Transportmetrica, Vol. 1, No. 2 (2005), 161-189

MEASURING RAILWAY PERFORMANCE WITH ADJUSTMENT OF ENVIRONMENTAL EFFECTS, DATA NOISE AND SLACKS

LAWRENCE W. LAN[1] AND ERWIN T.J. LIN[2]

Received 3 November 2004; received in revised form 18 February 2005; accepted 2 March 2005

Conventional data envelopment analysis (DEA) approaches (e.g., CCR model, 1978; BCC model, 1984) do not adjust the environmental effects, data noise and slacks while comparing the relative efficiency of decision-making units (DMUs). Consequently, the comparison can be seriously biased because the heterogeneous DMUs are not adjusted to a common platform of operating environment and a common state of nature. Although Fried et al. (2002, *Journal of Productivity Analysis*, **17**, 157-174) attempted to overcome this problem by proposing a three-stage DEA approach, they did not account for the slack effects and thus also led to biased comparison. In measuring the productivity growth, Färe et al. (1994, *American Economic Review*, **84**, 66-83) proposed a method to calculate the input or output distance functions. Similarly, they did not take environmental effects, statistical noise and slacks into account and thus also resulted in biased results. To correct these shortcomings, this paper proposes a four-stage DEA approach to measure the railway transport technical efficiency and service effectiveness, and a four-stage method to measure the productivity and sales capability growths, both incorporated with environmental effects, data noise and slacks adjustment. In the empirical study, a total of 308 data points, composed of 44 worldwide railways over seven years (1995-2001), are used as the tested DMUs. The empirical results have shown strong evidence that efficiency and effectiveness scores are overestimated, and productivity and sales capability growths are also overstated, provided that the environmental effects, data noise and slacks are not adjusted. Based on our empirical findings, important policy implications are addressed and amelioration strategies for operating railways are proposed.

KEYWORDS: Four-stage DEA, productivity, railway transport, sales capability, service effectiveness, technical efficiency

1. INTRODUCTION

Rail transport has long played an important role in the economic development for a country. However, many railways in the world have been facing keen competition from other modes such as highway and air carriers over the past few decades. Some railways have even suffered from major decline in the market share and failed to adopt effective strategies to correct the decline situation. Taking the freight transport as an example, the market share (ton-km) for China Railway (CR) has declined from 40% in 1990 to 32% in 1998 (Xie et al., 2002). The market share for European Union (EU) rails has declined from 32% in 1970 to 12% in 1999 (Lewis et al., 2001). As Fleming (1999) pointed out, truckers can deliver furniture from Lyon, France to Milan, Italy in eight hours, while railways need forty-eight hours; the decline of railway market could be attributed to relatively higher level-of-service of other competitive modes or to rail's poor performance itself in technical efficiency and/or service effectiveness. Without in-depth analysis, one can hardly gain insights into the main causes of the decline. In addition, enhancing the technical efficiency and service effectiveness as well as the productivity and sales capability should always be viewed as an important issue for the railway transport industry to remain competitive and sustainable in the market. If one could scrutinize the sources of inefficiency and ineffectiveness by making a clear distinction between efficiency and effectiveness or between productivity and sales capability, one

[1] Institute of Traffic and Transportation, National Chiao Tung University, 4F, 114 Chung Hsiao W. Rd., Sec. 1, Taipei, Taiwan 10012. Corresponding author (Email: lawrencelan@mail.nctu.edu.tw).
[2] Bureau of High Speed Rail, Ministry of Transportation and Communications, Taiwan.

would perhaps be capable of proposing more practical strategies to ameliorate the rail transport operation.

For ordinary commodities, measures of technical efficiency (a transformation of outputs from inputs) and technical effectiveness (a transformation of consumptions from inputs) are essentially the same because the commodities, once produced, can be stockpiled for consumption. Nothing will be lost throughout the transformation from outputs to consumptions if one assumes that all the stockpiles are eventually sold out. For non-storable commodities such as transport services, however, technical efficiency and technical effectiveness usually represent two distinct measurements. When such commodities are produced and a portion of which are not consumed right away, the technical effectiveness (with combined effects of both technical efficiency and sales effectiveness) would be less than the technical efficiency. In other words, it would make more sense if one could separate technical efficiency from sales effectiveness in evaluating the performance of non-storable commodities. More importantly, it would provide lucid sources of any poor performance so that appropriate enhancement strategies could be proposed accordingly. Therefore, to elucidate the non-storable nature of railway transport, it is important to expand the technical efficiency and productivity measurements to service effectiveness and sales capability measurements.

In the evaluation of mass transport performance, Fielding et al. (1985) proposed a concept of cost-efficiency, service-effectiveness and cost-effectiveness by indexing the ratios of appropriate factors drawn from output/input, consumption/output and consumption/input, respectively. Following their concept, this paper measures the railway's technical efficiency and productivity by corresponding appropriate outputs to inputs, and service effectiveness and sales capability by corresponding appropriate consumptions to outputs as depicted in Figure 1. For technical efficiency evaluation we use input-oriented data envelopment analysis (DEA) which measures the maximum possible proportional reduction in all inputs, keeping all outputs fixed; for service effectiveness evaluation we use consumption-oriented DEA which measures the maximum possible proportional expansion in all consumptions, also keeping all outputs unchanged. Likewise, for productivity evaluation we use input-based Malmquist productivity index; for sales capability evaluation we use consumption-based Malmquist sales index.

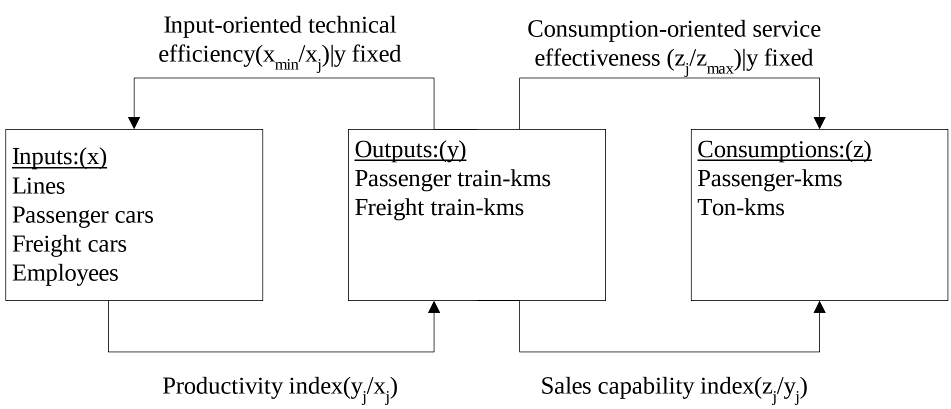


FIGURE 1: A framework for measuring the non-storable railway transport performance

In measuring the technical efficiency, conventional DEA approaches neither consider the environmental effects and data noise nor account for the slack effects; thus, the comparison is frequently seriously biased. The main reason is because all the decision making units (DMUs) are not placed on a common platform of operating environment and a common state of nature. In measuring the change in productivity, previous studies often calculate the distance functions without taking environmental effects, statistical noise and slacks into account; thus, the estimated productivity growth is often biased. To correct these shortcomings, this paper proposes a four-stage DEA approach to measure the railway transport technical efficiency and service effectiveness and also proposes a four-stage method to measure the productivity and sales capability growths. Both of four-stage DEA approach and four-stage method have considered the effects of environmental factors, data noise and slacks. Details of our proposed four-stage DEA approach, four-stage method, the empirical analysis and important policy implications will be elaborated in the subsequent sections.

2. LITERATURE ON RAILWAY PERFORMANCE MEASURES

The methods for measuring the efficiency or productivity of rail systems are generally classified into two categories: non-parametric and parametric techniques (e.g. Coelli et al. (1998) and Oum et al. (1999)). Depending on whether or not the inefficiency is accounted for, each category can be further divided into frontier and non-frontier approaches. Methods of index number and least squares are attributed to non-frontier approaches since they ignore the technical inefficiency. While data envelopment analysis (DEA) and stochastic frontier analysis (SFA) are regarded as frontier approaches because they consider the technical inefficiency. Oum et al. (1999) undertook an overall survey on these four categories of methods that have been used in the railway industry. Freeman et al. (1985) applied the index number method to measure and compare the total factor productivity of the Canadian Pacific (CP) and Canadian National (CN) railways over the period of 1956-81. Tretheway et al. (1997) also conducted the same study with the index number method; but they extended the data to 1991 and found that although CP and CN sustained modest productivity growth throughout the period of 1956-1991, their performance slipped over the past decade, partly because of slower output growth. The cost function can also be used to measure the productivity. Caves et al. (1981) specified the variable cost function and adopted the least squares method to estimate the productivity growth of US railroads. They concluded that the behavioral assumptions underlying cost function analysis had important implications for the measurement of productivity growth. Friedlaender et al. (1993) selected labor, equipment, fuel, and materials and supplies as the inputs, ton-miles as the output, and then used the least squares method to estimate the short-run variable cost function of US class I railroads. They concluded that the institutional barriers to capital adjustment might be substantial; therefore, with respect to capital stock adjustment, the rail industry still had a long way to go. McGeehan (1993) also employed the least squares method to estimate the railway cost functions and found that the Cobb-Douglas function would not be appropriate for describing the production structure of Ireland railways. Atkinson and Cornwell (1998) proposed an alternative econometric framework for estimating and decomposing the productivity and then applied it to measure the productivity change for twelve US class I railroads over the period 1951 to 1975. The results concluded that a likelihood ratio test rejected the standard non-frontier specification. Total factor productivity (TFP) can be derived from a cost function since Caves et al. (1981). More recently, Loizides and Tsionas (2004) specified a translog cost function, using Monte Carlo simulation methods, to derive the exact distribution of productivity growth of ten European railways over the period 1969 to 1993, and to explore in detail how the productivity growth distribution shifts as a result of changes in input prices and output.

Oum and Yu (1994) adopted a two-stage DEA approach to evaluate the efficiency of 19 OECD countries' railways over the period of 1978 to 1989. The first stage was to measure efficiency by using DEA method and the second stage is to find out the factors that influence efficiency by using Tobit regression. The results indicate that the efficiency measures may not be meaningfully compared across railways without controlling for the effects of the differences in operating and market environments. Chapin and Schmidt (1999) used the DEA approach to measure the efficiency of US Class I railroad companies and found that the efficiency had been improved since deregulation, but not due to mergers. Cowie (1999) also applied the DEA method to compare the efficiency of Swiss public and private railways by constructing technical and managerial efficiency frontiers and then measured both efficiencies. It was found that private railways had higher technical efficiency than the public ones (89% versus 76%). Lan and Lin (2003) compared the difference of technical efficiency and service effectiveness for 76 worldwide railway systems with different DEA approaches, including conventional DEA, exogenously fixed inputs DEA (EXO DEA), and categorical DEA (CAT DEA) models. Their results showed that the efficiency and effectiveness estimated by EXO DEA and CAT DEA models were somewhat higher than those estimated by conventional DEA models because the environmental factors have been taken into account. Cantos and Maudos (2000) estimated productivity, efficiency and technical change for 15 European railways by using the SFA approach. Their results showed that the most efficient companies were those with higher degrees of autonomy. Cantos and Maudos (2001) also employed SFA to estimate both cost efficiency and revenue efficiency for 16 European railways. They concluded that the existence of inefficiency could be explained by the strong policy of regulation and intervention. Lan and Lin (2002) compared the performance difference for 85 worldwide railway systems measured by SFA and DEA approaches. The results indicated that different approach has led to different results and the Spearman rank correlation matrix of technical efficiency for SFA and DEA was 0.81. More recently, Lan and Lin (2004) proposed various stochastic distance function models to carry out performance evaluation for 46 worldwide railways by distinguishing the technical efficiency from the service effectiveness over the period of 1998-2000. The results showed that the percentage of electrified lines, population density, per capita gross national income and line density were the main factors affecting technical efficiency; while per capita gross national income, population density, ratio of passenger train-kilometers to total train-kilometers and line density were the main factors affecting service effectiveness. Kennedy and Smith (2004) applied two parametric techniques (COLS and SFA) to assess the relative efficiency of Railtrack's zones over the period 1995/96 to 2001/02. The results demonstrated that zonal differences in scale, technology, and other environmental factors are relatively small compared with external benchmarking studies.

