for each output

$\overline{R_i}$ = arithmetic mean of output weights

W_n = weights for each input

$\overline{W_n}$ = arithmetic mean of input weights

$\ln Y_i$ = unit measure of output

$\overline{\ln Y_i}$ = geometric mean of unit measure of output

$\ln X_n$ = unit measure of input

$\overline{\ln X_n}$ = geometric mean of unit measure of input.

The preferred input weights are the cost shares and the preferred output weights are the elasticities of cost with respect to outputs. In practice, the absence of such elasticities has led to the use of revenue shares as proxies. This is strictly a valid assumption only where there are constant returns to scale across all outputs and all outputs are priced at marginal cost. Without knowledge of cost elasticities it is not possible to distinguish changes in TFP due to scale effects from other sources of productivity gain. We shall return to this issue later. In the meantime, in Table 5 we present the TFPs for our 14 railways based on two outputs (passenger and freight) and four inputs (labour, materials, fuel and capital). It should be noted that our data on capital costs is particularly crude, being based on data on historic cost depreciation and interest collated by the Union Internationale de Chemins de Fer (UIC) and does not fully take into account differing accounting conventions.

Table 5 Total Factor Productivity—based on revenue shares

	Supply Measure	Demand Measure
BR	1.00	1.00
CFF	1.07	1.21
CIE	0.83	0.74
CP	0.58	0.88
DB	1.06	1.17
DSB	0.83	0.78
FS	0.45	0.70
NS	1.25	1.24
NSB	1.35	1.12
OBB	0.64	0.71
RENFE	1.19	1.38
SU/BV	1.96	2.74
SNCB	0.93	0.61
SNCF	0.86	1.34

Two TFP measures are developed based on intermediate, supply based outputs (train kms) and final, demand based outputs (passenger kms and freight tonne kms). In terms of the supply based TFP measure our base railway, BR, is ranked seventh, with SJ being the top performer with a score of 1.96 and FS being the worst performer with a score of 0.45. In terms of the demand based TFP measure BR is ranked eighth, with SNCF overtaking it in the rankings. SJ remains the top performer with its score increasing to 2.74, whilst the worst performer is now SNCB with a score of 0.61.

TRANSLOG COST MODEL

In this section we present the results of a translog cost model developed for the 14 railways in our sample (plus the Finnish operator, VR) for the period 1971 to 1990, based principally on data published by the UIC. The model estimates total operating costs as a function of three input prices (labour, energy, materials) and three outputs (passenger train kms, freight train kms and length of route) and is described in more details in Preston (1994b). The model took the following form:

$$\begin{aligned}\ln RTC = {} & \alpha_0 + \sum_i \alpha_i \ln Y_i + \sum_j \beta_j \ln P_j + \frac{1}{2}\sum_i \sum_k \delta_{ik} \ln Y_i \ln Y_k \\ & + \frac{1}{2}\sum_j \sum_m \gamma_{jm} \ln P_j \ln P_m + \sum_i \sum_j P_{ij} \ln Y_i \ln P_j + \sum_n \theta_n D_n + \phi T + \varepsilon\end{aligned}$$

where

Y_{ik} = output measures

P_{jm} = factor prices

D_n = railway specific dummy variable

T = time trend

ε = error term.

For homogeneity of degree one in input prices, we require that the following restrictions be satisfied:

$$\sum_j \beta_j = 1; \sum_j \gamma_{jm} = \sum_m \gamma_{jm} = 0; \sum_i P_{ij} = \sum_j P_{ij} = 0$$

Input cost shares can then be derived using Shepard's lemma. In general:

$$W_j = \frac{P_jX_j}{C} = \frac{P_j\partial C}{C\partial P_j} = \frac{\partial \ln C}{\partial \ln P_j}$$

where

W_j = cost share of input i

X_j = quantity of input i.

So for the translog:

$$W_j = B_j + \sum_m \gamma_{jm}\ln P_m + \sum_i P_{ij}\ln Y_i$$

Given n factor prices, n-1 cost share equations may be estimated jointly with the translog cost function. The estimation method used was the Seemingly Unrelated Regression (Zellner 1962) procedure provided by the Statistical Analysis Systems Computer Package (SAS 1988). Statistical tests indicated that autocorrelation, heteroscedasticity and multicollinearity were not significant problems in the resultant model. The econometric problems that arise from the use of pooled data have been reduced by the use of firm-specific dummy variables and a time trend variable, and may be thought of as a form of the covariance model advocated by Pindyck and Rubinfeld (1991: 224). It should be noted that of the 57 parameter values estimated in the three model system, only 27 are significant at the 5% level. It should also be noted that we can not readily compute standard errors for the elasticity measures, but we would anticipate that they would be large.

