

Efficiency Measures and Output Specification: The Case of European Railways

PEDRO CANTOS

JOSÉ M. PASTOR

LORENZO SERRANO

Universidad de Valencia

ABSTRACT

This study analyzes the sensitivity of the efficiency indicators of a sample of European railway companies to different alternatives in output specification. The results vary according to the specification selected. However, investigating the causes of these differences reveals that the efficiency indicators obtained with different specifications can be brought substantially closer, particularly when the efficiency indicators obtained by considering freight and passenger train-kilometers as output variables are corrected to account for the impact of the load factor.

INTRODUCTION

The literature on productivity and efficiency frequently reports different rankings in terms of both productivity and efficiency indicators, depending on the output variables used in the construction of the model.¹ In the case of railways, there are very few studies, besides that by Oum and Yu (1994), in which this phenomenon has been tested since most

¹ Berg et al. (1992) and Grifell et al. (1993) analyze the levels of efficiency for a sample of banks and show the sensitivity of the results obtained to the specification adopted for the output.

Pedro Cantos, Departamento de Análisis Económico, Universidad de Valencia, Edificio Departamental Oriental, Campus del Tarongers, s/n 46022 Valencia, Spain. Email: Pedro.Cantos@uv.es.

estimates use a single specification for output. Recently, Oum et al. (1999) published a complete review of productivity and efficiency estimates in rail transport in which it is clear that the results of these estimates are very sensitive to output specification.

Other papers have estimated technical efficiency levels for European railways on the basis of a deterministic production function (Perelman and Pestieau 1988) or a stochastic one (Gathon and Pestieau 1995). Cantos et al. (1999) obtained efficiency indicators using a non-parametric approach, and Cowie and Riddington (1996) used alternative methodologies. According to the latter, accurate measurement of efficiency is not possible although research is able to indicate good and bad performers. In most of the studies quoted, such as in Cantos et al. (1999), some companies are generally very efficient, as in the case of Sweden's SJ, Holland's NS, or Switzerland's CFF. However, others are very inefficient, such as the Greek CH, the Danish DSB, and the Irish CIE.

A notable feature of the railway industry is its multi-product character: there are various types of passenger railway output (long distance, urban, high speed, etc.) and freight output (general, intermodal, parcels, etc.). However, due to the shortage of data, most studies restrict the output vector to two aggregate dimensions, passenger and freight. The measurements most commonly used are the number of passenger-kilometers and ton-kilometers (see Caves et al. 1980, 1982, and 1985; McGeehan 1993; and Cantos et al. 1999). These demand-related measurements for output enable an assessment of the level of user consumption and the value they place on the service. As indicated by Oum and Yu (1994), this specification is recommended when there is little government control, such as when the restrictions imposed on the level of service (frequency) or prices are of little importance. In that case, the indices of passenger-kilometers and ton-kilometers adequately reflect the efficient productive behavior of the various production units.

On the other hand, if there is a high degree of government control over decisions about pricing or frequency, the above specification will not adequately reflect the greater or lesser efficiency of the companies since output will be influenced by these regulatory measures. In this case, supply-related or

intermediate measurements for output which place the emphasis on the degree of capacity or service level supplied by the companies are more suitable. For this reason, Nash (1985), Deprins and Simar (1989), Preston (1996), and Cantos and Maudos (2000) use the number of freight and passenger train-kilometers as output. These types of measurements isolate the effect of governmental control measures. Nevertheless, the use of this second type of measurement may lead to paradoxical results, such as situations in which companies with very low indices of load factor but with high levels of train-kilometers run are even more efficient than companies with high indices of load factor and low levels of train-kilometers run.