From the literature review we found that most previous railway performance studies did not distinguish technical efficiency from technical effectiveness. Some others did not make distinction between cost efficiency and technical efficiency or between cost effectiveness and technical effectiveness. None have been endeavored to evaluating the service effectiveness and sales capability. As explained in the introduction, to elucidate the non-storable nature of railway transport, it is necessary to distinguish efficiency from effectiveness and to distinguish productivity from sales capability so that one could clearly diagnose the sources of any poor performance in order to propose more practical improvement strategies. In the context of international comparison, different countries' currencies may not be ready to convert into common currency due to copious fluctuations of exchange rates; or different railways' factor prices and sale revenues are often difficult to collect. Under this circumstance, one could only compare the technical efficiency (effectiveness) rather than the cost efficiency (effectiveness).

3. METHODOLOGIES

Conventional DEA approaches, such as CCR model proposed by Charnes et al. (1978) or BCC model proposed by Banker et al. (1984), have become increasingly widespread in the efficiency measurement in the past two decades. However, these conventional DEA approaches may lead to biased comparison among DMUs. First, they do not consider the difference of efficiency scores caused by environmental diversity across the DMUs. Second, they do not take statistical errors of data into consideration. Third, when measuring the efficiency, there are usually slacks in inputs or outputs, but conventional DEA approaches do not account for the slack effects. To explain the slacks, Figure 2 demonstrates with four DMUs (A, B, C, and D) that all produce a certain level of output y with two inputs x_1 and x_2. DMUs C and D are assumed efficient and located on the piecewise frontier (isoquant) composed of a vertical line ending at C, a line segment connecting C and D, and a horizontal line starting at D. DMUs A and B are assumed inefficient and can proportionally (in radial direction) reduce both of their inefficient inputs towards the frontier at E and F, respectively, to become "efficient." The point E is essentially efficient because it is a combination of two efficient points C and D, but the point F may not be efficient. In Figure 2, obviously, F can further curtail the input x_1 by S_2 and still produce the same amount of output y. In DEA literature, S_1 is termed as radial slack (measuring the magnitude of radial inefficiency for input x_1) and S_2 is defined as non-radial slack (measuring the magnitude of non-radial inefficiency for input x_1).

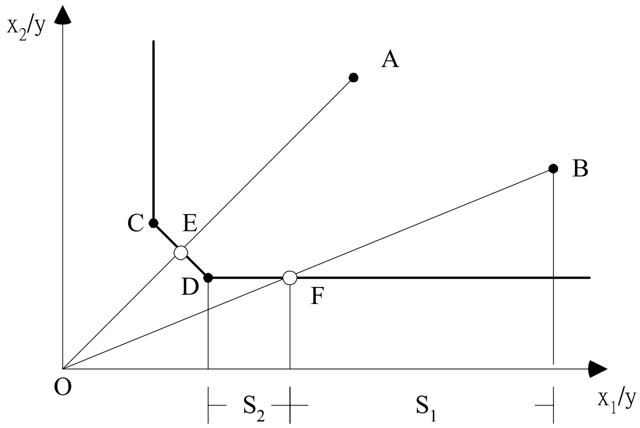


FIGURE 2: An illustration of radial and non-radial slacks by input-oriented DEA

The above shortcomings can significantly bias the relative efficiency scores, thus some researchers have devoted to improve the conventional DEA models. For instance, to take the non-discretionary environmental factors into account, Banker and Morey (1986a, 1986b) proposed an exogenously fixed inputs and outputs DEA model. They also introduced a categorical DEA model in which the DMUs are classified into several reference sets based on the operating environments. A specific DMU is then compared to other DMUs at the same rating of operating environments. To consider the effects of external operating environments, Fried et al. (1993) adopted conventional DEA approach to evaluate the performance of U.S. credit unions in the first stage and then regressed the sum of radial and non-radial slacks on some explanatory variables by using seemingly unrelated regression (SUR) in the second stage. Fried et al. (1999) also introduced a procedure to obtain the measure of managerial efficiency that controls for the exogenous features of operating environments. To further decompose the slacks into environmental effect, statistical noise, and managerial efficiency, Fried et al. (2002) proposed a three-stage DEA approach. In the first stage, conventional DEA is applied to measure the preliminary efficiency score for each DMU. In the second stage, the total slacks (radial and non-radial slacks) are regressed by the environmental factors using stochastic frontier analysis (SFA), which can decompose the slacks into environmental effect, managerial efficiency and statistical noise. In the third stage, input or output data (depending on the orientation used in the first stage) are adjusted and then the performance is re-evaluated by DEA. Although Fried's three-stage DEA has taken the environmental effects and statistical noise into account, they did not adjust the slack effects, thus the results can still be biased. In order to overcome this problem, this paper proposes a four-stage DEA approach, which is elaborated as follows.

3.1 Technical efficiency measurement

In the first stage, we use input-oriented DEA (measuring the maximum possible proportional reduction in all inputs, keeping all outputs fixed) to measure the technical efficiency (a transformation of inputs to outputs). Assume that there are J DMUs, each of which produces K products by utilizing M input factors; the input-oriented BCC model is specified as follows (Banker et al., 1984).

$$\underset{\theta,\lambda}{\text{Minimize}} \quad \theta$$

subject to

$$\begin{aligned} -y_j + \sum\nolimits_{j\in J} y_{kj}\cdot\lambda_j \geq 0, \quad k = 1,\ldots,K, \\ \theta\cdot x_j - \sum\nolimits_{j\in J} x_{mj}\cdot\lambda_j \geq 0, \quad m = 1,\ldots,M, \\ \sum\nolimits_{j\in J}\lambda_j = 1, \quad \lambda_j \geq 0, \quad j = 1,\ldots,J, \end{aligned} \tag{1}$$

where x_{mj} is the mth input and y_{kj} is the kth output for the jth DMU, respectively; λ_j is a constant and θ is a scalar standing for efficiency of the jth DMU. Solving this LP, one obtains the efficiency score for each DMU. As illustrated in Figure 2, the slack problem arises because model (1) uses piecewise linear segments to represent the efficient frontier.

In the second stage, factors affecting the slacks (the magnitudes of inefficiency for inputs) are further investigated. We regress the sum of radial and non-radial slacks on potential environmental factors by using SFA (Aigner et al., 1977). Thus, the sum of slacks can be decomposed into environmental influences, managerial inefficiency and statistical error (data noise) terms by the following:

$$S_{mj} = f_m(\omega_{ij};\delta_{mi}) + v_{mj} + u_{mj}, \quad m = 1,\ldots,M;\ j = 1,\ldots,J, \tag{2}$$

where dependent variables S_{mj} are the sum of radial and non-radial slacks estimated in the first stage; ω are the corresponding environmental factors and δ are the parameters to be estimated; f_m (ω_{ij}; δ_{mi}) is the deterministic slack frontier of mth input; v_{mj} is the statistical noise and u_{mj} represents the managerial inefficiency. Assume that v_{mj} follows a normal distribution with zero mean and variance σ_v^2 and u_{mj} is a positive half-normal distribution with mean μ and variance σ_u^2, and that v_{mj} is independent of u_{mj}.

In the third-stage, the adjusted inputs are constructed from the estimated results of (2) by using

$$x_{mj}^A = x_{mj} + \left[max_j(\omega_{ij}\hat{\delta}_{mj}) - \omega_{ij}\hat{\delta}_{mj}\right] + \left[max_j(\hat{v}_{mj}) - \hat{v}_{mj}\right], \quad m = 1,\ldots,M;\ j = 1,\ldots,J \ , \tag{3}$$

where x_{mj}^A and x_{mj} are adjusted and observed input quantities, respectively. This adjustment will put all DMUs into a common platform of operating environment and a common state of nature (Fried et al., 2002). The DEA-based efficiency for each DMU can be re-estimated again by substituting the adjusted data into (1) with which the environmental and statistical effects have been incorporated. However, such inputs adjustment in the third stage still does not account for the slack effects and thus a slack adjustment is further required (see, Sueyoshi (1999), Sueyoshi et al. (1999), Hibiki and Sueyoshi (1999), Sueyoshi and Goto (2001)).

In the fourth-stage, we further adjust the effect of slacks. The slack-adjusted (SA) model as shown in (4) counts the slacks in one dimension (Sueyoshi, 1999); however, the results are likely biased if slacks occur in two or more dimensions. To avoid this problem, we adopt Coelli's (1998) multi-stage model to estimate efficiency and slacks, and then substitute the results into the objective function of (4) to get the slack-adjusted technical efficiencies.

$$\text{Minimize} \quad \theta - \frac{1}{M+K}\left[\left(\sum_{m=1}^{M}(s_m^- / R_m^-)\right) + \left(\sum_{k=1}^{K}(s_k^+ / R_k^+)\right)\right]$$

subject to

$$\begin{aligned} &- y_{kj} + \sum\nolimits_{j\in J} y_{kj}\lambda_j - s_k^+ = 0, \quad k = 1,\ldots,K, \\ &\theta \cdot x_{mj} - \sum\nolimits_{j\in J} x_{mj}\lambda_j - s_m^- = 0, \quad m = 1,\ldots,M \ , \\ &\sum\nolimits_{j\in J}\lambda_j = 1, \quad \lambda_j \ge 0, \quad j = 1,\ldots,J, \theta : free \ , \end{aligned} \tag{4}$$

where s_m^- and s_k^+ are input and output slacks, respectively, $R_m^- = \max_j x_{mj} (m = 1,...,M)$ and $R_k^+ = \max_j y_{kj} (k = 1,...,K)$.

