MEASURING FIRMS' INPUT CONGESTION WITH CONSIDERATION OF ENVIRONMENTAL FACTORS: THE CASE OF EUROPEAN RAILWAY TRANSPORT

Erwin LIN

In practice, input congestion effects appear in railway transport due to the difficulties of disposing of unnecessary input factors. This study measures the output-oriented technical efficiency and input congestion with consideration of categorical variables for railway transport by using the DEA extension approach. The empirical results from 24 European railway companies show that in 12 railways, the presence of weak congestion can be proved. Based on the results of identifying the source(s) of input congestion and further determining its amount, one can obtain more insights into railways' operation and thus propose more effective strategies for improvement.

Keywords: Data Envelopment Analysis, input congestion, weak congestion, rail transport, categorical variables

JEL classification indices: C60, D04, D24, L25

1. INTRODUCTION

Rail transport has long played an important role in the economic development of a country and thus the enhancement of its operating efficiency is important to be sustainable in a competitive context. Many researchers have sought to measure rail transport performance during the past few decades, but one common drawback is that they did not take the input congestion effects into consideration. In

Erwin *Lin*, Associate Professor, Department of Marketing and Logistics, MingDao University, Peetou Township, Changhua County, Taiwan. E-mail: erwintjlin@mdu.edu.tw

practice, the "input congestion effects" are apparent in the railway transport due to the difficulties of disposing unnecessary input factors (e.g. overstaffing). Thus, it would be helpful if one could identify the possible sources of input congestion in order to propose more specific improvement strategies for any specific railway company. As such, evaluating input congestion effects for rail transport deserves an in-depth investigation.

In this study, we attempt to evaluate the output-oriented technical efficiency and then investigate the strong and weak congestion for 24 European rail companies in 2008. It is hypothesized that there may exist strong and/or weak input congestions in some railways and the main objective of this study is to test this hypothesis.

Moreover, based on the previous literature, we note that some methods for measuring input congestion have been developed, but these methods do not take into account environmental factors. In practice, due to the heterogeneity of operating environments in different countries, it is more reasonable to investigate technical efficiency as well as input congestion by dividing the samples into several categories based on the operation environment faced by each company. This study thus investigates strong and weak congestions by dividing the samples into two groups: Eastern Europe and Western Europe. More specifically, the methodology for measuring input congestion in this study is a novel approach, which is modified by imposing a categorical constraint on data envelopment analysis (DEA) models used in a three-stage approach. At the first stage, we evaluate the output-oriented technical efficiency by using the DEA model. The strong congestion is detected at the second stage. At the third stage, we then investigate weak congestion by adopting the method proposed by Tone – Sahoo (2004).

It should be noted that all of the three-stage models used have been modified by adding a categorical constraint to the models. Our results indicate that measuring input congestion with consideration of a dummy variable, which is used to classify the sample into two categories, seems more reasonable and reliable than those that ignore environmental effects. The major contribution of this study is that we identify the sources of input congestion in the railway transport performance measurement by extending the DEA method and incorporating the environmental effects into the analysis, from which one can obtain more insights into railways' operation and thus propose more effective strategies for improvement.

The remaining part of this article is organised as follows. Section 2 reviews previous studies, while Section 3 describes the methodology used in this study. The results of the empirical analysis for some selected railways are presented in Section 4, and finally some conclusions and recommendations drawn from the empirical study are discussed in Section 5.

2. LITERATURE REVIEW

The DEA method, first developed by Charnes et al. (1978), has been widely used in performance evaluation for both public and private sectors (e.g. Knežević et al. 2015; Kočišová 2015). Many studies, including theoretical developments and practical applications, have been published in literature in the past three decades. Since the DEA method has some advantages (e.g., one can handle multi-output, multi-input production technologies without the need of specifying a functional form in prior), many researchers measure the technical efficiency of railways by different DEA models. For example, Bookbinder – Qu (1993) estimated the performance of two Canadian and five US Class I railways by using DEA. Oum – Yu (1994) applied DEA to evaluate the efficiency of 19 OECD countries' rail companies over the period of 1978 to 1989. Chapin – Schmidt (1999) used the DEA to measure the efficiency of US Class I railroad companies since deregulation. Cowie (1999) also applied DEA to compare the efficiency of Swiss public and private railways by constructing technical and managerial efficiency frontiers, and then measured both efficiencies. Lan – Lin (2005) further proposed a four-stage DEA model to evaluate rail transport efficiency, effectiveness, productivity, and marketing capability. More recently, due to the unstorability of transport services, Yu (2008) assessed technical efficiency, service effectiveness, and technical effectiveness for the world's railways by using the Network DEA method. The aforementioned works have sought to measure the performance of rail transport in the past; however, none has taken the input congestion effect into account.