The problem with grouping companies from different countries is that the degree of governmental intervention and control is very different, complicating choice of measurement type. Oum and Yu (1994) estimate efficiency indicators on the basis of a Data Envelopment Analysis (DEA) model, using two different sets of measurements, passenger-kilometers and ton-kilometers on the one hand and passenger train-kilometers and freight train-kilometers on the other. Their results confirm that levels and rankings of efficiency differ, depending on which measurement is used. Thus, some companies such as the Spanish RENFE or the Norwegian NSB are clearly inefficient when measured by passenger-kilometers and ton-kilometers; however, with the other type of measurement, both companies notably improve their efficiency indicators. Therefore, the choice of output specification used continues to be a problem in studies of the estimation of efficiency and productivity.

Our study aims to analyze the differences in the efficiency indicators for the railway sector when different variables are specified as output. For this purpose, we use a non-parametric DEA model to calculate the efficiency indicators of a sample of European railway companies using the two types of output mentioned above. We then regress the difference between the efficiency indicators obtained on the indices of load factor of the supplied trains and show that when one of the efficiency indicators is corrected for the effect of these variables, the efficiency indicators of the two types of output become similar.

Our results, then, demonstrate that the differences in efficiency indicators can be explained by the differences in the output specification used, suggesting that efficiency indicators are compatible once differences in output specification are considered. When passenger and freight train-kilometers are specified as output, the efficiency is analyzed only as a function of the level of capacity or service supplied in terms of the volume of kilometers traveled. Meanwhile, when the number of passenger-kilometers and ton-kilometers is used, efficiency is evaluated as a function of the degree of use of the capacity or service supplied. Our study shows that the levels and rankings of efficiency obtained on the basis of different output specifications can be approximated by analyzing the differences between the output variables used.

METHODOLOGY, DATA, AND RESULTS

Methodology

In this study, we use the non-parametric technique, Data Envelopment Analysis (DEA), to estimate the technical efficiency of railway companies. DEA has two advantages over other techniques. First, it does not require specification of any functional form for production, avoiding the bias produced by an incorrect functional form. Second, DEA is better than parametric techniques at assessing the productive efficiency of railway companies since it can handle the multi-product nature of some companies.²

We calculate efficiency indicators with DEA by constructing a frontier through mathematical programming. A comparison of the companies relative to this production frontier gives us the measurements of individual effectiveness. Unlike parametric techniques, this technique does not estimate a previously specified functional form but instead calculates a convex frontier that “envelops” the

² In this respect, some authors, such as Cowie and Riddington (1996), analyze the productive efficiency of railway companies by using parametric techniques as well as DEA. However, since parametric techniques only allow specification of a production function with a single output, these authors chose the number of passenger train-kilometers as the output, without considering that the companies also carry freight. The consideration of a single output causes a bias in the efficiency measurements obtained, undervaluing the efficiency of those companies that specialize in freight.

observations. In this sense, the data themselves “dictate” the profile of the frontier. This technique’s flexibility (it makes few assumptions) and applicability have led to its use in a large number of studies in recent years.³

To illustrate this technique,⁴ let us suppose that the N companies forming the sample ($i = 1, \dots, N$) use a vector of input $x_i = (x_{i1}, \dots, x_{in})^T \in R^n +$ to produce a vector of output $y_i = (y_{i1}, \dots, y_{im})^T \in R^m +$. The measurement of the efficiency of company j (θ_j) is obtained by comparing this company’s performance with a linear combination of the N companies of the sample:

$$\begin{aligned} & \text{Max}_{\theta, \lambda} \theta_j \\ & \text{such that: } Y\lambda \geq \theta_j y_j, \\ & X\lambda \leq x_j, \\ & \lambda \geq 0 \end{aligned} \tag{1}$$

where x_j and y_j are vectors of dimensions $(n \times 1)$ and $(m \times 1)$, respectively; λ is a vector of dimension $(n \times 1)$, while X and Y are matrices of dimensions $(n \times N)$ and $(m \times N)$, respectively.