3.2 Effectiveness measurement

Similar to the aforementioned efficiency measurement, a four-stage DEA approach is also applied to the service effectiveness measurement (a transformation of outputs to consumptions). We measure the service effectiveness for each DMU by employing consumption-oriented DEA (measuring the maximum possible proportional expansion in all consumptions while all outputs remaining unchanged). In the first-stage, assume that K outputs (y_k) are transformed to Q consumptions (z_q), the consumption-oriented BCC model is then specified as follows.

$$\underset{\phi,\lambda}{\text{Maximize}} \quad \phi$$

subject to

$$-\phi \cdot z_j + \sum\nolimits_{j \in J} z_{qj} \cdot \lambda_j \geq 0, \quad q = 1, \ldots, Q, \tag{5}$$

$$y_j - \sum\nolimits_{j \in J} y_{kj} \cdot \lambda_j \geq 0, \quad k = 1, \ldots, K,$$

$$\sum\nolimits_{j \in J} \lambda_j = 1, \quad \lambda_j \geq 0, \quad j = 1, \ldots, J,$$

where z_{qj} is the qth consumption of jth DMU, y_j and λ_j are defined as (1); ϕ denotes proportional increase in consumptions, ranging from one to infinity, which could be achieved by the jth DMU without changing the output levels; $1/\phi$ defines the service effectiveness of each DMU, which varies between zero and one. DMU is effective if $1/\phi$ is equal to one and is ineffective if $1/\phi$ is less than one.

In the second- and third-stage, same procedures as the aforementioned efficiency measurement are applied. In the fourth-stage, the SA model as shown in (6) is used to adjust the slacks. Likewise, we also adopt Coelli's (1998) multi-stage model to estimate the effectiveness and slacks and then substitute the results into the objective function of (6) to get the slack-adjusted service effectiveness.

$$\text{Minimize} \quad \phi + \frac{1}{K+Q}\left[\left(\sum_{k=1}^{K}(s_k^- / R_k^-)\right) + \left(\sum_{q=1}^{Q}(s_q^+ / R_q^+)\right)\right]$$

subject to

$$-\phi \cdot z_{qj} + \sum\nolimits_{j \in J} z_{qj}\lambda_j - s_q^+ = 0, \quad q = 1, \ldots, Q, \tag{6}$$

$$y_{kj} - \sum\nolimits_{j \in J} y_{kj}\lambda_j - s_k^- = 0, \quad k = 1, \ldots, K,$$

$$\sum\nolimits_{j \in J} \lambda_j = 1, \quad \lambda_j \geq 0, \quad j = 1, \ldots, J, \theta : \textit{free},$$

where s_k^- and s_q^+ are output and consumption slacks, respectively, $R_k^- = \max_j y_{kj} (k = 1,\ldots,K)$ and $R_q^+ = \max_j z_{qj} (q = 1,\ldots,Q)$.

3.3 Productivity measurement

Malmquist index was first proposed in the consumer context (Malmquist, 1953). Caves et al. (1982) further introduced two theoretical indexes, named Malmquist input and output productivity indexes. Färe et al. (1989) exploited the fact of Malmquist indexes as ratios of distance functions and the distance functions to be reciprocal to Farrell's (1957) measurement of technical efficiency. Färe et al. (1994) assumed the production technology to be constant returns to scale and free disposability for inputs and outputs, thus an input-based Malmquist productivity index (MPI), denoted as m_I, could be expressed as follows.

$$m_I(y_s, x_s, y_t, x_t) = \left[\frac{d_I^s(y_t, x_t)}{d_I^s(y_s, x_s)} \times \frac{d_I^t(y_t, x_t)}{d_I^t(y_s, x_s)}\right]^{1/2}, \tag{7}$$

where y_s, y_t, x_s, x_t represent outputs (y) and inputs (x) at periods s and t, respectively. We adopt Färe's et al. input-based MPI rather than output-based one since our objective is to look for a minimal proportional contraction of the input vector, given an output vector. Thus $d_I^t(y_t, x_t)$ in (7) stands for the input-oriented distance between the observation (y_t, x_t) at period t and the production frontier at period t. The m_I can further be decomposed into two terms: efficiency change (ΔI) and productive technology change (ΔP), as shown in (8).

$$m_I(y_s, x_s, y_t, x_t) = \frac{d_I^t(y_t, x_t)}{d_I^s(y_s, x_s)} \times \left[\frac{d_I^s(y_t, x_t)}{d_I^t(y_t, x_t)} \times \frac{d_I^s(y_s, x_s)}{d_I^t(y_s, x_s)} \right]^{1/2} . \quad (8)$$

The first term, ΔI, captures the catching-up effect; the second term, ΔP, measures the movement of the frontier. To measure the m_I, Färe et al. (1994) proposed to calculate four distance functions by using linear programming technique (hereafter, called FGNZ method). It should be noted, however, that when solving for the four LPs one would employ the CCR model (see Charnes et al. (1978)) rather than BCC model. The reasons for adopting CCR model can be found in Färe et al. (1994, 1997). Also note that there are few shortcomings in FGNZ method where the solutions of LPs frequently contain slacks that are typically ignored. When slacks are present, radial efficiency measures will overstate the true efficiency and thus affects the productivity index in an unknown way. For example, assume that there is no technical change between period t and t+1, namely the DMUs face the identical frontier, and that the measured DMU is located on the frontier in both t and t+1 periods with non-radial slacks of S_t and S_{t+1} ($S_t > S_{t+1}$) respectively. The conventional DEA-like Malmquist index method will lead to a result that there has no productivity improvement. However, the definition of productivity tells us that this result is biased. In addition, the FGNZ method does not take environmental effects and statistical noise into account.

To measure MPI more precisely, we solve four distance functions by substituting the adjusted data, directly obtained from the third-stage of the four-stage DEA efficiency measurement, and adopting SA model (4) (hereafter, called four-stage method in contrast to FGNZ method). Consequently, the effects of environmental factors, statistical noise and slacks are all considered in our proposed four-stage method. While measuring the productivity of non-storable rail transport service, some previous studies utilized passenger-km and ton-km as "outputs" (in fact they are "consumptions"). In this paper, we would measure the productivity by the input-based Malmquist productivity index, thus passenger-train-km and freight-train-km will be used as outputs rather than passenger-km and ton-km.

3.4 Sales capability measurement

The sales capability index will be used to define the transformation ability of a railway outputs to consumptions. The relationship between sales capability index and productivity index is similar to the relationship between service effectiveness and technical efficiency. Productivity index, corresponding to technical efficiency, can be viewed as a ratio of outputs to inputs; while sales capability index, corresponding to service effectiveness, can be viewed as a ratio of consumptions to outputs. Since we look for a maximal proportional expansion of the consumption vector, given an output vector, the consumption-based Malmquist sales capability index (MSI), denoted as m_C, can be defined as follows.

$$m_C(z_s, y_s, z_t, y_t) = \left[\frac{d_C^s(z_t, y_t)}{d_C^s(z_s, y_s)} \times \frac{d_C^t(z_t, y_t)}{d_C^t(z_s, y_s)} \right]^{1/2} \tag{9}$$

where z_s, z_t, y_s, y_t stand for consumptions (z) and outputs (y) at periods s and t, respectively; $d_C^t(z_t, y_t)$ represents the consumption-oriented distance between the observation (z_t ,y_t) at period t and the sales frontier at period t. Likewise, m_C can be decomposed into two terms: effectiveness change (ΔE) and sales innovation change (ΔS) as follows.

$$m_C(z_s, y_s, z_t, y_t) = \frac{d_C^t(z_t, y_t)}{d_C^s(z_s, y_s)} \times \left[\frac{d_C^s(z_t, y_t)}{d_C^t(z_t, y_t)} \times \frac{d_C^s(z_s, y_s)}{d_C^t(z_s, y_s)} \right]^{1/2} \tag{10}$$

Similarly, in order to measure MSI more precisely, we solve four distance functions by substituting the adjusted data, directly obtained from the third-stage of the four-stage DEA effectiveness measurement, into (10) and then measure the four distance functions by adopting SA model (6) (hereafter, also called four-stage method in contrast to FGNZ method). Again, our proposed four-stage method accounts for the effects of environmental factors, statistical noise and slacks simultaneously.

4. EMPIRICAL ANALYSIS

4.1 Data

In the present paper we focus on multi-product railways which provide both passenger and freight services. The single-product railways providing only passenger or freight service are not studied here. Since we attempt to investigate how external factors affecting the efficiency (effectiveness) measures, those railways with incomplete data, including two consumptions, two outputs, four inputs, two external and two internal variables, in our study horizon will not be analyzed. Our data set, drawn from International Railway Statistics published by the International Union of Railways (UIC), contains 350 panel data composed of 50 railways covering seven years (1995-2001). Since DEA measures the relative efficiency (effectiveness) of each observation to the most efficient (effective) DMUs, the results might be significantly affected by the influential observations (i.e., outliers). Therefore, it is important to detect the outliers from the samples. We conduct a boxplot test and identify six outliers. After removing these outliers, our final data set only contains 44 railways, including 308 data points.

Table 1 summarizes the descriptive statistics of these 308 data points, including two consumptions (passenger-kilometers and ton-kilometers), two outputs (passenger train-kilometers and freight train-kilometers), four inputs (length of lines, number of passenger cars, number of freight cars, and number of employees), two external (environmental) variables (per capita gross national income and population density), and two internal variables characterizing the railways (percentage of electrified lines and ratio of passenger train-kilometers to total train-kilometers). One can easily find that the data are rather heterogeneous. Take GNI as an example, the data ranges from 220 to 45,060 US dollars, and the standard deviation is 13,086 US dollars. It reveals that the

TABLE 1: Descriptive statistics of the 308 DMUs (44 railways over 7 years: 1995-2001)

Statistics	Consumptions		Outputs		Inputs				External variables		Internal variables	
	pax-km (10^6)	ton-km (10^6)	pax train-km (10^3)	freight train-km (10^3)	length of lines (km)	pax cars	freight cars	labors	GNI	PD	ELEC (%)	ROP (%)
Max.	457022	312371	739800	260594	62915	36621	467884	1602051	45060	615	1.000	0.964
Min.	74	265	553	832	220	40	162	1212	200	10	0.000	0.156
Mean	23995	21414	91782	32366	8179	4286	34124	86131	13604	127	0.387	0.666
Std. dev.	67626	49216	158003	52901	12190	7165	62917	240308	13085	116	0.285	0.171

Note: GNI denotes per capita gross national income (US dollar) and PD denotes population density (persons per square kilometer) of the country to which the railway belongs. ELEC represents the percentages of lines being electrified. ROP is defined as the ratio of passenger train-kilometers to total train-kilometers.

environments faced by different railways are quite varied; thus, we must consider the effects of environmental factors on the variation of efficiency (effectiveness) scores. Due to data availability, we do not consider such factors as state/private ownership or regulatory differences across the railways.

For measuring the rail technical efficiency, some studies selected passenger train-kilometers and freight train-kilometers as outputs, number of employees, number of cars and length of lines as inputs (for example, Coelli and Perelman (2000)). We do not directly use length of lines as an input factor for two reasons. First, for rail transport industry, line-related facilities such as tracks, signals, stations and yards should be viewed as sunk, which are attributed to "fixed" costs. In this paper, we attempt to measure the efficiency of "variable" input factors. Second, the length of lines for these 44 railways ranges from 220 to 62,915 kilometers, which are rather heterogeneous. To account for the heterogeneous network scale and for a more homogeneous set of DMUs, where comparison makes more sense, we measure the technical efficiency by selecting number of passenger cars per kilometer of lines, number of freight cars per kilometer of lines, and number of employees per kilometer of lines as input factors and passenger-train-kilometer per kilometer of lines and freight-train-kilometer per kilometer of lines as output variables. In measuring the service effectiveness, on the other hand, we choose passenger-kilometers and ton-kilometers as two consumptions and passenger train-kilometers and freight train-kilometers as two outputs.