Conventional DEA models implicitly assume that each firm can always freely dispose overused input factors. The treatment of the congestion of input factors in the technical efficiency measurement context was probably pioneered by Färe – Svensson (1980), who referred to Turgot's "law of variable proportions" and distinguished three different strengths of congestion of production factors, including output-limitational congested, monotone output-limitational congested, and output-prohibitive congested. Färe et al. (1985) subsequently proposed an operational input-oriented weak disposal DEA model for dealing with input congestion measurement. Using a slack-based DEA approach proposed by Cooper et al. (1996), Brockett et al. (1998) estimated the inefficiency and amounts of input congestion in the Chinese empirical production before and after the 1978 economy reforms. Their results indicated that the textiles, chemicals, and metallurgy industries all exhibited input slacks due to the huge labour employment problem and all slacks of labour input in each industry were attributed to congestion. Cooper et al. (2001) also adopted same approach

to identify congestive inputs in Chinese textiles and automobiles industries. Their results showed that the elimination of managerial inefficiencies could have led to output augmentation without reducing employment, and even in the heavily congested textile industry, these output augmentations could have been accompanied by reductions in the amounts of capital used. Viton (1997) applied the DEA approach proposed by Färe et al. (1985) and used US data to investigate the input congestion for bus transit; the results indicated no sign of the excess usage of inputs. Odeck (2006) also applied the method of Färe et al. (1985) to evaluate the operators' efficiency in the Norwegian bus industry; the results indicated that 92% of inefficiency is due to a formidable degree of congestion. Kao (2010) measured the input congestion effect of Taiwan forests within the framework of DEA and concluded that the theoretical method of deleting excessive inputs to improve efficiency is actually impractical in this case. Tone – Sahoo (2004) developed a new scheme to evaluate scale elasticity in the presence of input congestion and then applied their innovative method to a data set obtained from the Japan Chain Store Association (JCSA). Their results indicated that the JCSA had been operating under decreasing returns to scale (DRS) for the last ten years. These findings are useful for managers to reform their business.

Based on the previous literature, we note that at least three methods for measuring input congestion have been developed: the Färe-Grosskopf-Lovell (FGL) method by Färe et al. (1985), the Cooper-Thompson-Thrall (CTT) method by Cooper et al. (1996), and the Tone-Sahoo (TS) method (2004). These three methods differ fundamentally: the FGL method used a radial Farrell (1957) measure, while the CTT and TS methods used a slack-based measure to identify the input(s) responsible for congestion. The key difference between the CTT and TS methods lies in the point of view: the CTT method measures congestion in terms of excessive input(s), while the TS method in terms of output shortfall. In addition, the TS method measures scale elasticity before identifying congestion, while the CTT method does not. By definition, input congestion occurs whenever increasing one or more inputs will decrease one or more outputs without improving other inputs or outputs. In other words, it is meaningful and useful to identify the presence of input congestion by checking whether its scale elasticity is negative. In this sense, measuring congestion from the output point of view and calculating scale elasticity in advance seems more logical and reasonable. We thus adopt the TS method in this study, which will be described in detail in the following section.

3. METHODOLOGY

3.1. Measuring technical efficiency and slacks

The production technology can be represented by the production possibility set containing all feasible input and output vectors: $P = \{(x,y) \mid x \text{ can produce } y\}$. That is, one can define the production possibility set as $P = \{(x,y) \mid (x \geq X\lambda, y \leq Y\lambda, \lambda \geq 0) \in T \}$, where $x = (x_1, x_2,\ldots, x_m) \in R_+^m$, and $y = (y_1, y_2,\ldots, y_k) \in R_+^k$. P is closed, bounded, convex, and satisfies strong disposability. Once the   production possibility set is defined, the efficiency can be measured by the distance from the observed data point to the best feasible practice (frontier), which can be solved by the linear programming (LP) technique.

Assume that there are J decision making units (DMUs) to be evaluated, $J = \{1,\ldots,J\}$, each DMU produces K outputs, $K= \{1,\ldots, K\}$, by utilizing M inputs, $M = \{1,\ldots, M\}$. Charnes et al. (1978) proposed to measure the technical efficiency for DMU_i underestimated by using the mathematical program known as the CCR model in an output-oriented form. The CCR model implicitly assumes that all DMUs are operated at their optimal scale, that is, under the assumption of constant returns to scale (CRS). However, it is believed that this may not be the case in practice. To relax this assumption, Banker et al. (1984) proposed the following model to account for variable returns to scale (VRS), known as the BCC model.

$$Max_{\varphi,\lambda} \quad \varphi_i + \varepsilon (\sum_{k=1}^{K} s_k^+ + \sum_{m=1}^{M} s_m^-)$$

$$s.t. \quad -\varphi_i \cdot y_{ik} + \sum_{j=1}^{J} \lambda_j y_{jk} - s_k^+ = 0, \quad k = 1,\ldots,K,$$

$$x_{im} - \sum_{j=1}^{J} \lambda_j x_{jm} - s_m^- = 0, \quad m = 1,\ldots,M, \tag{1}$$

$$\lambda_j \geq 0,$$

where x and y are input and output variables, respectively, λ is weight, φ_i is the maximal proportional expansion in all outputs, $1/\varphi_i$ is the efficiency of DMU_i to be evaluated, and s^+ and s^- are the output and input slacks, respectively. Solving (1) for each of J firms, one obtains the efficiency score for each firm. To measure s^+ and s^-, Ali – Seiford (1993) suggested employing a two-stage method: obtain φ_i by solving (1) and then solve the following LP, keeping the objective function value at optimal φ_i.

$$Max_{s_k^+, s_m^-} \quad (\sum_{k=1}^{K} s_k^+ + \sum_{m=1}^{M} s_m^-)$$

$$s.t. \quad \varphi_i \cdot y_{ik} = \sum_{j=1}^{J} \lambda_j y_{jk} - s_k^+, \quad k=1,...,K, \tag{2}$$

$$x_{im} = \sum_{j=1}^{J} \lambda_j x_{jm} + s_m^-, \quad m=1,...,M,$$

$$\lambda_j \geq 0,$$

where the notations are the same as in (1) and φ_i is the optimal solution computed from (1). The aforementioned models do not take into account environmental factors; in practice, the environment faced by each DMU would never be the same. In other words, it is almost inevitable that the operating environments of particular units will be non-homogeneous. For example, when measuring the relative efficiency for Starbucks coffee shops, one could always argue that homogeneity due to the different site each coffee shop is located in. It is unfair to compare the performance of an unfavourable DMU to a favourable one. To rectify, many researchers thus proposed some modified models to account for the effects of environmental factors on efficiency measurement. By referring to previous works (for instance, among others, Banker – Morey 1986a, 1986b; Ruggiero 1996, 1998; Førsund 2001), we propose to use the models as follows.

