CHAPTER 3. METHODOLOGY AND RESEARCH FRAMEWORK

This chapter presents two frontier approaches for measuring efficiency and effectiveness, Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA). In addition to the DEA and SFA, a non-parametric method for measuring productivity index and sales force index are also included in this chapter. Once the approaches have been described, the research framework is then presented. The chapter is organized as follows, 3.1 presents the methodologies, including DEA, SFA and the comparison between two methods. The methods for measuring productivity and sales force indexes are presented in 3.2, and the research framework follows.

3.1 The Methods for Measuring Efficiency and Effectiveness

In the neoclassical production economics, a production technology may be represented in many ways. The economists usually relate outputs and inputs by estimating production function or its duality: cost function. A shortcoming in specifying production or cost function is its strong assumption of profit maximization or cost minimization. In many cases, due to some reasons, the firms may not produce the outputs efficiently. In other words, the behaviors of DMU dose not satisfy the strong assumption. Another drawback is the data availability. The cost function depends on output quantities and input prices. It is impossible to estimate cost function if the price information is not available. Since this research attempts to compare the efficiency, effectiveness, productivity, and sales force of some selected worldwide railways, in consideration of the data availability and/or monetary conversion for the input factor prices among different countries, therefore, DEA and SFA are adopted in this research. The followings describe each of these methods in turn.

3.1.1 Data Envelopment Analysis

The data envelopment analysis begins with Edwardo Rhodes's Pd. D. disseration research. Since that, many DEA models have been developed. In this research, CCR, BCC and SZ models are applied to measure efficiency and effectiveness, which can be briefly described as follows.

3.1.1.1 CCR model

In 1978, Charnes, Cooper and Rhodes (CCR) introduced a mathematical programming method to measure the relative efficiency for organizations or firms and termed as date envelopment analysis (DEA). The advantage of DEA is that no explicit functional forms need to be imposed on the data. Thus, the use of DEA has become increasingly widespread since then. The CCR DEA model is a fractional programming, which can be transformed to linear programming (LP) as follows:

$$\text{(3-1)}\quad \begin{aligned} & Min_{\theta,\lambda}\theta \\ & s.t.\ \ -y_i + Y\cdot\lambda \geq 0, \\ & \qquad \theta\cdot x_i - X\cdot\lambda \geq 0,\quad \lambda \geq 0 \end{aligned}$$

Where X and Y are the K×N input matrix and the M×N output matrix, respectively. For the ith firm these are represented by the vector x_i and y_i, respectively. λ is a N×1 vector of constant and θ is a scalar, which stands for efficiency of ith firm. Solve this LP for each of the N firms; one obtains the efficiency score for each firm. One can easily transform model (3-1) to output orientation DEA forms as shown in model (3-2).

$$\text{(3-2)}\quad \begin{aligned} & Max_{\phi,\lambda}\phi \\ & s.t.\ \ -\phi\cdot y_i + Y\cdot\lambda \geq 0 \\ & \qquad x_i - X\cdot\lambda \geq 0 \\ & \qquad \lambda \geq 0 \end{aligned}$$

Where Y, X, x_i, y_i and λ are defined as previous; ϕ denotes proportional increase in output, which ranges from one to infinity; $1/\phi$ defines the service effectiveness of firm, which varies between zero and one.

3.1.1.2 BCC model

Note that model (3-1) is an input orientation DEA model under the assumption of constant returns to scale (CRS) technology. Banker, Charnes and Cooper (BCC, 1984) relaxed the restriction of CRS to account for variable returns to scale (VRS) technologies by adding convexity constraint to model (3-1). The BCC input orientation DEA model then becomes:

$$\text{(3-3)}\quad \begin{aligned} & Min_{\theta,\lambda}\theta \\ & s.t.\ \ -y_i + Y\cdot\lambda \geq 0, \\ & \qquad \theta\cdot x_i - X\cdot\lambda \geq 0, \\ & \qquad \sum\lambda = 1,\quad \lambda \geq 0 \end{aligned}$$

Similarly, one can easily transform input-oriented BCC model into output-oriented BCC model as shown in (3-4).

$$
\begin{aligned}
& Max_{\phi,\lambda}\phi \\
(3\text{-}4) \quad & s.t. \quad -\phi \cdot y_i + Y \cdot \lambda \geq 0 \\
& \qquad x_i - X \cdot \lambda \geq 0 \\
& \qquad \sum \lambda = 1, \lambda \geq 0
\end{aligned}
$$