4.2 Results

For the purpose of comparison, the efficiency and effectiveness scores are estimated by three DEA approaches: BCC model, Fried's et al. three-stage DEA approach and our proposed four-stage DEA approach. The DEA is solved by DEAP version 2.1 (Coelli, 1996a) and checked by GAMS computer software (Brooke et al., 1998). The SFA is estimated by FRONTIER 4.1 (Coelli, 1996b). The detailed results for each DMU by these three DEA approaches are presented in Appendix 2, which reports the average scores during the study horizon from 1995 to 2001. Table 2 further summarizes the distribution of efficiency and effectiveness scores by these three DEA approaches. Based on the results and some extended analyses, we draw important findings as follows.

TABLE 2: Frequency distribution of efficiency and effective scores by three different DEA approaches

Range of scores	Efficiency measurement			Effectiveness measurement		
	BCC	3-stage	4-stage	BCC	3-stage	4-stage
Less than 0.2	15	0	0	16	0	0
0.200~0.299	16	0	0	89	2	2
0.300~0.399	37	0	0	56	4	5
0.400~0.499	53	0	2	20	6	5
0.500~0.599	22	0	3	22	2	3
0.600~0.699	33	0	27	23	1	0
0.700~0.799	23	6	59	23	16	18
0.800~0.899	28	91	64	15	40	42
0.900~0.999	32	178	80	26	217	213
1.000	49	33	32	18	20	20
Max.	1.000	1.000	1.000	1.000	1.000	1.000
Min.	0.143	0.752	0.409	0.177	0.247	0.223
Mean	0.639	0.924	0.849	0.497	0.923	0.917
Std. Dev.	0.269	0.054	0.109	0.271	0.130	0.135

Finding 1. Efficiency (effectiveness) scores by BCC model are relative low and varied among regions

Based on the BCC model, in general, rail transport services are characterized with rather low efficiency (effectiveness) scores. For the whole industry, the average efficiency score is only 0.639, while average effectiveness score is 0.497 (Table 2). We further adopt Kruskal-Wallis rank test to examine whether or not the scores vary among regions. The samples are divided into four regions, which are West Europe, East Europe, Asia (Oceania included), and Africa (Mid-East included). The statistic proposed by Hays (1973) is used for the rank test:

$$H = \frac{12}{J(J+1)}\left[\sum_{p}\frac{T_p^2}{n_p}\right] - 3(J+1) \tag{11}$$

where T_p is the sum of ranks for group p, n_p is the number of data points in the group p and J is total number of data points, that is 308. The testing result indicates that the null hypothesis of scores invariance among regions should be rejected; that is, both efficiency and effectiveness scores vary among these four regions. We find that, on average, African railways have the worst performance while West European railways have the best performance in both technical-efficiency and service-effectiveness measurements.

Finding 2. Some efficient (effective) DMUs are rather robust (insensitive) but some others are very sensitive to data change

Many researchers criticize the robustness of DEA because the efficiency scores may be very sensitive to data change, for example, Charnes and Neralic (1990), Charnes et al. (1992), Zue (1996), Seiford and Zue (1998a,b). To investigate which DMUs are sensitive to possible data change, Seiford and Zue (1998b) consider the case when all data are changed simultaneously by solving the following LP model.

$$\beta^* = \text{Min}\ \beta$$

subject to (12)

$$\sum_{j=1,\, j\neq O}^{J}\lambda_j x_{mj} \le \beta_{mO} x_{mO}, \quad \sum_{j=1,\, j\neq O}^{J}\lambda_j y_{kj} \ge y_{kO}, \quad \sum_{j=1,\, j\neq O}^{J}\lambda_j = 1, \quad \beta, \lambda_j \ge 0, (j \neq O)$$

They show that under the circumstance of $1 \le \sqrt{\beta^*}$, where β^* is the optimal value to (12), an efficient DMU_O with efficiency score equal to 1.000 will still remain efficient, provided that the percentages increase in all inputs for the DMU_O are less than $g_O = \sqrt{\beta^*} - 1$ and the percentages decrease in all inputs for the remaining DMUs are less than $g_{-O} = (\sqrt{\beta^*} - 1)/\sqrt{\beta^*}$. The upper-bound levels (g_O, g_{-O}) can be viewed as the sensitivity indexes. The results of Seiford and Zue's sensitivity analysis for efficiency measurement are indicated in Table 3. For instance, the efficient DMU 149 (CFF, 98), DMU 281 (CFF, 2001) and DMU 306 (TRA, 2001) are rather robust (stable) because their sensitivity indexes are relative large (higher than 15%), suggesting that they are not sensitive to possible data change. In contrast, the efficient DMU 44 (QR, 95), DMU 125 (TRC, 97), DMU 176 (QR, 98), DMU 179 (DSB, 99), DMU 191 (NSB, 99), DMU 220 (QR, 99), DMU 257 (TRC, 2000), DMU 264 (QR, 2000), DMU 278 (SJ, 2001), DMU

279 (NSB, 2001) and DMU 286 (GYSEV, 2001) are very sensitive to possible data change because they have relatively small sensitivity indexes (less than 1%).

TABLE 3: Sensitivity indexes of efficient DMUs by input-oriented DEA (BCC model)

Railway	g_O	g_{-O}	Railway	g_O	g_{-O}	Railway	g_O	g_{-O}
DMU10	4.18%	4.02%	DMU176	0.14%	0.14%	DMU242	4.52%	4.33%
DMU11	6.71%	6.28%	DMU179	0.90%	0.89%	DMU257	0.54%	0.54%
DMU14	4.59%	4.39%	DMU191	0.89%	0.88%	DMU264	0.49%	0.49%
DMU42	8.15%	7.53%	DMU192	6.53%	6.13%	DMU265	8.56%	7.89%
DMU44	0.99%	0.98%	DMU198	6.51%	6.12%	DMU267	11.93%	10.66%
DMU58	8.10%	7.50%	DMU213	5.83%	5.51%	DMU275	5.71%	5.40%
DMU66	2.52%	2.45%	DMU216	10.11%	9.18%	DMU278	0.27%	0.27%
DMU81	2.61%	2.54%	DMU220	0.01%	0.01%	DMU279	0.84%	0.83%
DMU102	5.48%	5.20%	DMU221	5.40%	5.13%	DMU280	3.52%	3.40%
DMU110	1.82%	1.79%	DMU223	6.10%	5.75%	DMU281	15.81%	13.65%
DMU125	0.06%	0.06%	DMU226	2.70%	2.63%	DMU286	0.92%	0.91%
DMU128	1.67%	1.64%	DMU231	2.25%	2.20%	DMU301	2.65%	2.58%
DMU139	5.34%	5.07%	DMU234	3.86%	3.72%	DMU304	12.20%	10.87%
DMU147	3.19%	3.09%	DMU235	4.82%	4.60%	DMU306	15.50%	13.42%
DMU148	2.16%	2.11%	DMU236	4.37%	4.19%	DMU308	3.14%	3.05%
DMU149	16.03%	13.82%	DMU237	2.37%	2.32%			
DMU169	7.13%	6.65%	DMU241	4.23%	4.06%			

Note: g_O denotes the percentages increase in all inputs for the DMU_O, and g_{-O} denotes the percentages decrease in all inputs for the remaining DMUs

Similarly, consider the following LP model

$$\alpha^* = \text{Max } \alpha$$

subject to (13)

$$\sum_{j=1, j\neq O}^{J} \lambda_j y_{kj} \le y_{kO}, \quad \sum_{j=1, j\neq O}^{J} \lambda_j z_{qj} \ge \alpha \cdot z_{qO}, \quad \sum_{j=1, j\neq O}^{J} \lambda_j = 1, \quad \alpha, \lambda_j \ge 0, (j \neq O)$$

Seiford and Zue (1998b) also show that under the circumstance of $\sqrt{\alpha^*} \le 1$, where α^* is the optimal value to (13), an efficient DMU_O will remain efficient, provided that the percentages decrease in all outputs for the DMU_O are less than $h_O = 1 - \sqrt{\alpha^*}$ and the percentages increase in all outputs for the remaining DMUs are less than $h_{-O} = (1-\sqrt{\alpha^*})/\sqrt{\alpha^*}$. The upper-bound levels (h_O, h_{-O}) are the sensitivity indexes. The results of Seiford and Zue's sensitivity analysis for effectiveness measurement are indicated in Table 4. For example, the effective DMU 36 (CFM, 95), DMU 66 (GYSEV, 96), DMU 81 (TRC, 96) and DMU 227 (CH, 2000) are robust because their sensitivity indexes are rather large (higher than 15%), implying that they are not sensitive to possible data change. In contrast, the effective DMU 84 (JR, 96), DMU 251 (UZ, 2000) and DMU 295 (UZ, 2001) are very sensitive to possible data change because they have relatively small sensitivity indexes (less than 1%).

Finding 3. The total slacks and average slacks by three-stage DEA approach are smaller than those by BCC model

The input-oriented (consumption-oriented) DEA approach imposes a piecewise linear production (consumption) frontier to input-output (output-consumption) data set, thus both radial and non-radial slacks may simultaneously appear in the estimated results. Table 5 summarizes the results of slack analysis by BCC model and Fried's three-stage DEA approach. It shows that both input- and consumption-oriented estimation results exhibit a large amount of input and consumption slacks. Taking the BCC effectiveness measurement as an example, the consumption slacks for passenger-kilometer and ton-kilometer are 7,247,057 (6,608,582 in radial plus 638,475 in non-radial) and 7,079,282 (7,011,008 in radial plus 68,274 in non-radial), respectively. As anticipated, the total slacks and average slacks (TS and AS in Table 5) by three-stage DEA approach are smaller than those by BCC model, suggesting that the estimated results are seriously biased if one were not to consider the effects of environmental factors and statistical noise.