$$Max_{\varphi, \lambda} \quad \varphi_i + \varepsilon (\sum_{k=1}^{K} s_k^+ + \sum_{m=1}^{M} s_m^-)$$

$$s.t. \quad -\varphi_i \cdot y_{ik} + \sum_{j=1}^{J} \lambda_j y_{jk} - s_k^+ = 0, \quad k=1,...,K,$$

$$x_{im} - \sum_{j=1}^{J} \lambda_j x_{jm} - s_m^- = 0, \quad m=1,...,M,$$

$$\sum_{j=1}^{J} \lambda_j = 1, \tag{3}$$

$$\sum_{j=1}^{J} \lambda_j = 0, \text{ if } z_{jr} > z_{ir}, \quad r=1,...R,$$

$$\lambda \geq 0, j=1,...,J$$

where z is the binary variable, in which 1 stands for the DMU with more favoura-
ble environment. The constraint explicitly restricts the comparison set to exclude
DMUs that face a more favourable environment in measuring efficiency for those
DMUs with an unfavourable environment. After φ_i has been measured, one uses
model (4) to compute the input and output slacks.

$$Max_{s_k^+, s_m^-} \quad (\sum_{k=1}^{K} s_k^+ + \sum_{m=1}^{M} s_m^-)$$

$$s.t. \quad \varphi_i \cdot y_{ik} = \sum_{j=1}^{J} \lambda_j y_{jk} - s_k^+, \quad k = 1,...,K,$$

$$x_{im} = \sum_{j=1}^{J} \lambda_j x_{jm} + s_m^-, \quad m = 1,...,M, \qquad (4)$$

$$\lambda_j \geq 0,$$

$$\sum_{j=1}^{J} \lambda_j = 0, \text{ if } z_{jr} > z_{ir,} \quad r = 1,...R,$$

Again, the notations in (4) are the same as in (3). By following Tone – Sahoo
(2004), once φ_i, s^+ and s^- have been computed, four conditions can be used to
identify congestion: (1) if $\varphi^* = 1$, $s^{-*} = 0$ and $s^{+*} = 0$, then the DMU is efficient
and not congested. (2) if $\varphi^* = 1$, $s^{-*} \neq 0$ and $s^{+*} = 0$, then the DMU is technically
inefficient, but not congested because there is no output shortfall. (3) if $\varphi^* = 1$ and
$s^{+*} \neq 0$, or (4) if $\varphi^* > 1$, then the DMU exhibits congestion. In this study, we will
use these four conditions to identify the existence of input congestion. For more
detail, refer to Tone – Sahoo (2004).

3.2. Measuring scale elasticity

Following Banker et al. (1984), the dual form to (1) is expressed as follows:

$$Min \quad \sum_{m=1}^{M} v_m x_{om} - \omega_o$$

$$s.t. \quad -\sum_{m=1}^{M} v_m x_{jm} + \sum_{k=1}^{K} u_k y_{jk} + \omega_j \leq 0, j = 1,...,J, \qquad (5)$$

$$\sum_{k=1}^{K} u_k y_{ok} = 1, \quad v \geq 0, \; u \geq 0, \; \omega_o : free \; in \; sign$$

Banker et al. (1984) showed that the optimal solution of ω^* can be used to derive the scale economies. In many instances, however, the optimal ω^* is not uniquely determined; in this case, Tone – Sahoo (2004) proposed to find the upper bound ($\overline{\omega}$) and lower bound ($\underline{\omega}$) by solving the following LP.

$$\overline{\omega}\ (\underline{\omega}) = Max\ (Min)\ \omega_o$$

$$s.t.\ \ -\sum_{m=1}^{M} v_m x_{jm} + \sum_{k=1}^{K} u_k y_{jk} + \omega_j \leq 0, j = 1,...,J,$$

$$-\sum_{m=1}^{M} v_m x_{om} + \sum_{k=1}^{K} u_k y_{ok} + \omega_o = 0, \tag{6}$$

$$\sum_{k=1}^{K} u_k y_{ok} = 1,$$

$$v \geq 0,\ u \geq 0,\ \omega_o : free\ in\ sign$$

Once the upper (lower) bound of ω has been calculated, the upper (lower) scale elasticity in production $\overline{\rho}$ ($\underline{\rho}$) can be determined by

$$\overline{\rho} = 1 + \overline{\omega}\ \ and\ \ \underline{\rho} = 1 + \underline{\omega}. \tag{7}$$

3.3. Measuring strong congestion

If a DMU is identified as congestive by using the conditions mentioned above, we may need to further determine the source(s) of congestion. First, by following Tone – Sahoo (2004), we define strong congestion as "an increase (decrease) in all inputs will lead to a decrease (increase) in all outputs." Second, the efficiency under the strong disposability assumption can be measured by the following linear programming:

$$Max_{\theta,\lambda}\theta_i$$

$$s.t.\ \ -\theta_i \cdot y_{ik} + \sum_{j=1}^{J} \lambda_j y_{jk} - q_k^+ = 0,\ \ k = 1,...,K,$$

$$x_{im} - \sum_{j=1}^{J} \lambda_j x_{jm} = 0,\ \ m = 1,...,M, \tag{8}$$

$$\sum_{j=1}^{J} \lambda_j = 0,\ if\ z_{jr} > z_{ir,}\ \ r = 1,...R,$$

$$\lambda_j \geq 0,$$

where q_k^+ is output shortfall. Note that the difference between eq. (8) and eq. (3) is that the inequality is replaced by equality in the second constraint. To find out q_k^+, one has to use a two-stage method in the same manner as in eq. (2).