Three key results from the translog model are presented in Table 6. By taking the exponential of the firm specific dummy variables, cost efficiency may be assessed relative to the base operator, RENFE. All other things being equal, seven railway companies would have the same costs as RENFE, one railway would have lower costs (SJ) and six would have higher costs (OBB, SNCB, CP, FS, CFF and DSB).

Table 6 Cost efficiency (1971 - 1990), returns to density (1990) and returns to scale (1990)

	Operators comparisons	Returns to Density	Returns to Scale	Train Km per annum (000)	Length of line (km)	Density (train km per line km)
BR	1.17*	1.11	0.72	445060	16584	26837
CFF	1.46	0.81	1.22	122394	2978	41099
CIE	0.86*	6.00	1.23	14237	1944	7324
DB	1.03*	1.33	0.64	603797	26949	22405
DSB	1.38	1.10	1.26	52160	2344	22252
FS	1.39	1.48	0.69	314255	16066	19560
NS	1.17*	0.77	1.25	117314	2798	41928
NSB	0.99*	4.44	1.02	36705	4044	9076
OBB	2.00	1.25	0.88	117201	5624	20839
SJ	0.70	6.04	0.76	99634	10801	9225
SNCB	1.82	1.00	1.10	92802	3479	26675
SNCF	0.73*	2.37	0.60	487670	34070	14313
VR	0.94*	53.83	0.89	41026	5867	6993
CP	1.65	1.98	1.05	33693	3064	10996
RENFE	1.00	2.08	0.73	168960	12560	13452

Note:
* not significantly different from 1.0 at the 5% level.

Returns to density were estimated as:

$$RTD = (\partial \ln RTC / \partial \ln TKT)^{-1}$$

where

RTC = Total operating costs

TKT = Total train kilometres.

Twelve of the fifteen railways exhibit increasing returns to density and this finding is particularly marked for the four railways with a traffic density of less than 10,000 train km per annum per km of line (VR, CIE, NSB and SJ). The Belgian railway (SNCB) exhibits constant returns to density, with a traffic density of almost 27,000 train km per km of line. The two most densely used rail networks in western Europe (NS and CFF, both with traffic densities of over 40,000 train km per km of line) exhibit decreasing returns to density, suggesting existing infrastructure is congested and that investment plans to expand rail capacity in both the Netherlands and Switzerland may be justified.

Returns to scale were estimated as:

$$RTS = \left(\frac{\partial \ln RTC}{\partial \ln TKT} + \frac{\partial \ln RTC}{\partial \ln LL} \right)^{-1}$$

where

LL = Length of line.

From Table 6, it can be seen that only seven of the fifteen railways exhibit returns to scale greater than one and of these NSB, CP and SNCB may be characterised as having broadly constant returns to scale, whilst the remaining four (CIE, DSB, NS and CFF) may be characterised as having increasing returns. Of the eight railways with returns to scale that are less than one, VR may be characterised as having constant returns and SNCF, DB, FS, RENFE, BR, SJ and OBB as having decreasing returns to scale.

In Table 7, we attempt to explain the variation of the three key indices presented in Table 6 through regression analysis. In earlier work (Preston and Nash 1993), we hypothesised that much of the variation in operator efficiency was due to managerial autonomy. However, when we used the indices of managerial autonomy developed by Gathon and Pestieau (1991) for European railways, it was found that this only explained 4% of the variation in operator cost efficiency. An hypothesis that accounting conventions explain much of the variation has little empirical support, as the proportion of total costs that are capital costs only explains 1% of variation. However, cost recovery explains 18% of variation in cost efficiency, although the parameter value is not quite significant at the 5% level. Nonetheless, the parameter value implies that the elasticity of costs with respect to subsidy is 0.4, a finding consistent with that of TRRL (1980). We interpret this as evidence for the existence of x-inefficiency, albeit relatively weak evidence. However, it is likely that geographic factors such as the need for sea-crossings (DSB) and the prevalence of mountainous terrain (CFF) may explain at least as much variation.

Table 7 also shows that 16% of the variation in returns to density was explained by traffic density and 54% of the variation in returns to scale was explained by train kilometres. Our results imply an optimal density of around 28,000 train km per km of line per annum and optimal output level of around 120 million train km per annum. However, the best explanation for variation in returns to scale was provided by length of line. Graphical analysis of the relationship between returns to scale and length of line indicated that the rectangular hyperbola was the most appropriate function. The reciprocal of length of line explains 93% of variation in returns to scale and suggests an optimal network size of around 3,900km.