TABLE 4: Sensitivity indexes of effective DMUs by consumption-oriented DEA (BCC model)

Railway	h_O	h_{-O}	Railway	h_O	h_{-O}	Railway	h_O	h_{-O}
DMU11	1.18%	1.20%	DMU80	7.45%	8.05%	DMU251	0.20%	0.20%
DMU31	2.16%	2.21%	DMU81	17.39%	21.05%	DMU260	1.43%	1.45%
DMU36	16.61%	19.92%	DMU84	0.10%	0.10%	DMU285	7.48%	8.08%
DMU37	2.26%	2.31%	DMU110	8.11%	8.82%	DMU295	0.81%	0.82%
DMU44	1.19%	1.21%	DMU227	15.80%	18.76%	DMU305	3.63%	3.76%
DMU66	16.13%	19.23%	DMU250	10.58%	11.83%	DMU308	5.92%	6.29%

Note: h_O denotes the percentages decrease in all consumptions for the DMU_O, and h_{-O} denotes the percentages increase in all consumptions for the remaining DMUs

Finding 4. The significant external and internal factors affect the input and consumption slacks

We regress the input and consumption slacks (TS values of BCC model in Table 5) on the external and internal factors (defined in Table 1), respectively, by using SFA (2). The estimated results are reported in Table 6, from which we find that most parameters are significant to the magnitude of slacks (i.e., the inputs inefficiency or consumptions ineffectiveness). It should be noted that negative sign represents an opposite direction to the magnitude of slacks. For the input slacks, higher percentage of electrified lines or higher ratio of passenger service can lower the magnitude of input slacks. Positive sign in the coefficient of length of line (LINE) indicates that larger scale of railway will increase the magnitude of input slacks. On the other hand, for the consumption slacks, negative sign in the coefficient of PD implies that higher population density can lower the magnitude of consumption slacks. Positive sign in the coefficient of GNI indicates that higher income per capita will increase the magnitude of consumption slacks. This reflects the fact that higher GNI will generally lead to higher private car ownership thus lower the public transport usage. Similar to the input slacks; positive sign in the coefficient of LINE implies that larger scale of railway generally creates greater consumption slacks both in passenger and freight services.

Finding 5. Efficiency (effectiveness) scores by three-stage DEA approach are considerably higher than those by BCC model

Once the parameters (Table 6) are estimated, the input and consumption data can then be adjusted by (3). We therefore use the adjusted data to re-estimate the efficiency (effectiveness) scores by (1). Table 2 indicates that the efficiency and effectiveness

TABLE 5: Input and consumption slacks by BCC model and 3-stage DEA approach

DEA model		Input slacks						Consumption slacks			
		Employee		Pax-cars		Fre-cars		Pax-km		Ton-km	
		Radial	Non-rad.	Radial	Non-rad.	Radial	Non-rad.	Radial	Non-rad.	Radial	Non-rad.
BCC	No.	260	96	260	11	260	141	290	57	290	19
	TS	1182.3	161.5	65.1	16.6	726	141.3	6,608,582	638,475	7,011,008	68,274
	AS	3.839	0.524	0.211	0.054	2.357	0.459	21,456	2,073	22,763	222
3-stage	No.	275	44	275	67	275	20	288	45	288	13
	TS	327.8	159.5	41.5	8.1	262.8	18.9	2,711,095	392,360	2,855,532	2,911
	AS	1.064	0.518	0.135	0.026	0.853	0.061	8,802	1,273	9,271	9.5

Note: No., TS, and AS stand for number of DMUs with slacks, total slacks, average slacks (defined as TS / 308), respectively.

scores re-estimated from the adjusted data (Fried's three-stage DEA approach) are considerably higher than those estimated from the unadjusted data (BCC model), 0.924 vs. 0.639 and 0.923 vs. 0.497, respectively. We also note that the standard deviation of efficiency (effectiveness) scores has decreased from 0.269 (0.271) to 0.054 (0.130) and the number of high efficient (effective) railways has drastically increased after the data being adjusted. For instance, the number of DMUs with efficiency (effectiveness) scores greater or equal to 0.9 is changed from 81 (44) by BCC model to 211 (237) by three-stage DEA approach. Obviously, the results by three-stage approach are more reasonable than those by BCC model because both the environmental factors and statistical noise have been taken into account.

TABLE 6: Factors affecting input and consumption slacks by SFA

Input slacks				Consumption slacks		
Parameters	Employee	Pax-cars	Fre-cars	Parameters	Pax-km	Ton-km
Constant	1.457* (10.472)	0.731* (5.867)	-1.217* (-5.272)	Constant	-5.092* (-6.854)	-2.745* (-5.593)
ln(ELEC)	-0.327* (-4.561)	-0.255* (-3.873)	-0.462* (-4.741)	ln(PD)	-2.183* (-3.409)	-0.258* (-5.383)
ln(ROP)	-2.546* (-5.551)	-0.106 (-0.344)	-3.200* (-5.760)	ln(GNI)	0.605* (14.461)	0.297* (8.551)
ln(LINE/1000)	0.195* (6.156)	0.060* (1.991)	0.055 (1.215)	ln(LINE/1000)	1.076* (15.688)	1.315* (26.333)
σ^2	15.639* (2.450)	9.397* (3.662)	14.621 (1.140)	σ^2	10.390* (5.729)	10.559* (2.685)
γ	0.996* (413.306)	0.997* (555.945)	0.989* (112.241)	γ	0.987* (275.729)	0.987* (251.639)
μ	-7.893* (-1.974)	-6.121* (-2.886)	-6.445 (-0.745)	μ	-6.404* (-3.935)	-6.456* (-1.799)
Log likelihood function	-329.023	-259.455	-355.538	Log likelihood function	-410.812	-403.307
LR one-sided test	98.370	106.256	61.975	LR one-sided test	129.93	101.97

Note: t-values in parentheses, asterisks (*) represent significant at the 0.05 level. Also note that $\sigma^2 = \sigma_u^2 + \sigma_v^2, \gamma = \sigma_u^2 / \sigma^2$

Finding 6. Efficiency (effectiveness) scores by three-stage DEA approach are slightly overestimated in comparison with our proposed four-stage DEA approach

Table 5 shows the evidences that although the total and average slacks have been decreased by three-stage DEA approach, there still exist slack problems in both inputs and consumptions. Therefore, we further employ the proposed four-stage DEA approach to re-estimate the efficiency and effectiveness scores and the results are also presented in Appendix 2 and Table 2. Compared with Fried's three-stage approach, our four-stage DEA approach has 52 (199) DMUs remaining unchanged in the efficiency (effectiveness) measurement. On average, the efficiency and effectiveness scores estimated by four-stage approach are slightly less than those by three-stage approach. In other words, the efficiency and effectiveness scores are slightly overestimated by the three-stage DEA approach because the slacks are not adjusted.

Finding 7. Productivity growth measured by FGNZ method is overestimated in comparison with our proposed four-stage method

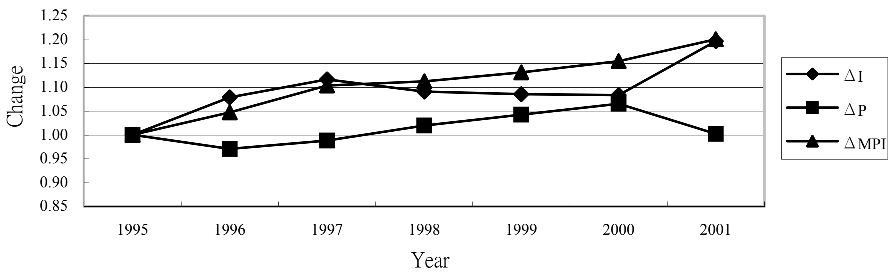
We measure the change in Malmquist productivity index (MPI) and its components for the 44 railway companies over the period of 1995-2001 by using both FGNZ method and our proposed four-stage method. The results are indicated in Appendix 3 and summarized in Table 7, and the time trends are presented in Figures 3a and 3b. It reveals that the productivity measured by FGNZ method is overestimated because of ignoring the slacks adjustment. These 44 railways have an average productivity growth of 20.2 percent over 1995-2001 by the FGNZ method; while the actual average productivity growth is only 6.6 percent by our four-stage method. The results also reveal that the productivity growth is due to improvements in efficiency (ΔI) rather than productive technology change (ΔP).

TABLE 7: Changes in Malmquist productivity index and its components (base year 1995)

Year	FGNZ method			Four-stage method		
	ΔI	ΔP	ΔMPI	ΔI	ΔP	ΔMPI
1995	1.000	1.000	1.000	1.000	1.000	1.000
1996	1.079	0.971	1.047	1.071	0.911	0.976
1997	1.117	0.988	1.104	1.083	0.918	0.994
1998	1.091	1.020	1.112	1.110	0.874	0.970
1999	1.086	1.043	1.131	1.190	0.851	1.013
2000	1.083	1.065	1.155	1.197	0.851	1.019
2001	1.197	1.003	1.202	1.126	0.947	1.066

Note: ΔI, ΔP and ΔMPI represent efficiency change, productive technology change and Malmquist total factor productivity change, respectively.

(a) FGNZ method



(b) Four-stage method

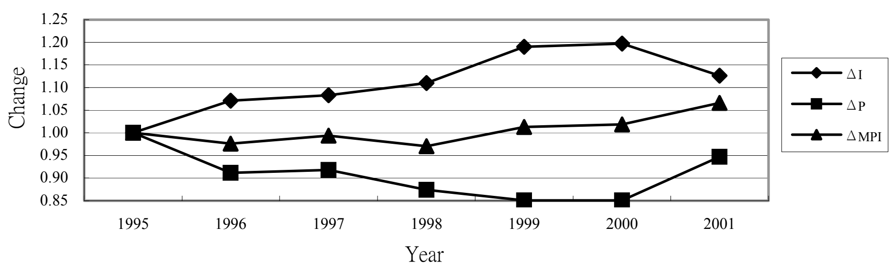


FIGURE 3: Changes in productivity index and its components

Finding 8. Sales capability growth measured by FGNZ method is slightly overestimated in comparison with our proposed four-stage method

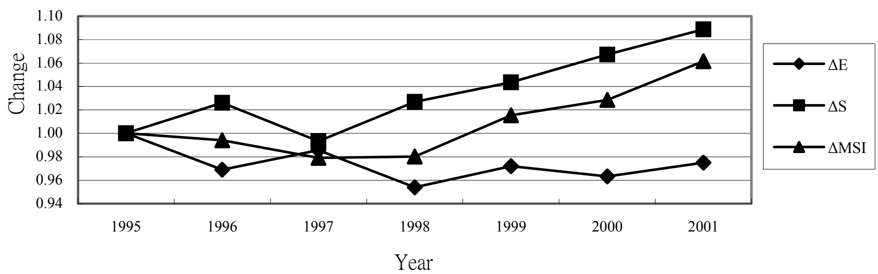
The Malmquist sales capability indexes are reported in Appendix 3 and summarized in Table 8, and the time trends and its components are depicted in Figures 4a and 4b. Based on the results, on average, sales capability grows at a rate of 7.3 percent over the period of 1995 to 2001 when adopting the FGNZ method. However, if we adjust the slacks by adopting the four-stage method, it becomes 6.1 percent. The results indicate that sales capability index is slightly overestimated if one does not take slacks adjustment into account. The results also reveal that the sales capability growth is due to sales innovation change (ΔS) rather than improvements in effectiveness (ΔE).

TABLE 8: Changes in Malmquist sales capability index and its components (base year 1995)

Year	FGNZ method			Four-stage method		
	ΔE	ΔS	ΔMSI	ΔE	ΔS	ΔMSI
1995	1.000	1.000	1.000	1.000	1.000	1.000
1996	0.969	1.026	0.994	0.978	1.014	0.992
1997	0.985	0.993	0.979	0.992	1.019	1.010
1998	0.954	1.027	0.980	0.990	1.030	1.019
1999	0.972	1.043	1.015	0.988	1.042	1.030
2000	0.963	1.067	1.029	0.998	1.058	1.055
2001	0.985	1.089	1.073	0.995	1.067	1.061

Note: ΔE, ΔS and ΔMSI stand for effectiveness change, sales innovation change and Malmquist sales capability change, respectively.