From the resolution of this problem for each of the N companies of the sample, we obtain N weightings (λ) and N optimum solutions (θ^*). Each optimum solution θ^* is the parameter of efficiency of each company that, by construction, satisfies $\theta^* \geq 1$. Companies with $\theta > 1$ are considered inefficient, while those with $\theta = 1$ catalogued as efficient are those that stand at the frontier. The inherent virtues of the DEA technique have encouraged studies comparing this methodology with alternative techniques, with varying results.⁵

³ Seiford and Thrall (1990) counted more than 400 articles on the application of DEA between 1978 and 1990. More recently, Førsund and Seiford (1999) count the empirical applications of this technique in the thousands.

⁴ See details in Charnes, Cooper, and Rhodes (1978).

⁵ See Banker, Conrad, and Strauss (1986); Gong and Sickless (1992); Ferrier and Lovell (1990); Bjurek, Hjalmarsson, and Førsund (1990); Pastor (1996); Cowie and Riddington (1996); etc. However, the precision of the estimation of efficiency with DEA can only be assessed on the basis of simulated data where the efficiency is known in advance. In this respect, Banker et al. (1988) compare the results of a translogarithmic function, using simulated data for a known underlying technology, concluding that

From an intuitive viewpoint, to analyze the efficiency of the productive scheme of company j (y_j, x_j) the problem constructs a feasible scheme as a linear combination of the schemes of the N companies of the samples that produce $\theta_j y_j$, using a lower or equal amount of input. Therefore, $(\theta_j - 1)$ indicates the maximum radial expansion to which the vector of the output of company j can be subjected without needing to increase the level of input. When $\theta_j = 1$, no linear combination of companies producing more with less input can be found, so the company is catalogued as efficient. In the other cases, $\theta_j > 1$, and so a feasible alternative scheme which obtains a higher amount of output using the same input does exist.

Data

A panel of 17 European companies over the period of 1970 to 1995 was selected. The information was taken mostly from the reports published by the Union International des Chemins de Fer and was completed with the data published in the companies' statistical memoranda. Table 1 provides a set of the main characteristics of the railways used. Two sets of output were selected: 1) the number of passenger-kilometers (PKT) and ton-kilometers (TOKT) and 2) the passenger train-kilometers (PTK) and freight train-kilometers (FTK). For both, we estimate the efficiency indicators of the European companies using a non-parametric frontier approach (DEA). The variables used as input were 1) number of workers, 2) consumption of energy and materials,⁶ 3) number of locomotives,

the predominance of DEA over parametric methods with regard to lesser deviation from the true values is due to DEA's greater flexibility of approach to the true functional form. Banker et al. (1988) also verify that the accuracy of the DEA results is greater when the size of the sample is increased, suggesting that DEA estimators show the property of consistency, subsequently shown theoretically by Banker (1993). In this same sense, Gong and Sickles (1992) conclude that the disadvantages of DEA relative to other methods depend on the choice of functional form. If the chosen specification coincides with the underlying one, parametric methods work better. On the other hand, the advantages of DEA are more evident when errors of specification exist.

⁶ This variable was converted into U.S. currency using the Purchasing Power Parity Index obtained from the Organisation for Economic Cooperation and Development (OECD) reports (2000) and deflated to constant 1975 value.

4) number of passenger carriages, 5) number of freight cars, and 6) number of track-kilometers.⁷

It should be noted that there are other factors that can affect the level of efficiency. The different indices of the quality of service or of infrastructure may bias the results if they are not taken into account. Another important factor is the degree of circuitousness. For example, if the infrastructure is expanded to allow for less circuitous routes, the number of passenger-kilometers or ton-kilometers will decrease even though the outcome is unchanged. The lack of relevant information on this type of variable makes it impossible to consider them in our study.

Results

The individual average inefficiency indicators for the period are shown in table 2.⁸ INEF refers to the results obtained using the number of passenger-kilometers and ton-kilometers as output, and INEG refers to results obtained using passenger train-kilometers and freight train-kilometers. Each type of measurement refers to different aspects of the efficiency in the use of input, as noted in the previous section. The average correlation indices measured by the Pearson coefficient and the Spearman ranking coefficient between INEF and INEG are respectively 0.62 and 0.76, each with a standard error of 0.16.