Table 7 Explanatory regressions (t-statistics in brackets)

Dependent Variable	Independent Variable	Constant	Slope	R^2
Operators Comparison	Autonomy Index	1.501 (3.60)	-0.005 (-0.70)	0.04
Operators Comparison	Capital Costs	1.335 (4.11)	-0.007 (-0.38)	0.01
Operators Comparison	Cost Recovery	1.734 (5.441)	-1.105 (-1.63)	0.18
Returns to Density	Density	15.099 (2.21)	-0.0005 (-1.57)	0.16
Returns to Scale	Length of Line	1.142 (22.85)	-0.00002 (-5.68)	0.71
Returns to Scale	Train Km	1.108 (17.87)	-0.0000009 (-3.90)	0.54
Returns to Scale	Length of Line^{-1}	0.618 (20.80)	1486.567 (13.06)	0.93

In Table 8 we look at the sensitivity of our results concerning returns to density and scale to assumptions about capital costs. We tested two broad assumptions. Firstly, that capital costs are fixed. For returns to density we assume that a change in TKT will not affect capital costs and for returns to scale we assume that a change in TKT and LL will not affect capital costs (or put another way we assume that the elasticity of capital costs with respect to TKT and LL is zero).

Secondly, we assume that capital costs are variable. For returns to density we assume that a proportionate change in TKT will lead to a proportionate change in capital costs (that is the elasticity of capital costs with respect to TKT is one). For returns to scale we assume initially that a proportionate change in TKT and LL will lead to a proportionate change in capital costs; that is the elasticity of capital costs with respect to TKT and LL is one. An alternative assumption is that the elasticity of capital costs with respect to TKT is one and the elasticity of capital costs with respect to LL is one, implying a return to scale with respect to capital costs of 0.5.

Table 8 Sensitivity of returns to density and scale to assumptions concerning capital costs

Operator	Returns to Density		Returns to Scale		
	Capital Costs Fixed	Capital Costs Variable	Capital Costs Fixed	Capital Costs Variable (1)	Capital Costs Variable (2)
BR	1.17	1.10	0.76	0.73	0.70
CFF	1.01	0.84	1.52	1.17	0.95
CIE	6.71	3.66	1.41	1.20	1.04
DB	1.59	1.26	0.76	0.68	0.61
DSB	1.55	1.06	1.79	1.18	0.88
FS	1.82	1.35	0.86	0.73	0.64
NS	0.91	0.80	1.47	1.20	1.02
NSB	5.05	3.38	1.13	1.02	0.93
OBB	1.39	1.22	0.98	0.89	0.81
RENFE	2.65	1.69	0.93	0.78	0.66
SJ	7.14	3.17	0.92	0.79	0.87
SNCB	1.16	1.00	1.28	1.08	0.94
SNCF	3.06	1.82	0.77	0.66	0.57
CP	2.42	1.66	1.30	1.04	0.87
VR	62.50	5.15	1.09	0.91	0.78

Analysis of Table 8 indicates that our results are sensitive to assumptions concerning capital costs. For returns to density, if we assume capital costs are fixed, compared to Table 6, returns increase so that all railways exhibit increasing returns except NS (decreasing returns) and CFF (constant returns). If capital costs are assumed to be variable, then there is a tendency for returns to converge towards unity, but with most railways exhibiting increasing returns, with the exceptions of CFF and NS (decreasing returns) and DSB and SNCB (constant returns).

For returns to scale, if we assume capital costs are fixed then, compared to Table 6, returns to scale increase so that OBB now appears to be the railway with the optimal size network as it exhibits constant returns. Under our first assumption concerning variable capital costs (returns to scale with respect to capital costs equals one), returns converge towards unity, with CP and NSB exhibiting constant returns. Under our second assumption concerning variable capital costs (returns to scale with respect to capital costs equals 0.5), returns reduce so that NS and CIE exhibit constant returns and all other railways exhibit decreasing returns.

We believe that capital costs are likely to have a large fixed element (50% plus) and the most plausible estimates of returns to density and scale may be between those given in Table 6 and by the capital costs fixed columns of Table 8.

In Table 9 we re-estimate the TFP measures of Table 5, but replace the revenue share weights with cost elasticity weights. The cost elasticities have been calculated on the basis of capital costs for each railway being fixed and on the basis that the elasticity measures with respect to final output are the same as those with respect to intermediate output. The results in Table 9 are thus meant to be illustrative rather than definitive. In terms of TFP with respect to our supply-side measure, CIE is now the top performer although FS remains the worst performer, with BR (our base case) ranked eighth, equal with SNCB. In terms of TFP with respect to demand-side measures, SJ becomes the best performer, FS remains the worst performer and BR's ranking falls to ninth.