Similar to Banker et al. (1984), the dual problem to eq. (6) can be formulated as follows:

$$Min \quad \sum_{m=1}^{M} v_m x_{om} - w_o \tag{9}$$

$$s.t. \quad -\sum_{m=1}^{M} v_m x_{jm} + \sum_{k=1}^{K} u_k y_{jk} + w_j \leq 0, j = 1,...,J,$$

$$\sum_{k=1}^{K} u_k y_{ok} = 1,$$

$$u \geq 0, \quad v, \quad w_o : free \ in \ sign.$$

Note that in eq. (9) v is free in sign, which allows the frontier to bend downward. In other words, the frontier can have negative slopes. The multiple optima in eq. (9) can also be dealt with by the following:

$$\overline{w} = Max \ w_o$$

$$s.t. \quad -\sum_{m=1}^{M} v_m x_{jm} + \sum_{k=1}^{K} u_k y_{jk} + w_j \leq 0, j = 1,...,J,$$

$$-\sum_{m=1}^{M} v_m x_{om} + \sum_{k=1}^{K} u_k y_{ok} + w_o = 0, \tag{10}$$

$$\sum_{k=1}^{K} u_k y_{ok} = 1, \ u \geq 0, \quad v, \quad w_o : free \ in \ sign.$$

The upper bound of scale elasticity can be calculated by $\overline{\rho} = 1 + \overline{w}$. Following TS (2004), in this paper we define strong congestion as "a DMU is identified as strong congestion if and only if its upper bound of scale elasticity is negative."

3.4. Measuring weak congestion

Since the strong congestion defined in the previous subsection is too restricted, Tone – Sahoo (2004) thus relax the restriction to deal with the weak congestion, which is defined as "weak congestion occurs whenever an increase (decrease) in one or more inputs will lead to a decrease (increase) in one or more outputs."

Notice, that strong congestion implies weak congestion, but not *vice versa*. To investigate weak congestion, we calculate t_k^+ and t_m^- by a two-stage method; that is, solve the following model (11) and then model (12).

$$Max \quad \frac{1}{K} \sum_{k=1}^{K} \frac{t_k^+}{y_{ko}}$$

$$s.t. \ x_{mo} = \sum_{j=1}^{n} \lambda_j x_{ij} + t_m^-, (m=1,...,M)$$

$$y_{ko} = \sum_{j=1}^{n} \lambda_j y_{kj} - t_k^+, (k=1,2,...,K) \tag{11}$$

$$\sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0, \ t_m^-, \ t_k^+ \geq 0.$$

$$\sum_{j=1}^{J} \lambda_j = 0, \ if \ z_{jr} > z_{ir,} \quad r=1,...R,$$

$$Max \quad \frac{1}{K} \cdot \sum_{k=1}^{K} \frac{t_k^+}{y_{ko}} + \frac{1}{M} \cdot \sum_{m=1}^{M} \frac{t_m^-}{x_{mo}}$$

$$s.t. \ x_{mo} = \sum_{j=1}^{n} \lambda_j x_{mj} + t_m^-, (m=1,...,M)$$

$$y_{ko} = \sum_{j=1}^{n} \lambda_j y_{kj} - t_k^+, (k=1,2,...,K) \tag{12}$$

$$\sum_{j=1}^{n} \lambda_j = 1, \ \lambda_j \geq 0, \ t_m^-, \ t_k^+ \geq 0,$$

$$\sum_{j=1}^{J} \lambda_j = 0, \ if \ z_{jr} > z_{ir,} \quad r=1,...R.$$

To solve model (11) and (12), we use a two-stage process, more specifically, we solve (11) first, and then maximize $\frac{1}{M} \cdot \sum_{m=1}^{M} \frac{t_m^-}{x_{mo}}$ by model (12), while keeping $\frac{1}{K} \cdot \sum_{k=1}^{K} \frac{t_k^+}{y_{ko}}$ at the optimum value obtained from (11). It should be noted that the major difference between our proposed models (11), (12) and TS models (model 79 and 80 in Tone – Sahoo 2004) is that we impose the last constraint on the models (11) and (12), which restrict the reference set only to those DMUs with the same and unfavourable environment. Once t_k^+ and t_k^- have been calculated, we have the following two cases: (1) if $t_k^{+*} = 0$, no weak congestion is detected since a decrease in inputs cannot increase any outputs, (2) if $t_k^{+*} \neq 0$ and $t_k^{-*} \neq 0$, then DMU

is identified as with weak congestion. Furthermore, after data projection, we can apply the same procedure as in (6) and (7) for the projected DMU, and then decide the nature of degree of economies (DSE) after calculation via equation (7).