(a) FGNZ method



(b) Four-stage method

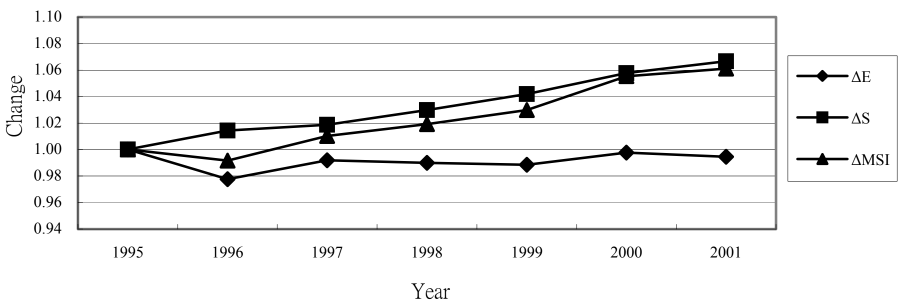


FIGURE 4: Changes in sales capability index and its components

5. POLICY IMPLICATIONS

In order to propose appropriate improvement operational strategies for different railways, we construct effectiveness-efficiency matrices as shown in Figures 5a (BCC model) and 5b (four-stage DEA approach). As anticipated, the number of DMUs in the third quadrant (both efficiency and effectiveness scores less than the mean values) in Figure 5b has been significantly decreased because the original heterogeneous DMUs have been adjusted to a common platform of operating environment and a common state of nature by our proposed 4-stage DEA approach. Since we adopt input-oriented DEA to measure the relative efficiency of railways, those railways in the second quadrant with low efficiency but high effectiveness should consider strategies of input factors curtailing to increase the technical efficiency. Our empirical analysis shows that (Table 5) the total slack of employee is 1,344 persons per kilometer of lines (1,182.3 in radial and 161.5 in non-radial), which is larger (in terms of the magnitude of value) than the total slacks of the other two input factors (82 passenger-cars per kilometer of lines and 867 freight-cars per kilometer of lines), hence, reducing the excess number of employees is perhaps more urgent than reducing the excess number of freight-cars than reducing the excess number of passenger-cars, provided input factor cutting strategies are to be considered.

Our results also indicate that percentage of electrified lines is a significant factor affecting the magnitude of input slacks as well as technical efficiency. In general, the efficient DMUs are those with high percentages of electrified lines. For example, the percentages of electrified lines of NS (Netherlands), SJ (Sweden) and BLS (Switzerland) are 0.727, 0.748, and 1.000 and their average efficiency scores in the study period are 0.972, 0.991 and 0.958, respectively, based on the BCC model. In contrast, the average efficiency scores of CFM (E) (Moldova), ONCFM (Morocco) and CFS (Syria) are 0.164, 0.400, and 0.337, and their percentages of electrified lines are all zero. The policy implication suggests that a railway company can enhance its technical efficiency by introducing more electrified lines.

Since a higher ratio of passenger train-kilometers to total train-kilometers (ROP) will generally lower the input slacks and as a result higher the technical efficiency. Our results indicate that some DMUs such as NS (Netherlands), DSB (Denmark) and JR (Japan) orient their rail service toward passenger transport (with average ROP values of 0.925, 0.874 and 0.899, respectively) and they experience significantly higher efficiency than those DMUs with low ROP values. This can be partly explained by the fact that the speeds (including loading and unloading at terminals) or frequencies of freight trains are generally much lower than the passenger trains. It could also be due to the national policy to provide guideway passenger transport to attract more private cars in these countries. Although the implication for raising the rail technical efficiency is to increase the share of passenger service rather than freight; yet railway is still the most effective freight mode in land transport, particularly for the low-valued bulky commodities such as raw materials, intermediate and final products. Rail freight service is rather labor intensive and time consuming, especially at the terminals where loading and unloading take place. Hence, expediting the process of freights at terminals by introducing fast loading and unloading equipment and advanced information and communication technologies would be critical to make the rail service more compatible with the trucking service. The intercity passenger trains or high-speed trains can also provide line-haul service for high-valued compact freights, such as express parcels, provided it is well integrated with the local pickup and delivery logistics.

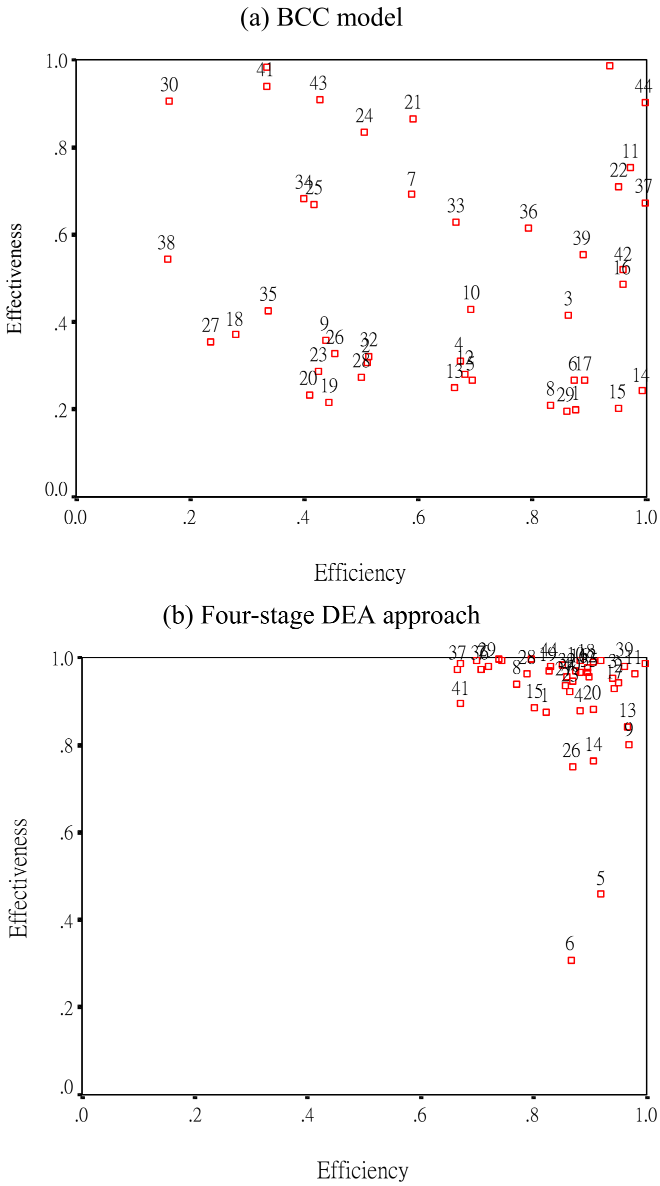


FIGURE 5: Effectiveness versus efficiency matrix

The strategies for improving the service effectiveness can be quite different from those for raising the technical efficiency. Since we adopt the consumption-oriented DEA approaches to measure the service effectiveness, those firms in the fourth quadrant with relative high efficiency but low effectiveness should devote to raising the consumption in passenger or freight or both to enhance the effectiveness. Our slack analysis shows that the total slack (radial and non-radial) of passenger-km is greater than that of ton-km, thus priority should be given in promoting the passenger services rather than the freight, which concurs with the implication of technical efficiency analysis by increasing the share of passenger service rather than freight. Our results also show that per capita gross national income (GNI) and population density (PD) are the two external factors significantly affecting the service effectiveness of railways. Although the operators can hardly control these two external factors to level up the service effectiveness, they can still consider various operational strategies, including increasing the punctual rate, replacing the over-aged assets (tracks and rolling stocks), rescheduling the trains better matching the demands, improving the booking system, and providing discounts to frequent users, to attract more patronages from competitive modes. Our results explicate that the selected 44 railways have an average of positive progress in both efficiency and effectiveness of recent years. The decline of rail market share in these countries would be attributed to higher level-of-service of other competitive modes, not to rail's poor performance in technical efficiency or service effectiveness.

In Figure 6, we further construct a similar matrix in which the changes in each railway's sales capability and productivity are indicated. We note that quite a number of railways have exhibited deterioration in productivity growth over 1995-2001. Since the MPI can be decomposed into efficiency change and productive technology change, it is necessary to find out the determinants causing productivity decline. If the source comes from efficiency drop, the strategies for improving efficiency described above are applicable. If the determinant is due to productive technology change, then introducing innovative production technologies should be a correct direction. In our analysis, the cumulative efficiency change, productive technology change, and Malmquist total factor productivity change over 1995-2001 are 1.197, 1.003 and 1.202 respectively based on the FGNZ method, and 1.126, 0.947 and 1.066, respectively based on the proposed four-stage method. In other words, the source of productivity growth is due to improvements in efficiency rather than productive technology change. Its policy implication strongly suggests improvement of productive technology be a critical direction for raising the productivity. Such strategies as improving the line geometry or introducing tilting trains to increase the train operating speed can be considered. Construction of high-speed rails, application of new technologies in signaling and traffic controls, upgrading the infrastructures (such as tracks) and facilities (such as loading and unloading equipment) can also be promising in raising the rail productivity.

From Figure 6 we also notice that several companies have revealed a decrease in sales capability over the same period. Similar to MPI, the MSI can be decomposed into effectiveness change and sales innovation change. Therefore, for those with sales capability decline, one requires further investigating the determinants of recession. If the effectiveness recession is the source, then the strategies for improving effectiveness described above may be applicable. If the deterioration is due mainly to sales problem, then improving effectiveness would be a wrong way. In this case, introducing innovative marketing techniques, such as new dispatching management information systems, automatic ticketing by vending machine, seat booking by internet and alliance with other firms, convenience stores or tourist agencies, could be good strategies. Our empirical analysis shows that the cumulative effectiveness change, sales innovation change, and Malmquist sales capability change over the seven years are 0.983, 1.092 and 1.073 respectively based on the FGNZ method and 0.994, 1.067, and 1.061 respectively based on the four-stage method. In other words, the source of sales capability growth is due to sales innovation change rather than effectiveness change. Its policy implication strongly suggests improvement of effectiveness be a critical direction for raising the sales capability. Therefore, the strategies for improving effectiveness described above can be applied to raise the sales capability.