Alternatively, a parametric test was made of the similarity of the two measurements, using ordinary least squares (OLS) to regress the inefficiency indicators obtained in INEF against the indicators obtained in INEG. The value of the parameter estimated was 0.937, with a standard error of 0.008. In this case, the null hypothesis that the parameter is equal to one can be rejected; in other words, it can be rejected that both measurements of efficiency are statistically equal (student's t is 7.80).

⁷ A more detailed discussion of the data used in this study can be found in Cantos et al. (1999).

⁸ We will follow Farrell's (1957) definition of the technical efficiency of a company: it is not possible to produce more output with less input. In the results of table 2, a company is technically efficient in this way when the index has a value of 1, whereas if the index is higher than 1, the company would be able to increase output without needing to increase input.

TABLE 1 Average Values for Variables: 1970 to 1995

| Company | Country | PKT (mill.) | TOKT (mill.) | LT (km) | PTK (thous.) | FTK (thous.) | PT | TT |
|-----------------|-------------|-------------|--------------|---------|--------------|--------------|--------|--------|
| BR ^a | UK | 30,917 | 18,200 | 17,387 | 345,279 | 78,297 | 90.14 | 239.24 |
| CFF | Switzerland | 9,586 | 7,067 | 2,960 | 75,880 | 27,847 | 127.17 | 253.64 |
| CFL | Luxemburg | 234 | 658 | 272 | 3,335 | 1,520 | 74.53 | 435.71 |
| CH | Greece | 1,687 | 695 | 2,494 | 13,787 | 2,998 | 123.82 | 238.13 |
| CIE | Ireland | 1,007 | 590 | 2010 | 8,413 | 4,231 | 117.91 | 140.05 |
| CP | Portugal | 4,940 | 1,272 | 3,434 | 27,477 | 6,639 | 178.64 | 183.36 |
| DB ^b | Germany | 41,977 | 60,638 | 28,094 | 40,049 | 198,841 | 100.62 | 305.12 |
| DSB | Denmark | 4,081 | 1,793 | 2,227 | 40,919 | 7,982 | 99.89 | 227.25 |
| FS | Italy | 40,679 | 18,912 | 16,257 | 227,877 | 60,584 | 177.46 | 309.88 |
| NS | Holland | 9,928 | 3,114 | 2,851 | 98,918 | 13,156 | 98.40 | 238.07 |
| NSB | Norway | 2,097 | 2,694 | 4,180 | 23,235 | 10,217 | 90.30 | 263.94 |
| ÖBB | Austria | 7,536 | 11,203 | 5,781 | 66,622 | 36,501 | 114.56 | 304.63 |
| RENFE | Spain | 15,361 | 11,441 | 13,083 | 101,791 | 44,353 | 155.33 | 255.86 |
| SJ | Sweden | 5,741 | 16,595 | 11,110 | 59,839 | 40,042 | 95.52 | 414.47 |
| SNCB | Belgium | 6,940 | 8,235 | 3,975 | 68,786 | 21,667 | 102.05 | 379.89 |
| SNCF | France | 54,967 | 60,519 | 34,716 | 290,831 | 189,179 | 188.75 | 320.02 |
| VR | Finland | 3,027 | 7,617 | 5,941 | 24,651 | 17,816 | 122.65 | 426.18 |