3.1.1.3 SZ model

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

$$
\begin{aligned}
& \beta^* = Min\beta \\
(3\text{-}5) \quad & subject\ to \\
& \sum_{\substack{j=1\\ j\neq k}}^{n} \lambda_j x_{ij} \leq \beta_{ik} x_{ik}, \sum_{\substack{j=1\\ j\neq k}}^{n} \lambda_j y_{rj} \geq y_{rk}, \sum_{\substack{j=1\\ j\neq k}}^{n} \lambda_j = 1,\ \beta, \lambda_j \geq 0, (j \neq k)
\end{aligned}
$$

Seiford and Zue (1998b) show that under the circumstance of $1 \leq \sqrt{\beta^*}$, where β^* is the optimal value to (3-5), an efficient DMU_k with efficiency score equal to 1.000 will still remain efficient, if the percentages increase in all inputs for the DMU_k are less than $g_k = \sqrt{\beta^*} - 1$ and the percentages decrease in all inputs for the remaining DMUs are less than $g_{-k} = (\sqrt{\beta^*} - 1)/\sqrt{\beta^*}$. The upper-bound levels (g_k, g_{-k}) can be viewed as the sensitivity indexes. Similarly, consider the following LP model

$$
\begin{aligned}
& \alpha^* = Max\alpha \\
(3\text{-}6) \quad & subject\ to \\
& \sum_{\substack{j=1\\ j\neq k}}^{n} \lambda_j x_{ij} \leq x_{ik}, \sum_{\substack{j=1\\ j\neq k}}^{n} \lambda_j y_{rj} \geq \alpha y_{rk}, \sum_{\substack{j=1\\ j\neq k}}^{n} \lambda_j = 1,\ \alpha, \lambda_j \geq 0, (j \neq k)
\end{aligned}
$$

Seiford and Zue (1998b) also show that under the circumstance of $\sqrt{\alpha^*} \leq 1$,, where α^* is the optimal value to (3-6), an efficient DMU_k will remain efficient, if the percentages decrease in all outputs for the DMU_k are less than $h_k (= 1 - \sqrt{\alpha^*})$ and the percentages increase in all outputs for the remaining DMUs are less than $h_{-k} (= (1-\sqrt{\alpha^*})/\sqrt{\alpha^*})$.

The upper-bound levels (h_k, h_{-k}) are the sensitivity indexes.

3.1.2 Stochastic Frontier Analysis

SFA is an alternative method for measuring efficiency of production. The most popular functional forms used in the literature are stochastic production and cost function, which are briefly described as follows.

3.1.2.1 Stochastic production frontier

Inspiration of Farrell's concepts, Aigner and Chu (1968) proposed a method for estimating a parametric frontier production function. Their model is as follows.

(3-7) $\ln(y_i) = f(x_i;\beta) - u_i,\ i = 1,..., N$

Where $\ln(y_i)$ is the logarithm of the output for the i-th DMU; x_i is a vector , whose elements are the logarithms of the K-input quantities used by the i-th DMU; β is a vector of unknown parameters to be estimated; and u_i is non-negative random variable, associated with technical inefficiency in production of DMUs. Instead of econometric technique, they proposed to use so-called Parametric Linear Programming (PLP) for estimating production function and thus measuring the efficiency of firms, that is, technical efficiency of i-th DMU becomes:

(3-8) $TE_i = \dfrac{y_i}{\exp(f(x_i;\beta))} = \dfrac{\exp(f(x_i;\beta) - u_i)}{\exp(f(x_i;\beta))} = \exp(-u_i), i = 1, 2, \ldots, N$

This efficiency measure is an out-oriented Farrell measure of technical efficiency. Aigner *et al.* (1977) proposed a composite error to count technical efficiency and statistical noise. The model can be defined as

(3-9) $y_i = f(x_i;\beta)\times\exp(v_i)\times\exp(-u_i) = f(x_i;\beta)\times\exp(v_i)\times TE_i$

Where y_i is the output of *i-th* firm, v_i is symmetric random error term. Aigner *et al.* (1977) assume that v_i follows a normal distribution with zero mean and constant variance, and u_i is non-negative independent and identical distributed (*i.i.d.*) random variable, which counts technical inefficiency of firms. Then,

(3-10) $TE_i = \exp(-u_i) = \dfrac{y_i}{f(x_i;\beta)\times\exp(v_i)}, i = 1, 2, \ldots, N$