4. EMPIRICAL ANALYSIS

4.1. The data

Our data set is drawn from International Railway Statistics, published by the International Union of Railways (UIC 2008). Note, that when applying DEA to evaluate performance, the DMUs under consideration should be homogeneous (Golany – Roll 1989). In this study, we attempt to analyse the efficiency and identify possible source(s) of input congestion of railways; we thus selected 24 European railway companies. All of these railways provide both passenger and freight services. Railways providing only passenger or freight transport service such as SJ (Sweden), NS (Netherlands), DSB (Denmark), and NSB (Norway) are excluded from our data set.

We choose passenger-kilometre per unit length of lines and ton-kilometre per unit length of lines as output variables. As for the input variables, some studies selected number of employees, number of cars, and length of lines as inputs. However, we do not use length of lines because for rail transport industry, line-related facilities such as tracks, signals, stations, and yards can be viewed as sunk cost, which are generally attributed to "fixed" costs. In this paper, we attempt to measure the efficiency of "variable" input factors so as to identify the input congestion. As for the input of rolling stock, because the passenger (freight) service cannot be produced by freight (passenger) cars, therefore, instead of the sum of cars generally used in previous studies, we select the available number of passenger cars per unit length of lines and the available number of freight cars per unit length of lines (both are in terms of cars/km) as the two input factors. In addition, we also choose the number of employees per kilometre of lines as an input variable. *Table 1* presents the descriptive statistics for the data, including three inputs, two outputs, and one scaling or normalizing factor (that is, length of lines).

4.2. The categorical variable

From *Table 1,* one can easily see that the data are rather heterogeneous. Take Lines as an example: the data ranges from 275 (CFL, Luxemburg) to 33,855 kilometres (DB, Germany). To make a meaningful comparison, we thus divide

Table 1. Descriptive statistics of the data

Statistics	Pax-km	Ton-km	Labors	Pcars	Fcars	Lines	GNI
Max	86,664	91,178	240,008	18,671	119,916	33,855	58,810
Min	273	103	1,972	187	168	275	11,090
Mean	13,722	14,646	40,049	3,596	20,258	7,862	27,830
Stdev.	23,419	19,196	57,998	4,890	27,329	9,027	12,064

Note: GNI is in the unit of million, and dollar, respectively, Pcars, Fcars are the number of cars used per kilometre of lines.

each output and input by length of line. In other words, we select the passenger-km per unit length of line, and ton-km per kilometre of lines as two output variables, and choose number of employees per kilometre, number of passenger cars per kilometre, and number of freight cars per kilometre as three input variables. Furthermore, *Table 1* also indicates that the environmental factor, GNI per capita, varies from country to country, the data ranges from USD 11,090 to 58,810, and the standard deviation is USD 12,064. It reveals that the environments faced by different railways are quite varied. Based on the underlying theory of transportation and some previous studies (see, for example, Lan – Lin 2005, 2006), we are aware that GNI is the factor which significantly affects transport service production. In addition, the average GNI of Western Europe and Eastern Europe in the sample are USD 37,334 and 18,866, respectively, suggesting that the operating environments between these two regions are different. We thus introduce a category variable, z, into the models; in which 0 stands for railways operating in Eastern Europe and 1 for those in Western Europe.

4.3. The efficiency and output slack measures

Firstly, we measure output-oriented technical efficiency for the 24 sampled railways by applying model (1): the results are presented in *Table 2*. To investigate whether the efficiencies of Western and Eastern European countries are different, we thus perform a hypothesis testing by using the t-test; the result of the p-value (0.01) reveals that the null hypothesis of no difference between the two regions cannot be accepted. In other words, classifying the samples into two categories will be more reasonable and reliable. We then re-estimate the efficiency of each railway by using model (3); the results are also displayed in *Table 2*. It should be noted that the categorical variable z is imposed by 1 (0) for the railways operating in Western (Eastern) European countries. As can be seen from *Table 2,* the average technical efficiencies of the samples are 1.404 and 1.503, with and without classifying the samples into two categories, respectively. As expected, the aver-

age efficiency of railways in Eastern European countries is 1/1.038 when imposing the categorical variable on the model, which is more efficient than those counterpart of without imposition (1/1.255). As for the output slack measures, nine railways (ÖBB, SNCB, VR, CP, BLS, CFF, HZ, CD, and MÁV Rt) are evaluated as with output slacks in freight service, while six railways (SNCF, DB AG, FS SpA, RENFE, PKP, and SZ) are measured as with output slacks in passenger

Table 2. Output-oriented efficiency and output slack measurements

Country	Railways	φ_{BCC}	s_1^+	s_2^+	φ_{CV}	$s_{1,cv}^+$	$s_{2,cv}^+$
Austria	*ÖBB*	2.106	0	2.046	2.106	0	2.046
Belgium	*SNCB*	1.889	0	3.249	1.889	0	3.249
Finland	*VR*	1.422	0	176.859	1.422	0	176.859
France	*SNCF*	2.566	0.106	0	2.566	0.106	0
Germany	*DB AG*	2.010	0.131	0	2.010	0.131	0
Greece	*OSE*	1.000	0	0	1.000	0	0
Ireland	*CIE*	1.000	0	0	1.000	0	0
Italy	*FS SpA*	2.408	0.702	0	2.408	0.702	0
Luxemburg	*CFL*	1.000	0	0	1.000	0	0
Portugal	*CP*	1.988	0	1.764	1.988	0	1.764
Spain	*RENFE*	1.827	2.089	0	1.827	2.089	0
Switzerland	*BLS*	1.356	0	8.227	1.356	0	8.227
Switzerland	*CFF*	1.703	0	5.153	1.703	0	5.153
Bulgaria	*BDZ*	1.000	0	0	1.000	0	0
Croatia	*HZ*	1.018	0	4.380	1.000	0	0
Czech Rep.	*CD*	1.837	0	0.082	1.172	0	3.271
Estonia	*EVR*	1.000	0	0	1.000	0	0
Hungary	*MÁV Rt.*	1.731	0	1.148	1.016	0	79.360
Latvia	*LDZ*	1.000	0	0	1.000	0	0
Lithuania	*LG*	1.000	0	0	1.000	0	0
Poland	*PKP*	1.859	1.257	0	1.233	0	1.732
Romania	*CFR*	1.000	0	0	1.000	0	0
Slovenia	*SZ*	1.363	6.530	0	1.000	0	0
Turkey	*TCDD*	1.000	0	0	1.000	0	0
Average		1.503			1.404		