Table 9 **Total Factor Productivity—based on cost elasticity**

	Freight Cost Elasticity	Passenger Cost Elasticity	TFPS	TFPD
BR	0.25	0.85	1.00	1.00
CFF	0.36	0.82	1.14	1.32
CIE	0.28	0.10	1.99	1.89
CP	0.26	0.56	0.83	1.20
DB	0.31	0.30	0.60	0.67
DSB	0.19	0.80	1.08	1.08
FS	0.09	0.22	0.30	0.41
NS	0.21	1.38	1.43	1.45
NSB	0.48	0.23	1.94	1.82
OBB	0.03	0.17	0.60	0.68
RENFE	0.31	0.28	1.06	1.26
SJ	0.23	-0.09	1.96	2.46
SNCB	0.39	0.73	1.00	0.68
SNCF	0.15	0.05	0.40	0.53

CONCLUSIONS

In Table 10 we compare the 14 railways' rankings in terms of eight measures of performance we have developed, along with four measures developed by Oum and Yu (1994) with DEA techniques and 1989 data. The figures in Table 10 are derived as follows. The measure of commercial performance used is mean market share, derived from Table 3. The measures of operating and financial performance used are train km per staff member and total receipts divided by total costs respectively, both of which are derived from Table 2. The TFP measures using revenue shares as output weights are derived from Table 5. The translog cost efficiency index is derived from Table 6, whilst the TFP measures using cost elasticities as output weights are derived from Table 9. Oum and Yu's DEA gross efficiency indices are derived from their Table 3 and the residual efficiency indices from their Table 6.

Table 10 Summary of ranking measures

	Partial Measures			TFP Revenue Share		T'log Cost	TFP Cost Elasticity		DEA Gross Efficiency		Residual Efficiency	
	Comm. Perf.	Op. Perf.	Fin. Perf.	Supp.	Dem.	EI	Supp.	Dem.	Supp.	Dem.	Supp.	Dem.
BR	11	5	1	7	8	7=	8=	9	1=	1=	5=	3=
CFF	2	6	3	5	5	11	5	5	9	11	13	13=
CIE	13	8	7=	10=	11	3	1	2	1=	1=	7=	1
CP	8	12	12	13	9	12	10	7	10=	1=	9=	6
DB	6	9	9	6	6	6	11=	12	8	13	9=	13=
DSB	7	7	7=	10=	10	9	6	8	12	10	7=	3=
FS	10	14	14	14	13	10	14	14	13	7	14	9=
NS	14	1	6	3	4	7=	4	4	1=	5	1	2
NSB	9	10	5	2	7	4	3	3	7	9	5=	8
OBB	1	13	11	12	12	14	11=	10=	10=	14	12	11
RENFE	12	3	10	4	2	5	7	6	1=	8	2	7
SJ/BV	3	2	2	1	1	1	2	1	1=	1=	3	5
SNCB	5	4	13	8	14	13	8=	10=	14	12	11	9=
SNCF	4	11	4	9	3	2	13	13	6	6	4	12

Notes:

Comm. Perf = Commercial Performance
Op. Perf = Operating Performance
Fin. Perf = Financial Performance
Suppl. = Supply
Dem. = Demand
T'log Cost EI = Translog Cost Efficiency Index

The main feature of Table 10 is the variability of each railway's rankings. The relative assessment of performance of a railway depends crucially on the measure used. However, some common themes do emerge. Sweden's railways are a consistent top performer, being in the top five for all twelve measures considered. At the other extreme, Italy's railways are a consistently poor performer being in the bottom five for ten of the measures considered. The Austrian and Belgian railways also stand out as consistently poor performers, particularly if our findings on commercial performance are discounted as not saying anything informative about productivity. Similarly, the Dutch railways emerge as a good performer if their poor commercial performance, in terms of market share, is ignored. The classification of the other railways in our sample is rather more difficult. The main message of Table 10 is that no one league table can be entirely informative.

Furthermore, although our performance measures can give some broad indications of good and poor performance, they are, as they stand, much less useful in explaining the causes of these performance differences. Clearly, scale effects are important, which are in turn largely determined by network size and density of utilisation. However, even when these scale effects are taken into account large variations in performance remain. In contrast to Oum and Yu (1994), we have not been able to determine managerial autonomy as a statistically significant explanatory variable although there is considerable qualitative evidence to suggest that this may be an important factor. We have, though, like Oum and Yu, uncovered some evidence that high levels of subsidy have leaked into higher costs and x-inefficiencies.