In order to estimate u_i , one has to impose a distribution form (such as half-normal, truncated-normal, gamma, etc.) on the model. For example, one specifies half-normal distribution, that is, assume (Kumbhakar and Lovell, 2000):

i) $v_i \sim$ *i.i.d.* $N(0, \sigma_v^2)$

ii) $u_i \sim$ *i.i.d.* $N^+(0, \sigma_u^2)$

iii) Both v_i *and* u_i *are independently distributed of each other, and of the regressors..*

Because v_i is independent of u_i, the joint p.d.f. of u_i and v_i are

(3-11) $f(\varepsilon)=\dfrac{2}{\sigma\sqrt{2\pi}}\exp\left[1-\Phi\left(\dfrac{\varepsilon\lambda}{\sigma}\right)\right]\times\exp\left(-\dfrac{\varepsilon^2}{2\sigma^2}\right)=\dfrac{2}{\sigma}\phi\left(\dfrac{\varepsilon}{\sigma}\right)\Phi\left(-\dfrac{\varepsilon\lambda}{\sigma}\right)$

Where, $\varepsilon = v - u$, $\sigma=\left(\sigma_u^2+\sigma_v^2\right)^{1/2}, \lambda=\dfrac{\sigma_u}{\sigma_v}, \phi(\cdot)\ and\ \Phi(\cdot)$ are standard normal cumulative distribution function and density function, respectively. The log likelihood function of $f(\varepsilon)$ is

(3-12) $\ln L = const - N\ln\sigma + \sum_{i=1}^{N}\ln\Phi\left(-\dfrac{\varepsilon_i\lambda}{\sigma}\right)-\dfrac{1}{2\sigma^2}\sum_{i=1}^{N}\varepsilon_i^2$

Then, one can estimate by using maximum likelihood estimation method. Jondrow *et al.* (1982) have derived

(3-13) $E\langle u_i\,|\,\varepsilon_i\rangle=\mu_{*i}+\sigma_*\left[\dfrac{\phi(-\mu_{*i}/\sigma_*)}{1-\Phi(-\mu_{*i}/\sigma_*)}\right]=\sigma_*\left[\dfrac{\phi(\varepsilon_i\lambda/\sigma)}{1-\Phi(\varepsilon_i\lambda/\sigma)}-\left(\dfrac{\varepsilon_i\lambda}{\sigma}\right)\right]$

Where, $\mu_{*i}=\dfrac{-\varepsilon\sigma_u^2}{\sigma^2}, \sigma_*^2=\dfrac{\sigma_u^2\sigma_v^2}{\sigma^2}$

The technical efficiency of firms then becomes

(3-14) $TE_i=\exp\left(-\hat{u}_i\right)=\exp\left(-E\langle u_i\,|\,\varepsilon_i\rangle\right)$

Battese and Coelli (1988) (hereafter BC) proposed another point estimator for TE_i

(3-15) $TE_i=E\left[\exp\langle -u_i\,|\,\varepsilon_i\rangle\right]=\left[\dfrac{1-\Phi(\sigma_*-\mu_{*i}/\sigma_*)}{1-\Phi(-\mu_{*i}/\sigma_*)}\right]\exp\left(-\mu_{*i}+\dfrac{1}{2}\sigma_*^2\right)$

For a nonlinear function $g(x)$, $E[g(x)]$ is not equal to $g(E[x])$, Kumbhakar and Lovell (2000) pointed out that BC is preferred. Hence, this research uses BC estimator.

3.1.2.2 Stochastic cost frontier

The stochastic production frontier described in 3.1.2.1, has a shortcoming, estimation of a production frontier requires that producers produce only a single output. This is not the case for many producers. To deal with the efficiency of a multiple-outputs and multiple-inputs firm, the stochastic cost frontier method than been developed. Based on the literatures, the cost efficiency, CE_i can be defined as

(3-16) $CE_i=\dfrac{c(y_i,w_i;\beta)}{E_i}=\dfrac{min.\ feasible\ cost}{observed\ cost}$.