^a Does not include information for period 1970–1973

^b Does not include information for 1995

PKT: number of total passenger-kilometers

TOKT: number of total ton-kilometers

LT: length of track

PTK: number of passenger train-kilometers

FTK: number of freight train-kilometers

PT: index of passengers per train

TT: index of tons per train

TABLE 2 Inefficiency Levels

| | INEF | | INEG | | INEGC |
|-------|-------|-------|-------|-------|-------|
| CIE | 1.703 | CH | 1.528 | CIE | 1.512 |
| CH | 1.475 | FS | 1.520 | CH | 1.476 |
| BR | 1.352 | SNCB | 1.449 | DSB | 1.375 |
| DSB | 1.269 | CIE | 1.428 | SNCB | 1.344 |
| NSB | 1.183 | DSB | 1.349 | BR | 1.339 |
| SNCB | 1.107 | CP | 1.296 | FS | 1.268 |
| RENFE | 1.045 | BR | 1.295 | CP | 1.188 |
| FS | 1.029 | RENFE | 1.109 | NSB | 1.109 |
| DB | 1.027 | NSB | 1.087 | CFL | 1.045 |
| CP | 1.022 | VR | 1.079 | NS | 1.013 |
| ÖBB | 1.020 | ÖBB | 1.062 | DB | 0.995 |
| SNCF | 1.014 | CFL | 1.054 | RENFE | 0.994 |
| CFL | 1.004 | SNCF | 1.037 | ÖBB | 0.992 |
| VR | 1.000 | DB | 1.031 | CFF | 0.941 |
| SJ | 1.000 | CFF | 1.000 | VR | 0.937 |
| CFF | 1.000 | NS | 1.000 | SJ | 0.928 |
| NS | 1.000 | SJ | 1.000 | SNCF | 0.851 |

INEF: inefficiency index when passenger-kilometers and ton-kilometers are used as output measurements

INEG: inefficiency index when passenger train-kilometers and freight train-kilometers are used as output measurements

INEGC: corrected inefficiency measures when passenger train-kilometers and freight train-kilometers are used as output measurements

We would expect the different degrees of utilization of trains to explain a large part of the differences. In particular, companies with high indices of load factor are much more efficient when passenger-kilometers and ton-kilometers are used as measures of output. See the values for variables representing the number of passengers per train (PT) and freight tons per train (TT) in table 1 for VR, SNCB, or FS. On the other hand, companies with low indices of load factor, such as NSB, are more efficient when output is expressed as train-kilometers.

In any event, due to the multi-product nature of railway companies and the wide range of input used, there is no simple transformation between the two output measurements or between the measurements of efficiency obtained in each case. This can only occur when there is a single output and a single input. In this case, if we know that the company offers only passenger services, we can use two measurements of output, passenger-kilometers (PKT) or passenger train-kilometers (PTK). If we only have one input (I), a measurement of productivity can be constructed from the ratio of PKT/I or PTK/I. Therefore, a simple transformation exists between the two measurements using the ratio PKT/PTK. However, this is not the case for the railway industry.

We define DINEF as the difference between the logs of INEG and INEF and regress it by OLS on the logs of the number of passengers per train (LPT) and the freight-tons per train (LTT). Thus, $DINEF = \log(INEG/INEF)$. The regression results, including time effects ($DUMMY_t$), follow⁹:

$$\begin{aligned} DINEF_{it} &= DUMMY_t + \\ &0.2493 LPT_{it} + 0.175 LTT_{it} \\ R^2 &= .2984; N=442,^{10} \end{aligned} \quad (2)$$

where the LPT coefficient has a t-statistic of 10.60 and the LTT coefficient has a t-statistic of 8.19.

⁹ Note that the regression does not include a constant since all the time effects were included in the estimation. Alternative specifications were also tried for the variables of the regression (semi-logarithmic transformation, estimation of levels, etc.). The results were very similar to those of equation (2), so the logarithmic specification was chosen due to the advantages of its ease of interpretation and the reduction of problems of heteroscedasticity.

¹⁰ The *F*-test for the joint significance for LPT and LTT is $F_{2,416} = 55.49$. However, the *F*-test for the significance for LPT, LTT, and the time effects is $F_{27,442} = 6.52$. In both cases, the null hypothesis of nonsignificance is clearly rejected.

Other reasons for the difference between these two measurements of efficiency may exist, such as the different passenger and freight traffic. In the case of passenger traffic, the companies that focus their production on urban services will carry a larger number of passenger-kilometers out of the same number of kilometers supplied than the companies focusing on long distance services. In this example, the lack of this type of information prevents a better fit of the regression given in equation (2).