We conclude that assessing the performance of the western European railway industry remains fraught with problems. There is still considerable work required to determine consistent data and methodologies. Nonetheless, it is work that is worth undertaking as it will help inform the public policy debate over the future organisation of Europe's railways and the proposed transition of at least some parts of the industry from the public to the private sector.

REFERENCES

Boxwell, R.J. (1994) *Benchmarking for Competitive Advantage*, McGraw-Hill, New York.

British Railways Board and University of Leeds (1979) *A Comparative Study of European Rail Performance,* BRB, London.

Caves, D.W., Christensen, L.R. and Diewert, W.E. (1982) Multilateral comparisons of output, input and productivity using superlative index numbers, *The Economic Journal* 92, 73-86.

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 US Railroads, in A.F. Daugherty (ed) *Analytical Studies in Transport Economics*, Cambridge University Press.

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.

Dodgson, J.S. (1985) A Survey of Recent Developments in the Measurement of Total Factor Productivity, in K.J. Button and D.E. Pitfield (eds) *International Railway Economics*, Gower, Aldershot.

Friedlander, A.F., Berndt, E.R., Chiang, J., Showalter, M. and Vellturo, C.A. (1993) Rail costs and capital costs in a quasi regulated environment, *Journal of Transport Economics and Policy* 27 (2), 31-152.

Ganley, J.A. and Cubbin, J.S. (1992) *Public Sector Efficiency Measurement Applications of Data Envelopment Analysis*, Elsevier Science, Amsterdam.

Gathon, H.J. and Pestieau, P. (1991) *Autonomy and Public Performance in Public Enterprises: The Case of European Railways*, CIRIEC Working Paper 91/04, Department of Economics, University of Liege.

Harris, R.G. (1977) Economics of traffic density in the rail freight industry, *Bell Journal of Economics* 8 (22), 556-63.

Hensher, D.A., Daniels, R. and DeMellow, I. (1992) *A Comparative Assessment of the Productivity of Australia's Rail Systems 1971/2-1990/1*, Working Paper ITS-WP-92-7, Institute of Transport Studies, University of Sydney.

Hensher, D.A. and Waters, W.G. (1993) *Using Total Factor Productivity and Data Envelopment Analysis for Performance Comparisons Among Government Enterprises: Concepts and Issues*, Working Paper ITS-WP-93-10, Institute of Transport Studies, University of Sydney.

Keeler, T.E. (1974) Railroad costs, returns to scale and excess capacity, *Review of Economics and Statistics* 61, 201-208.

Nash, C.A. (1985) European Railway Comparisons—What Can We Learn, in K.J. Button and D.E. Pitfield (eds) *International Railway Economics*, Gower, Aldershot.

Nash, C.A. and Preston, J.M. (1992) Assessing the performance of European railways, in *Lyon '92 Selected Proceedings of the 6th World Conference in Transport Research*, Vol II Transport Policies, pp. 1943-1954.

Obeng, K. (1985) Bus transit cost, productivity and factor substitution, *Journal of Transport Economics and Policy* 18, 183-203.

Oum, T.H. and Yu, C. (1994) Economic efficiency of railways and implications for public policy: A comparative study of the OECD countries' railways, *Journal of Transport Economics and Policy* 27 (2), 121-138.

Pindyck, R.S. and Rubinfield, D.L. (1991) *Econometric Models and Economic Forecasts,* McGraw Hill, New York.

Preston, J.M. (1994a) Cost recovery worsens as Europe's major railways lose market share, *Railway Gazette International* 150 (7), 465-466.

Preston, J.M. (1994b) Does size matter? A case study of Western European railways, *Proceedings of the UTSG Annual Conference*, University of Leeds.

Preston, J.M. and Nash, C.A. (1993) European railway comparisons: Lessons for policy, *Proceedings of the Third International Conference on Competition and Ownership in Surface Passenger Transport.*

Preston, J.M., Shires, J., Garlick, M., Hodgson, F. and Nash, C.A. (1994) *European Railway Comparisons: Final Report*, Working Paper 418, Institute for Transport Studies, University of Leeds.

SAS (1988) *SAS/ETS Users Guide Version 6*, SAS Inc, Cary, N. Carolina.

Talvitie, A. and Obeng, K. (1991) Productivity Measurement, in *Proceedings of the First International Conference on Competition and Ownership of Bus and Coach Services, Transportation Planning and Technology* 15, 169-176.

TRRL (1980) *The Demand for Public Transport: Report of the International Collaborative Study of the Factors Affecting Public Transport Patronage*, TRRL, Crowthorne.

Zellner, A. (1962) An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, *Journal of the American Statistical Association* S7, 348-368.