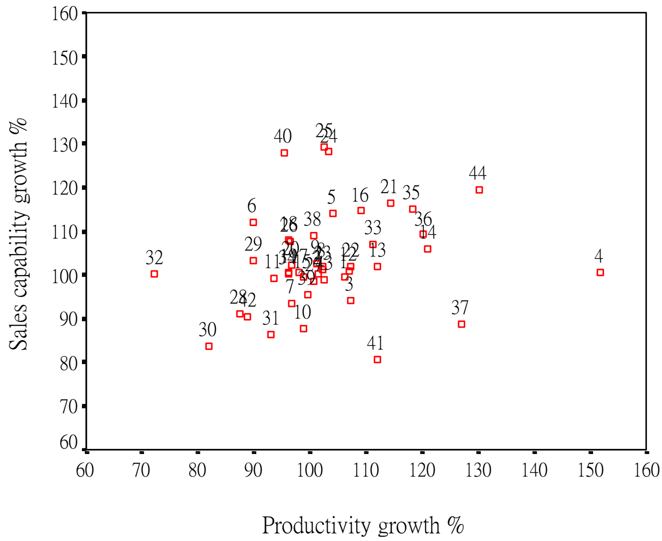


FIGURE 6: Sales capability growth versus productivity growth (Four-stage method)

6. CONCLUDING REMARKS

Conventional DEA approaches, such as CCR and BCC models, neither consider the environmental differences across the DMUs nor account for the statistical error (data noise) and slack effects. Thus, the comparison can be seriously biased because all DMUs are not brought into a common platform of operating environment and a common state of nature. To overcome these shortcomings, Fried et al. (2002) proposed a three-stage DEA approach with consideration of the environmental effects and statistical noise, but they still did not adjust the slack effects and thus the results could be biased as well. We propose a four-stage DEA approach by elaborating Fried’s three-stage DEA approach with further adjustment of slack effects. The empirical results show that our proposed four-stage DEA approach has slightly more reasonable efficiency and effectiveness scores than those measured by Fried’s three-stage DEA approach, which is far more reasonable than those measured by BCC model.

In measuring the productivity growth, FGNZ method (Färe et al., 1994) measured four distance functions without taking the environmental effects, statistical error and slack adjustment into consideration and thus the results could be biased. To overcome these shortcomings, we follow our four-stage DEA approach by proposing a four-stage method, which incorporates environmental factors, statistical noise and slacks into the MPI and MSI measurements. The empirical results reveal that the changes in MPI and MSI by our proposed four-stage method are somewhat less than those measured by the FGNZ method, indicating that the productivity growth or sales capability growth would be overstated if one were to ignore the effects of environmental factors, data noise and slacks.

In this study, passenger-train-kilometer and freight-train-kilometer are used as the two outputs, which implicitly assume that the average number of cars per train and average number of seats per car are the same in different companies and train sets. The reason for making this assumption is due to the detailed data not available. To measure the rail performance more in line with the reality, we might select seat-kilometer as passenger service output and car-kilometers as freight service output in the future research, provided that those data are available. In the present paper, we have ignored the effects of congestion and assumed strong disposability for inputs and outputs; namely, a firm can always freely dispose unwanted inputs and outputs. In reality, the excess of some inputs may not be fully controlled by the operators (e.g., laying-off the extra employees may be protected by the labor union) and some undesirable outputs such as air pollution, noise and accidents are often inevitable. The input congestion may occur in railway transport whenever increasing some inputs will decrease some outputs without improving other inputs or outputs, or conversely, whenever decreasing some inputs will increase some outputs without worsening other inputs or outputs (Cooper et al., 2001). It is of interest to measure the efficiency and effectiveness when congestion is present. Therefore, one possible avenue of future research is to measure the rail performance by further considering the effects of input congestion (such as labors) and output congestion (such as accidents).

ACKNOWLEDGEMENTS

The authors wish to thank three referees for their positive comments and very constructive suggestions to improve this paper.

REFERENCES

Aigner, D.J., Lovell, C.A.K. and Schmidt, P. (1977) Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, **6**, 21-37.

Atkinson, S.E. and Cornwell, C. (1998) Estimating radial measures of productivity growth: frontier vs non-frontier approaches. *Journal of Productivity Analysis*, **10**, 35-46.

Banker, R.D., Charnes, A. and Cooper, W.W. (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, **30**, 1078-1092.

Banker, R.D. and Morey, R.C. (1986a) Efficiency analysis for exogenously fixed inputs and outputs. *Operations Research*, **34**, 513-521.

Banker, R.D. and Morey, R.C. (1986b) The use of categorical variables in data envelopment analysis. *Management Science*, **32**, 1613-1627.

Brooke, A., Kendrick, D., Meeraus, A. and Raman, R. (1998) GAMS A User's Guide. GAMS Development Corporation, Washington DC.

Cantos, P. and Maudos, J. (2000) Efficiency, technical change and productivity in the European rail sector: a stochastic frontier approach. *International Journal of Transport Economics*, **27**, 55-76.

Cantos, P. and Maudos, J. (2001) Regulation and efficiency: the case of European railways. *Transportation Research*, **35A**, 459-472.

Caves, D.W., Christensen, L.R. and Swanson, J.A. (1981) Productivity growth, scale economies and capacity utilization in U.S. railroads, 1955-74. *The American Economic Review*, **71**, 994-1002.

Caves, D.W., Christensen, L.R. and Diewert, W.E. (1982) The economic theory of index numbers and the measurement of input, output and productivity. *Econometrica*, **50**, 1393-1414.

Chapin, A. and Schmidt, S. (1999) Do mergers improve efficiency? *Journal of Transport Economics and Policy*, **33**, 147-162.

Charnes, A., Cooper, W.W. and Rhodes, E. (1978) Measuring the efficiency of decision making units. *European Journal of Operational Research*, **2**, 429-444.

Charnes, A., Haag, S., Jaska, P. and Semple, J. (1992) Sensitivity of efficiency classification in the additive model of data envelopment analysis. *International Journal System Science*, **23**, 789-798.

Charnes, A. and Neralic, L. (1990) Sensitivity analysis of the additive model in data envelopment analysis. *European Journal of Operational Research*, **48**, 332-341.

Coelli, T. (1996a) A Guide to DEAP Version 2.1: A Data Envelopment Analysis Computer Program. CEPA Working Paper 96/08, Department of Econometrics, University of New England, Armidale.

Coelli, T. (1996b) A Guide to FRONTIER Version 4.1: A Computer Program for Frontier Production Function Estimation. CEPA Working Paper 96/07, Department of Econometrics, University of New England, Armidale.

Coelli, T. (1998) A multi-stage methodology for the solution of orientated DEA models. *Operations Research Letters*, **23**, 143-149.

Coelli, T.J. and Perelman, S. (2000) Technical efficiency of European railways: a distance function approach. *Applied Economics*, **32**, 1967-1976.

Coelli, T.J., Rao, D.S.P. and Battese, G. (1998) *An Introduction to Efficiency and Productivity Analysis*. Kluwer Academic Publishers, Boston, MA.

Cooper, W.W., Gu, B. and Li, S. (2001) Comparisons and evaluations of alternative approaches to the treatment of congestion in DEA. *European Journal of Operational Research*, **132**, 62-74.

Cowie, J. (1999) The technical efficiency of public and private ownership in the rail industry. *Journal of Transport Economics and Policy*, **33**, 241-252.

Färe, R., Grosskopf, S., Lindgren, B. and Roos, P. (1989) Productivity developments in Swedish hospitals: a Malmquist output index approach. In A. Charnes, W.W. Cooper, A. Lewin and L. Seiford (eds.), *Data Envelopment Analysis: Theory, Methodology and Applications*, Kluwer Academic Publishers, Boston.

Färe, R., Grosskopf, S. and Norris, M. (1997) Productivity growth, technical progress and efficiency change in industrialized countries: reply. *American Economic Review*, **87**, 1040-1043.

Färe, R., Grosskopf, S., Norris, M. and Zhang, Z. (1994) Productivity growth, technical progress and efficiency change in industrialized countries. *American Economic Review*, **84**, 66-83.

Farrell, M.J. (1957) The measurement of productive efficiency. *Journal of the Royal Statistical Society Series A*, **120**, 253-290.

Fielding, G.J., Babitsky, T.T. and Brenner, M.E. (1985) Performance evaluation for bus transit. *Transportation Research*, **19A**, 73-82.

Fleming, C. (1999) Train drain: in the unified Europe shipping freight by rail is a journey into the past. The Wall Street Journal, March 29, A1-A8.

Freeman, K.D., Oum, T.H., Tretheway, M.W. and Waters II, W.G. (1985) The total factor productivity of the Canadian class I railways: 1956-1981. *The Logistics and Transportation Review*, **21**, 249-276.

Fried, H.O., Lovell, C.A.K. and Eeckaut, P.V. (1993) Evaluating the performance of US credit unions. *Journal of Banking and Finance*, **17**, 251-265.

Fried, H.O., Lovell, C.A.K., Schmidt, S.S. and Yaisawarng, S. (2002) Accounting for environmental effects and statistical noise in data envelopment analysis. *Journal of Productivity Analysis*, **17**, 157-174.

Fried, H.O., Schmidt, S.S. and Yaisawarng, S. (1999) Incorporating the operating environment into a nonparametric measure of technical efficiency. *Journal of Productivity Analysis*, **12**, 249-267.

Friedlaender, A.F., Berndt, E.R., Chiang, J.S.W., Showalter, M. and Vellturo, C.A. (1993) Rail cost and capital adjustments in a quasi-regulated environment. *Journal of Transport Economics and Policy*, **27**, 131-152.

Hays, W.L. (1973) *Statistics*. Holt, Rinehart and Winston, New York.

Hibiki, N. and Sueyoshi, T. (1999) DEA sensitivity analysis by changing a reference set: regional contribution to Japanese industrial development. *Omega International Journal of Management Science*, **27**, 139-153.

Kennedy, J. and Smith, A.S.J. (2004) Assessing the efficient cost of sustaining Britain's rail network. *Journal of Transport Economic and Policy*, **38**, 157-190.

Lan, L.W. and Lin, E.T.J. (2002) Measuring technical efficiency and scale efficiency in rail industry: a comparison of 85 railways using DEA and SFA. *Proceedings of Asia Conference on Efficiency and Productivity Growth*, 19 July, Academia Sinica, Taipei.

Lan, L.W. and Lin, E.T.J. (2003) Technical efficiency and service effectiveness for railways industry: DEA approaches. *Journal of the Eastern Asia Society for Transportation Studies*, **5**, 2932-2947.

Lan, L.W. and Lin, E.T.J. (2004) Two-stage measurements of railway transport performance with stochastic distance function approaches. *Proceedings of Croucher Advanced Study Institute (ASI) Workshop on Advanced Modeling for Transit Supply and Passenger Demand*, 19-20 January, The Hong Kong Polytechnic University.

Lewis, I., Semeijn, J. and Vellenga, D.B. (2001) Issues and initiatives surrounding rail freight transportation in Europe. *Transportation Journal*, **41**, 23-31.

Loizides, J. and Tsionas, E.G. (2004) Dynamic distribution of productivity growth in European railways. *Journal of Transport Economic and Policy*, **38**, 45-76.

Malmquist, S. (1953) Index numbers and indifference surface. *Trabajos de Estadistica*, **4**, 209-242.

McGeehan, H. (1993) Railway costs and productivity growth: the case of the Republic of Ireland, 1973-1983. *Journal of Transport Economics and Policy*, **27**, 19-32.

Oum, T.H., Waters II, W.G. and Yu, C. (1999) A survey of productivity and efficiency measurement in rail transport. *Journal of Transport Economics and Policy*, **33**, 9-42.

Oum, T.H. and Yu, C. (1994) Economic efficiency of railways and implications for public policy. *Journal of Transport Economic and Policy*, **28**, 121-138.

Seiford, L.M. and Zue, J. (1998a) Stability regions for maintaining efficiency in data envelopment analysis. *European Journal of the Operational Research*, **108**, 127-139.