We can see that both variables are highly significant and positive. Therefore, estimates of efficiency that use indices of train-kilometers penalize the companies with high indices of load factor. Estimates that use indices of passenger-kilometers and ton-kilometers favor companies with high indices of load factor. A higher degree of load factor involves a higher level of inefficiency when only train-kilometers are used as a measurement of output. We can obtain a corrected measurement of INEG (INEGC) by taking into account the effect of the degree of load factor:

$$\begin{aligned} INEGC_{it} &= \exp(\log INEG_{it} - \hat{DINEF}_{it}), \\ \text{where } \hat{DINEF}_{it} &= DUMMY_t \\ &+ 0.2493 LPT_{it} + 0.175 LTT_{it}. \end{aligned} \quad (3)$$

With this, we aim to correct such a bias. The individual average levels of this corrected measurement of inefficiency are shown in table 2. The correlation between INEGC and INEF rises to 0.82, a value clearly higher than the 0.62 obtained for the original inefficiencies. In the case of Spearman's correlation coefficient, the growth is more modest, passing from an initial 0.76 to 0.84, with a standard error of 0.13. As for the alternative test of the two measurements of efficiency, if INEF is now regressed against INEGC, the parameter estimated for INEGC is 0.989, with a standard error of 0.006. In this case, the null hypothesis that these measurements are equal cannot be rejected (student's *t* is 1.63). The results indicate that once we take into account the different focus of each type of output measurements, the inefficiencies we obtain are largely consistent. The results show a similar view of the performance of European companies over the period and that in an analysis of efficiency it is not only important to know a company's position in the ranking but also its relative level of efficiency.

CONCLUSIONS

This paper verifies the sensitivity of the efficiency indicators to the output specification in the rail sector. Additionally, it shows that the results obtained with two different specifications for railway output can be harmonized. In particular, when the efficiency indicators obtained with one of the specifications, number of passengers and freight train-kilometers, are corrected to take the degree of utilization of the trains into account, the efficiency indicators obtained with this new specification are very similar to those obtained when the number of passengers and ton-kilometers are used as output measures. This study shows, therefore, that the analysis of the differences between the alternatives for the specification of measurements of output helps to explain the differences between the indicators of efficiency that such measurements can generate. Thus, this analysis serves as an additional means of testing the consistency of the efficiency results obtained.

ACKNOWLEDGMENTS

The authors are grateful for the comments and suggestions made by the anonymous referees of the *Journal of Transportation and Statistics* and for the financial support of the DCICYT SEC98-0895.