Note: φ_{BCC} and φ_{CV} denote the efficiency measures by BCC models without and with imposing categorical variable; s_1^{+*} and s_2^{+*} stand for output slacks of passenger and freight services, respectively. $s_{1,cv}^+$ and $s_{2,cv}^+$ represent the counterparts of slacks of imposing categorical variable.

service based on the assumption of the BCC model. On the other hand, eight rail-ways (ÖBB, SNCB, VR, CP, BLS, CFF, CD, and MÁV Rt) have output slacks in freight, and four railways (SNCF, DB AG, FS SpA, and RENFE) are evaluated as with output slacks of passenger service when classifying the sample into two categories. As we mentioned earlier, the operating environment will be more ho-mogeneous after dividing the samples into two groups based on their operating environment, thus the result seems more reasonable and reliable. It is noteworthy that such relative larger-scale railways as SNCF (France), DB AG (Germany), FS (Italy), and RENFE (Spain) are all evaluated as exhibiting output slacks in terms of passenger and/or freight services, no matter what model is used. This is one of the evidences indicating that larger railways should downsize in order to eliminate slacks.

To investigate the scale elasticity in production for each railway, we then meas-ure the upper bound and lower bound of ρ by means of model (6) and equation (7), and the results indicate that all railways are evaluated as with positive scale elasticity, indicating that there is no strong congestion in the sampling railways. As for the weak congestion, by definition, the weak congestion can be detected as being present in those DMUs with ($\varphi^* = 1$, and $S^{+*} \neq 0$) or $\varphi^* > 1$. As can be seen from *Table 2*, thirteen railways, ÖBB (Austria), SNCB (Belgium), VR (Fin-land), SNCF (France), DB (Germany), FS (Italy), CP (Portugal), RENFE (Spain), BLS (Switzerland), CFF (Switzerland), CD (Czech), MÁV Rt (Hungary), and PKP (Poland) are evaluated as exhibiting input congestion as their efficiency scores are all greater than unity. Thus, it deserves further investigation.

In addition to the BCC DEA model, we also measure the efficiency scores and output shortfall by using model (8) that is the technology based on the assump-tion of strong congestion and with categorical constraint. More specifically, we compute the efficiency and output shortfalls by using model (8), and then project the inefficient DMUs to the frontier by the following equation: $y^p = \theta^* \cdot y_o + q^{+*}$ while x remains unchanged. Next, we recalculate the efficiency and the scale elasticity for each DMU after projection. The results are presented in *Table 3*. As one can see from *Table 3,* as expected, the θ^* are all less than or equal to those φ measured by model (3), and the average efficiency is 1.102, in contrast to 1.404 measured by the BCC model. This is because of using the projected data and model (8) is more restrictive than model (3).

Similarly, we estimate the DSE for each railway under the assumption of strong congestion technology by using model (8), and the results are displayed in *Table 3*. By definition, strong congestion occurs if and only if $\bar{\rho} < 0$; as one can see from *Table 3,* there is no railway with negative scale elasticity, suggesting that no strong congestion is detected under the assumption of strong technology.

As mentioned in the above, strong congestion implies weak congestion, but not *vice versa*. Thus, we further investigate if the selected 24 railways have weak congestion by models (11) and (12), and the results of shortfalls in outputs as well as excesses in inputs are indicated in *Table 4*. From *Table 4,* we note that 12 railways are evaluated as exhibiting weak congestion since these railways have at least one shortfall in passenger or freight service (denoted as C in the last column of *Table 4*).

Table 3. Results of strong congestion

Country	Railways	φ_{BCC}	φ_{CV}	θ^*	q_1^+	q_2^+	$\bar{\rho}$
Austria	*ÖBB*	2.106	2.106	1.842	0	0	0.467
Belgium	*SNCB*	1.889	1.889	1.165	0	46	0.154
Finland	*VR*	1.422	1.422	1.422	0	177	7.472
France	*SNCF*	2.566	2.566	1.060	0	29	3.255
Germany	*DB AG*	2.010	2.010	1.000	0	0	1.240
Greece	*OSE*	1.000	1.000	1.000	0	0	7.294
Ireland	*CIE*	1.000	1.000	1.000	0	0	1.462
Italy	*FS SpA*	2.408	2.408	1.000	0	0	7.234
Luxemburg	*CFL*	1.000	1.000	1.000	0	0	3.121
Portugal	*CP*	1.988	1.988	1.825	0	0	0.156
Spain	*RENFE*	1.827	1.827	1.000	0	0	8.641
Switzerland	*BLS*	1.356	1.356	1.000	0	0	1.133
Switzerland	*CFF*	1.703	1.703	1.000	0	0	0.608
Bulgaria	*BDZ*	1.000	1.000	1.000	0	0	7.251
Croatia	*HZ*	1.018	1.000	1.000	0	0	74.298
Czech Rep.	*CD*	1.837	1.172	1.126	0	0	2.265
Estonia	*EVR*	1.000	1.000	1.000	0	0	8.022
Hungary	*MÁV Rt.*	1.731	1.016	1.000	0	0	1.616
Latvia	*LDZ*	1.000	1.000	1.000	0	0	2.565
Lithuania	*LG*	1.000	1.000	1.000	0	0	3.227
Poland	PKP	1.859	1.233	1.000	0	0	6.979
Romania	*CFR*	1.000	1.000	1.000	0	0	55.766
Slovenia	*SZ*	1.363	1.000	1.000	0	0	14.183
Turkey	*TCDD*	1.000	1.000	1.000	0	0	1.000
Average		1.503	1.404	1.102			