Seiford, L.M. and Zue, J. (1998b) Sensitivity analysis of DEA models for simultaneous changes in all the data. *Journal of the Operational Research Society*, **49**, 1060-1071.

Sueyoshi, T. (1999) DEA non-parametric ranking test and index measurement: slack-adjusted DEA and an application to Japanese agriculture cooperatives. *Omega International Journal of Management Science*, **27**, 315-326.

Sueyoshi, T. and Goto, M. (2001) Slack-adjusted DEA for time series analysis: performance measurement of Japanese electric power generation industry in 1984-1993. *European Journal of Operational Research*, **133**, 232-259.

Sueyoshi, T., Ohnishi, K. and Kinase, Y. (1999) A benchmark approach for baseball evaluation. *European Journal of Operational Research*, **115**, 429-448.

Tretheway, M.W., Waters II, W.G. and Fok, A.K. (1997) The total factor productivity of the Canadian class I railways: 1956-91. *Journal of Transport Economics and Policy*, **31**, 93-113.

UIC (1995, 1996, 1997, 1998, 1999, 2000, 2001) *International Railway Statistics*. Paris.

Xie, R., Chen, H. and Nash, C. (2002) Migration of railway transport from command economy to market economy: the case of China. *Transport Reviews*, **22**, 159-177.

Zue, J. (1996) Robustness of the efficient DMUs in data envelopment analysis. *European Journal of Operational Research*, **90**, 451-460.

APPENDIX 1. NOTATIONS

$d_I^s(x_t, y_t)$	input distance from period *t* to period *s*
$d_C^s(y_t, z_t)$	consumption distance from period *t* to period *s*
g_O, g_{-O}	sensitivity indexes of the *O*th DMU in input-oriented DEA model
h_O, h_{-O}	sensitivity indexes of the *O*th DMU in consumption-oriented DEA model
m_I	input-based Malmquist productivity index (MPI)
m_C	consumption-based Malmquist sales capability index (MSI)
n_p	number of DMUs in group *p*
p	number of groups in Kruskal-Wallis rank test
x	inputs
x_{mj}^A, x_{mj}	adjusted and observed *m*th input quantities for the *j*th DMU, respectively
y	outputs
z	consumptions
I	number of environmental factors
J	number of DMUs
K	number of outputs
M	number of input factors
Q	number of consumptions
S_{mj}	sum of radial and non-radial slacks
T_p	sum of ranks for group *p*
α^*, β^*	the optimal values in the sensitivity test models (12) and (13)
δ	parameters of environmental factors
θ	technical efficiency scores estimated by the input-oriented DEA model
λ	constant
u_{mj}	managerial inefficiency
v_{mj}	statistical error (data noise)
$1/\phi$	service effectiveness scores estimated by the consumption-oriented DEA model
ω_{ij}	the *i*th environmental factor of DMU *j*
ΔE	effectiveness change
ΔI	catching-up effect (efficiency change)
ΔP	movement of frontier (technical change)
ΔS	sales innovation change

APPENDIX 2. THE AVERAGE EFFICIENCY AND EFFECTIVENESS SCORES MEASURED BY THREE DEA APPROACHES OVER 1995-2001

DMUs			Efficiency scores			Effectiveness scores		
No.	Country	Railway	BCC	3-stage	4-stage	BCC	3-stage	4-stage
1	Austria	*ÖBB*	0.876	0.953	0.821	0.199	0.876	0.875
2	Belgium	*SNCB/NMBS*	0.510	0.951	0.951	0.308	0.944	0.944
3	Denmark	*DSB*	0.863	0.963	0.940	0.416	0.957	0.953
4	Finland	*VR*	0.673	0.929	0.882	0.312	0.877	0.877
5	France	*SNCF*	0.694	0.941	0.920	0.268	0.460	0.458
6	Germany	*DB AG*	0.872	0.934	0.866	0.268	0.347	0.308
7	Greece	*CH*	0.589	0.895	0.708	0.691	0.972	0.972
8	Ireland	*CIE*	0.832	0.860	0.770	0.210	0.938	0.938
9	Italy	*FS SpA*	0.437	0.971	0.969	0.357	0.799	0.799
10	Luxembourg	*CFL*	0.693	0.953	0.875	0.428	0.971	0.970
11	Netherlands	*NS N.V.*	0.972	0.996	0.979	0.755	0.984	0.962
12	Portugal	*CP*	0.682	0.936	0.896	0.282	0.966	0.966
13	Spain	*RENFE*	0.664	0.974	0.966	0.250	0.840	0.840
14	Sweden	*SJ AB*	0.991	0.920	0.906	0.242	0.764	0.763
15	Norway	*NSB BA*	0.950	0.883	0.802	0.202	0.884	0.884
16	Switzerland	*BLS*	0.958	0.967	0.881	0.488	0.965	0.965
17	Switzerland	*CFF/SBB/FFS*	0.890	0.983	0.943	0.268	0.929	0.929
18	Bulgaria	*BDZ*	0.279	0.926	0.895	0.370	0.977	0.976
19	Croatia	*HZ*	0.444	0.918	0.826	0.215	0.970	0.970
20	Czech Rep	*CD*	0.409	0.926	0.905	0.233	0.881	0.881
21	Estonia	*EVR*	0.591	0.917	0.797	0.864	0.998	0.998
22	Hungary	*GYSEV/RÖEE*	0.951	0.806	0.671	0.710	0.988	0.988
23	Hungary	*MÁV Rt.*	0.426	0.941	0.864	0.288	0.923	0.923
24	Latvia	*LDZ*	0.506	0.911	0.883	0.835	0.994	0.994
25	Lithuania	*LG*	0.418	0.927	0.882	0.669	0.992	0.992
26	Poland	*PKP*	0.454	0.916	0.868	0.328	0.765	0.751
27	Romania	*CFR*	0.237	0.899	0.855	0.355	0.937	0.937
28	Slovak	*ZSR*	0.501	0.924	0.788	0.274	0.962	0.962
29	Slovenia	*SZ*	0.860	0.885	0.719	0.197	0.980	0.980
30	Moldova	*CFM (E)*	0.164	0.892	0.743	0.906	0.994	0.994
31	Ukraine	*UZ*	0.333	0.871	0.738	0.984	0.996	0.996
32	Turkey	*TCDD*	0.514	0.887	0.859	0.321	0.957	0.957
33	Israel	*IsR*	0.667	0.922	0.850	0.630	0.983	0.983
34	Morocco	*ONCFM*	0.400	0.953	0.907	0.684	0.990	0.990
35	Syria	*CFS*	0.337	0.942	0.898	0.427	0.969	0.955
36	Mozambique	*CFM*	0.793	0.875	0.706	0.614	0.972	0.973
37	Tanzania	*TRC*	0.997	0.817	0.666	0.672	0.981	0.973
38	Azerbaijan	*AZ*	0.160	0.875	0.698	0.545	0.993	0.993
39	Korea	*KNR*	0.888	0.979	0.962	0.554	0.984	0.979
40	Japan	*JR*	0.935	0.971	0.869	0.987	0.994	0.945
41	India	*IR*	0.334	0.929	0.671	0.938	0.989	0.894
42	Taiwan	*TRA*	0.959	0.998	0.998	0.519	0.992	0.988
43	Turkmenistan	*TRK*	0.427	0.945	0.918	0.908	0.994	0.994
44	Australia	*QR*	0.997	0.873	0.829	0.903	0.980	0.980
Mean			0.639	0.924	0.849	0.497	0.923	0.917

APPENDIX 3. THE AVERAGE PRODUCTIVITY AND SALES CAPABILITY GROWTHS MEASURED BY TWO METHODS OVER 1995-200

DMUs			Productivity growth (%)		Sales capability growth (%)	
No.	Country	Railway	FGNZ method	Four-stage method	FGNZ method	Four-stage method
1	Austria	*ÖBB*	117.9	106.1	95.3	99.4
2	Belgium	*SNCB/NMBS*	107.4	101.4	101.2	100.2
3	Denmark	*DSB*	110.4	107.1	118.9	94.2
4	Finland	*VR*	104.5	151.7	98.6	100.7
5	France	*SNCF*	117.7	104.0	124.9	114.2
6	Germany	*DB AG*	105.9	89.8	120.4	112.1
7	Greece	*CH*	104.8	96.6	118.0	93.5
8	Ireland	*CIE*	90.4	102.2	108.2	101.8
9	Italy	*FS SpA*	116.0	101.1	113.5	102.6
10	Luxembourg	*CFL*	94.6	98.7	126.2	87.8
11	Netherlands	*NS N.V.*	106.9	93.5	97.9	99.2
12	Portugal	*CP*	108.9	106.9	86.9	100.8
13	Spain	*RENFE*	116.6	112.1	123.4	102.0
14	Sweden	*SJ AB*	115.7	121.0	90.0	106.0
15	Norway	*NSB BA*	124.6	98.7	110.5	99.4
16	Switzerland	*BLS*	112.2	109.0	121.2	114.6
17	Switzerland	*CFF/SBB/FFS*	119.7	98.0	103.7	100.6
18	Bulgaria	*BDZ*	120.7	96.2	85.4	108.1
19	Croatia	*HZ*	111.0	96.2	85.9	100.6
20	Czech Rep	*CD*	117.2	96.6	105.8	102.2
21	Estonia	*EVR*	120.6	114.3	127.4	116.3
22	Hungary	*GYSEV/RÖEE*	104.5	107.3	119.9	101.9
23	Hungary	*MÁV Rt.*	110.5	102.3	123.2	101.1
24	Latvia	*LDZ*	114.1	103.4	135.3	128.1
25	Lithuania	*LG*	108.9	102.5	125.1	129.2
26	Poland	*PKP*	97.3	96.3	93.6	107.7
27	Romania	*CFR*	94.8	100.7	97.6	98.6
28	Slovak	*ZSR*	123.3	87.4	94.5	91.2
29	Slovenia	*SZ*	111.9	89.9	99.8	103.3
30	Moldova	*CFM (E)*	98.0	81.8	111.7	83.8
31	Ukraine	*UZ*	115.4	93.1	107.5	86.3
32	Turkey	*TCDD*	99.9	72.2	105.8	100.2
33	Israel	*IsR*	128.8	111.3	101.1	107.0
34	Morocco	*ONCFM*	113.1	96.1	112.2	100.1
35	Syria	*CFS*	107.7	118.3	125.5	115.1
36	Mozambique	*CFM*	130.2	120.2	83.9	109.2
37	Tanzania	*TRC*	130.3	126.9	71.6	88.7
38	Azerbaijan	*AZ*	113.0	100.6	114.2	109.0
39	Korea	*KNR*	129.4	99.5	98.1	95.4
40	Japan	*JR*	121.7	95.3	105.4	128.0
41	India	*IR*	116.2	111.9	116.5	80.6
42	Taiwan	*TRA*	109.2	88.8	112.5	90.3
43	Turkmenistan	*TRK*	85.7	102.5	100.5	98.9
44	Australia	*QR*	116.4	130.1	108.3	119.5
Mean			120.2	106.6	107.3	106.1