REFERENCES

- Banker, R.D. 1993. Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation. *Management Science* 39, no. 10:1265–73.
- Banker, R.D., A. Charnes, W.W. Cooper, and A. Maindirata. 1988. A Comparison of Alternative Approaches to the Measurement of Productive Efficiency. *Applications of Modern Production Theory: Efficiency and Productivity*. Boston, MA: Kluwer Academic Publishers.
- Banker, R.D., R. Conrad, and R. Strauss. 1986. A Comparative Application of Data Envelopment Analysis and Translog Methods: An Illustrative Study of Hospital Production. *Management Science* 32, no. 1:30–44.
- Berg, S., F.R. Førsund, and E.S. Jansen. 1992. Technical Efficiency of Norwegian Banks: The Non-Parametric Approach to Efficiency Measurement. *The Journal of Productivity Analysis* 2, 127–42.
- Bjureck, H., L. Hjalmarsson, and F.R. Førsund. 1990. Deterministic Parametric and Non-Parametric Estimation of Efficiency in Service Production, A Comparison. *Journal of Econometrics* 46, 213–27.
- Cantos, P. and J. Maudos. 2000. Efficiency, Technical Change and Productivity in the European Rail Sector: A Stochastic Frontier Approach. *International Journal of Transport Economics* 27, no. 1:55–75.
- Cantos, P., J.M. Pastor, and L. Serrano. 1999. Productivity, Efficiency and Technical Change in the European Railways: A Non-Parametric Approach. *Transportation* 26, no. 4:337–57.
- Caves, D.W., L.R. Christensen, and J.A. Swanson. 1980. Productivity Growth, Scale Economies and Capacity Utilization in U.S. Railroads, 1955–1974. *American Economic Review* 71, December: 994–1002.
- Caves, D.W., L.R. Christensen, M.W. Tretheway, and R.J. Windle. 1985. Network Effects and Measurement of Returns to Scale and Density in U.S. Railroads. *Analytical Studies in Transport Economics*. Cambridge, MA: Cambridge University Press.
- Caves, D.W., L.R. Christensen, and W.E. Diewert. 1982. The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity. *Econometrica* 50, 1393–414.
- Charnes, A., W.W. Cooper, and E. Rhodes. 1978. Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research* 2, 429–44.
- Cowie, J. and G. Riddington. 1996. Measuring the Efficiency of European Railways. *Applied Economics* 28, 1027–35.
- Deprins, D. and L. Simar. 1989. Estimating Technical Inefficiencies with Correction for Environmental Conditions with an Application to Railway Companies. *Annals of Public and Co-operative Economics*, 81–101.
- Ferrier, G. and C.A.K. Lovell. 1990. Measuring Cost Efficiency in Banking: Econometric and Linear Programming Evidence. *Journal of Econometrics* 46, 229–45.
- Førsund, F. and L.M. Seiford. 1999. The Evolution of DEA Economics and Operations Research Perspectives. Paper presented at the Sixth European Workshop on Efficiency and Productivity Analysis, Copenhagen, Denmark, October.
- Gathon, H.J. and P. Pestieau. 1995. Decomposing Efficiency into its Managerial and its Regulatory Components: The Case of European Railways. *European Journal of Operational Research* 12, 500–7.
- Gong, B.H. and R.C. Sickles. 1992. Finite Sample Evidence on the Performance of Stochastic Frontiers and Data Envelopment Analysis Using Panel Data. *Journal of Econometrics* 51, 259–84.
- Grifell, E., D. Prior, and V. Salas. 1993. Efficiency Scores Are Sensitive to Variable Specification: An Application to Banking. Working Paper. Universitat Autònoma de Barcelona, Bellaterra.

- McGeehan, H. 1993. Railway Costs and Productivity Growth. *Journal of Transport Economics and Policy* 27, no. 1:19–32.
- Nash, C.A. 1985. European Rail Comparisons—What Can We Learn? *International Railway Economics: Studies in Management and Efficiency*. Aldershot, England: Ashgate Publishing Company.
- Organisation for Economic Cooperation and Development (OECD). 2000. Purchasing Power Parity Index. Available at www.oecd.org/statistics.
- Oum, T.H. and C. Yu. 1994. Economic Efficiency of Railways and Implications for Public Policy. *Journal of Transport Economics and Policy* 28, no. 2:121–38.
- Oum, T.H., W.G. Waters II, and C. Yu. 1999. A Survey of Productivity and Efficiency Measurement in Rail Transport. *Journal of Transport Economics and Policy* 33, no. 1:9–42.
- Pastor, J.M. 1996. Diferentes metodologías para el análisis de la eficiencia de los bancos y cajas de ahorros españoles. *Fundación Fondo para la Investigación Económica y Social (FIES)*. Documento de trabajo 123.
- Perelman, S. and P. Pestieau. 1988. Technical Performance in Public Enterprises: A Comparative Study of Railways and Postal Services. *European Economic Review* 32, 432–41.
- Preston, J. 1996. Economics of British Rail Privatisation: An Assessment. *Transport Reviews* 16, no. 1:1–21.
- Seiford, L.M. and R.M. Thrall. 1990. Recent Developments in DEA, The Mathematical Programming Approach to Frontier Analysis. *Journal of Econometrics* 46, 7–38.