Note: q_1^+ and q_2^+ stand for the output shortfall of passenger and freight services, respectively.

Table 4. Results of weak congestion measures with consideration of environmental factors

Country	Railways	t_1^{+*}	t_2^{+*}	t_1^{-*}	t_2^{-*}	t_3^{-*}	Congestion
Austria	*ÖBB*	0	1,491	0	0.181	1.370	C
Belgium	*SNCB*	0	4,607	0	0.628	0.151	C
Finland	*VR*	0	0	0	0	0	–
France	*SNCF*	1,078	0	0	0.358	1.228	C
Germany	*DB AG*	915	0	0	0.264	2.761	C
Greece	*OSE*	0	0	0	0	0	–
Ireland	*CIE*	0	0	0	0	0	–
Italy	*FS SpA*	1,197	0	0	0.415	1.335	C
Luxemburg	*CFL*	0	0	0	0	0	–
Portugal	*CP*	0	596	0	0.085	0.038	C
Spain	*RENFE*	512	98	0	0.165	0.087	C
Switzerland	*BLS*	1,987	0	0	0.394	6.534	C
Switzerland	*CFF*	554	8,810	0	1.356	0	C
Bulgaria	*BDZ*	0	0	0	0	0	–
Croatia	*HZ*	0	0	0	0	0	–
Czech Rep.	*CD*	0	265	0	0.024	0.151	C
Estonia	*EVR*	0	0	0	0	0	–
Hungary	*MÁV Rt.*	0	200	0	0.053	0	C
Latvia	*LDZ*	0	0	0	0	0	–
Lithuania	*LG*	0	0	0	0	0	–
Poland	*PKP*	0	602	0	0.047	1.392	C
Romania	*CFR*	0	0	0	0	0	–
Slovenia	*SZ*	0	0	0	0	0	–
Turkey	*TCDD*	0	0	0	0	0	–
Average		1,040	2,083	0	0.33	1.50	

Note: The "C" in the last column means that the railway is evaluated as exhibiting weak congestion. t_1^{+*}, t_2^{+*} denote the shortfall in passenger and freight services, respectively; t_1^{-*}, t_2^{-*}, t_3^{-*} represent the excess in the number of Labor, Pcars, Fcars, respectively.

Based on the results, one can obtain more insights into railways' operation and thus propose more effective strategies for improvement. More specifically, the railways evaluated as congested in passenger or freight service, or both, should expand their outputs accordingly. On average, those 8 railways (see *Table 4*) can improve their freight service by expanding 2,083 thousand ton-kilometres/km and those 6 railways evaluated as exhibiting congestion in passenger service should improve their performance by augmenting 1,040 thousand passenger-kilometres/

km so as to eliminate the weak congestion. As for the excessive use of inputs, based on our results, the available capacity of freight transport service in terms of the number of freight cars is the most congestive input factor and should thus be contracted urgently, followed by available passenger cars. This finding is somewhat different from the conclusions of some previous studies, which in general suggested that overstaffing is the most serious problem.

At the disaggregate level, *Table 4* shows that we also found that some railways are evaluated as exhibiting weak congestion since a decrease of one or two inputs will lead to an increase in one or two outputs. Of the 12 railways evaluated as exhibiting weak congestion, ÖBB (Austria), SNCB (Belgium), SNCF (France), DB AG (Germany), FS SpA (Italy), CP (Portugal), RENFE (Spain), BLS (Switzerland), and CFF (Switzerland) operate in Western Europe, while CD (Czech), MÁV Rt. (Hungary), and PKP (Poland) in Eastern Europe, indicating that railways in Western Europe are more congestive than railways in Eastern Europe. Looking at the sample, we find that the railways evaluated as exhibiting weak congestion are mostly larger-scale ones with an operating line greater than 10,000 kilometres such as DB AG (Germany), SNCF (France), FS SpA (Italy), RENFE (Spain), PKP (Poland), CD (Czech), and MÁV Rt (Hungary). The policy implication is that these railways should contract input accordingly in order to improve their operating performance, especially for those larger-scale railways with larger amount of output shortfall.

5. CONCLUSION

In recent years, many researchers have focused their attention on the investigation of the technical efficiency and scale elasticity of production; most of them adopted the BCC DEA model as a useful tool. However, Tone – Sahoo (2004) pointed out that the BCC model generally overstated true scale elasticity when congestion is present. To rectify this, Tone – Sahoo (2004) proposed a new scheme to investigate scale elasticity in the presence of congestion within a unified framework. It is thus meaningful to apply the proposed new scheme to investigate the sources of congestion for industries with financial losses such as rail transport.