In (3-16), the cost frontier $c\ (y_i,\ w_i,\ \beta)$ is deterministic, which can be transformed into a stochastic form by adding random term to it, as shown in (3-17).

$$(3\text{-}17)\quad CE_i = \frac{c(y_i, w_i; \beta)\cdot exp(v_i)}{E_i}$$

Equation (3-17) defines the cost efficiency as the ratio of minimum cost attainable in a random environment characterized by exp (v_i) to observed expenditure, E_i. CE_i is bounded between zero and unity. CE_i=1, if and only if the firm is cost efficient. If we assume that $c\ (y_i, w_i, \beta)$ takes the log-linear functional form, then

$$(3\text{-}18)\quad \ln E_i = \beta_0 + \sum \beta_m \ln y_m + \sum \beta_j \ln w_j + v_i + u_i\ .$$

Where, v_i is the two-sided random error term, and u_i is the non-negative term which captures the cost inefficiency. That is

$$(3\text{-}19)\quad CE_i = \exp(-u_i)\ .$$

Similar to production function case, one needs to make the distributional assumptions on the random error term and the inefficiency term, such as:

i) $v_i \sim i.i.d.\ N\ (0, \sigma_v^2)$

ii) $u_i \sim i.i.d.\ N^+(0, \sigma_u^2)$

iii) v_i *and* u_i *are independently distributed of each other, and of the regressors.*

Because v_i is independent of u_i, the joint p.d.f. of u_i and v_i are

$$(3\text{-}20)\quad f(\varepsilon) = \frac{2}{\sigma\sqrt{2\pi}}\exp\left[1-\Phi\left(\frac{-\varepsilon\lambda}{\sigma}\right)\right]\times\exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right) = \frac{2}{\sigma}\phi\left(\frac{\varepsilon}{\sigma}\right)\Phi\left(\frac{\varepsilon\lambda}{\sigma}\right)$$

Where, $\varepsilon = v + u$, $\sigma = \left(\sigma_u^2 + \sigma_v^2\right)^{1/2}, \lambda = \frac{\sigma_u}{\sigma_v}, \phi(\cdot)\ and\ \Phi(\cdot)$ are standard normal cumulative distribution function and density function, respectively. Once joint p.d.f. has been defined, one can estimate by using maximum likelihood method as described in 3.1.2.1.

3.1.3 The Comparison between DEA and SFA

Although both DEA and SFA are attributed to frontier methods, and both are inspirited by Farrell's (1957) efficiency measurement concept, however, DEA uses different techniques from SFA to construct frontiers. DEA uses the mathematical programming technique; while SFA applies the econometric technique to estimate the inefficiency of production or cost or revenue functions. The two approaches have some common merits and some monopolistic advantages or disadvantages, which can be summarized in Table 3-1.

Table 3-1 the comparison between DEA and SFA

Items to be compared.	D E A	S F A
Parametric/ nonparametric	Non-parametric frontier method	Parametric frontier method
Technique	Mathematical programming technique.	Econometric technique.
Can be use to measure	Technical efficiency. Scale efficiency. Scale economies. Allocative efficiency. Productivity (if panel data are available). Technical change and efficiency change (if panel data are available). Congestion efficiency.	Technical efficiency. Scale efficiency. Scale economies. Allocative efficiency. Productivity. Technical change and efficiency change.
Assumption	The frontier is constructed by the best firms and all deviations from frontier are imputed to inefficiency.	The frontier which firms faced is stochastic, it may influenced by weather, strike, luck, measurement errors, etc. Exclusion of these statistical noises, the remaining deviations from frontier can be attributed to inefficiency.
Merit	It does not assume that all firms are fully efficient. It can be used to deal with multi- output and multi-input problem. It does not require price information. It does not require specification of a functional form. It is not need to specify a distributional form for the inefficiency term. When measuring productivity, it permits TFP to be decomposed into technical change and efficiency change.	It does not assume that all firms are fully efficient. It accounts for noise, such as the effects of weather, strikes, luck, etc. It does not require price information. It can be used to conduct tests of hypotheses. When measuring productivity, it permits TFP to be decomposed into technical change and efficiency change.
Shortcoming	It does not account for noise. It cannot be used to conduct tests of hypotheses. The inclusion of outliers may reduce efficiency scores.	It is need to specify a functional form for production or cost function. It is need to specify a distributional form for the inefficiency term. It is sensitive to distribution form assumed.
Limitation	The efficiency scores are only relative to the best firms in the sample. Measurement error and other noise may influence the shape and position of the frontier. Exclusion of an important input or output may bias the results. Not accounting for environmental difference may give misleading results.	The efficiency scores are only relative to the best firms in the sample. Sufficient observations are needed so as to enlarge degree of freedom. Not accounting for environmental difference may give misleading results.