The production of railways, including passenger and freight service outputs, are significantly influenced by certain environmental factors such as GNI per capita, which vary from country to country. To take into account the environmental factor, in this study we go one step further by measuring the efficiency and output slack as well as input congestion with consideration of the category variable. We divide the sample into two subsets, Eastern and Western Europe, and then impose the constraint containing a dummy variable on the models proposed

by Tone – Sahoo (2004). Our empirical results indicate that of 24 observations, 13 railways are measured as technical inefficiency; none has been detected as strong congestion, nonetheless, and 12 railways are evaluated as exhibiting weak input congestion.

The major contribution of this study is that we identify the sources of input congestion in the railway transport performance measurement by extending the DEA method and incorporating the categorical variables into the analysis, from which one can obtain more insights into railways' operation and thus propose more effective strategies for improvement.

Different railway technologies (electrified *vs*. diesel or steam traction power) and different ownerships (public *vs*. private) are frequently associated with different levels of output slacks as well as input congestions. One of the possible avenues in future studies is to investigate the effects of these internal factors on slacks and congestions, and to apply the proposed model to the longer panel data for rail transport performance evaluation.

REFERENCES

Ali, A. I. – Seiford, L. M. (1993): The Mathematical Programming Approach to Efficiency Analysis. In: Fried, H. O. – Lovell, C. A. K. – Schmidt, S. S. (eds): *The Measurement of Productive Efficiency: Techniques and Applications.* Boston: Oxford University Press, pp. 120–159.

Banker, R. D. – Charnes, A. – Cooper, W. W. (1984): Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science,* 30: 1078–1092.

Banker, R. D. – Morey, R. C. (1986a): Efficiency Analysis for Exogenously Fixed Inputs and Outputs. *Operations Research,* 34(4): 513–521.

Banker, R. D. – Morey, R. C. (1986b): The Use of Categorical Variables in Data Envelopment Analysis. *Management Science,* 32(12): 1613–1627.

Bookbinder, J. H. – Qu, W. W. (1993): Comparing the Performance of Major American Railroads. *Transportation Research Forum,* 33(1): 70–85.

Brockett, P. L. – Cooper, W. W. – Wang, Y. – Shin, H. (1998): Inefficiency and Congestion in Chinese Production before and after the 1978 Economic Reform. *Socio-Economic Planning Science,* 32(1): 1–20.

Chapin, A. – Schmidt, S. (1999): Do Mergers Improve Efficiency? *Journal of Transport Economics and Policy,* 33(2): 147–162.

Charnes, A. – Cooper, W. W. – Rhodes, E. (1978): Measuring the Efficiency of Decision Making Units. *European Journal Operational Research,* 2(6): 429–444.

Cooper, W. W. – Gu, B. – Li, S. (2001): Comparisons and Evaluations of Alternative Approaches to the Treatment of Congestion in DEA. *European Journal of Operational Research,* 132(1): 62–74.

Cooper, W. W. – Thompson, R. G. – Thrall, R. M. (1996): Introduction: Extension and New Developments in DEA. *Annals of Operations Research,* 66(1): 3–45.

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. – Lovell, C. A. K. (1985): *The Measurement of Efficiency of Production*. Boston: Kluwer-Nijhoff Publishing.

Färe, R. – Svensson, L. (1980): Congestion of Production Factors: *Econometrica,* 48: 1745–1753.

Førsund, F. R. (2001): Categorical Variables in DEA. *Working Paper*, Department of Economics, University of Oslo, Norway.

Kao, C. (2010): Congestion Measurement and Elimination under the Framework of Data Envelopment Analysis. *International Journal of Production Economics,* 123(2): 257–265.

Lan, L. W. – Lin, E. T. J. (2005): Measuring Railway Performance with Adjustment of Environmental Effects, Data Noise and Slacks. *Transportmetrica,* 1: 161–189.

Lan, L. W. – Lin, E. T. J. (2006): Performance Measurement for Railway Transport: Stochastic Distance Functions with Inefficiency and Ineffectiveness Effects. *Journal of Transport Economics and Policy,* 40: 383–408.

Odeck, J. (2006): Congestion, Ownership, Region of Operation, and Scale: Their Impact on Bus Operator Performance in Norway. *Socio-Economic Planning Science,* 40: 52–69.

Oum, T. H. – Yu, C. (1994): Economic Efficiency of Railways and Implications for Public Policy. *Journal of Transport Economics and Policy,* 28: 121–138.

Ruggiero, J. (1996): On the Measurement of Technical Efficiency in the Public Sector. *European Journal of Operational Research,* 90: 553–565.

Ruggiero, J. (1998): Non-Discretionary Inputs in Data Envelopment Analysis. *European Journal of Operational Research,* 111: 461–469.

Tone, K. – Sahoo, B. K. (2004): Degree of Scale Economies and Congestion: A Unified DEA Approach. *European Journal of Operational Research,* 158: 755–772.

UIC (2010): *International Railway Statistics: Year of 2008*. International Railway Union, Paris.

Viton, P. A. (1997): Technical Efficiency in Multi-mode Bus Transit: A Production Frontier Analysis. *Transportation Research Part B,* 31: 23–39

Yu, M. M. (2008): Assessing the Technical Efficiency, Service Effectiveness, and Technical Effectiveness of the World's Railways through NDEA Analysis. *Transportation Research Part A: Policy and Practice,* 42: 1283–1294.