Both DEA and SFA are by now two well-established non-parametric and parametric techniques, respectively, widely employed in management science. However, some researchers criticize that SFA methods do not process practicality in real applications. If one goes away from academics will there be people using SFA as a tool for running a firm? Färe (1996) mentioned at the Advanced Research Workshop on Efficiency Measurement, held at Odense University, May 22-24, 1995, "I have seen DEA being used but I have never seen anyone use SFA." On the other hand, Lovell (1996) criticized that there is no way to construct confidence intervals for the results based on DEA measurements. He argued at the same workshop, “how confident are we in the numbers we are generating from DEA?” Although none of methods is perfect, however, DEA and SFA provide substantially better measures of efficiency than simple partial measures. In application fields, one frequently asked question is: which method should one use? The answer to this question often depends upon the application being considered. If one is using farm level data where measurement error, missing variables, for example, data on an input is not available or not suitably measured, whether, etc. are likely to play a significant role, then the assumption that all deviations from the frontier are due to inefficiency, which is made by DEA, may be a brave assumption (Coelli, 1995). In other words, when measuring the efficiency of the firms or industries, in which their data are indeterminist, then SFA is more suitable.

DEA and SFA have become the two important and widespread analytical tools in the applications of efficiency measurement. With increasing applications in the empirical studies, some questions become more critical. Seiford (1996) has pointed out two major issues. The first one is that how do we know that the model selected best approximates the reality of the problem? The second one is that how do we know the results are correct? To valid the results of DEA, one needs some statistical tests for model specification. Besides, to overcome the shortcoming of data errors, perhaps one needs some innovative technique, which incorporate DEA and SFA. Lewin and Lovell (1990) also pointed out that two issues are necessary to be researched in the future. In the theory area, ongoing work in pursuit of a convergence of the two techniques can be expected both by making DEA stochastic and by relaxing parametric restrictions in the econometric models. In the applications area, we see a need for the simultaneous use of both techniques, particularly when important policy recommendations might be sensitive to the choice of analytical method.

3.2 The Method for Measuring Productivity and Sales Force

As mentioned in previous chapter, Malmquist (1953) defined input quantity index in the consumer context, Caves *et al*. (1982) defined the analogous productivity index, which is named after Malmquist (1953). Unlike Caves *et al*. (1982), Färe *et al*. (1994) calculated the Malmquist index directly by using the technical efficiency measures developed by Farrell (1957). In fact, the Farrell’s (1957) measurement in efficiency is based on the input distance function; which is introduced by Shephard (1953, 1970), and which can be described as follows.

Let $x = (x_1,...,x_N) \in R_+^N$ denote a vector of inputs, and $y = (y_1,...,y_M) \in R_+^M$ denote a vector of output. The production technology T is defined by $T = \{(x, y) : x\, can\, producey\}$, and it consists of all input-output vectors that are technically feasible. The input distance function is defined on the technology T as

$D_i(x,y)=\sup\{\lambda:(\frac{x}{\lambda},y)\in T\}$. In words, that is to find the maximal feasible contraction ratio of *x*, keeping all observed data point remains in the feasible set. Similarly, the output distance function is defined as $D_o(x,y)=\inf\{\theta:(x,y/\theta)\in T\}$. Following Färe *et al*. (1994), to define the Malmquist index; one needs to define distance functions with respect to two different time periods such as

(3-21) $D_O^t(x^{t+1},y^{t+1})=\inf\{\theta:(x^{t+1},y^{t+1}/\theta)\in T^t\}$

This distance function measures the maximal proportional change in outputs required to make (x^{t+1}, y^{t+1}) feasible in relation to the technology at period t. Note that, if the production (x^{t+1}, y^{t+1}) occurs outside the set of feasible production in period t, that is, technical change has occurred. Similarly, one may define a distance function that measures the maximal proportional change in output required to make (x^t, y^t) feasible in relation to the technology at period t+1, which Färe *et al*. (1994) called $D_o^{t+1}(x^t, y^t)$. Caves *et al*. (1982) defined the Malmquist productivity index as

(3-22) $M_{CCD}^t=\dfrac{D_O^{t+1}(x^{t+1},y^{t+1})}{D_O^t(x^t,y^t)}$

In (3-22), technology in period t is the reference technology. Similarly, the period t+1-based Malmquist productivity index can be defined as

(3-23) $M_{CCD}^{t+1}=\dfrac{D_O^{t+1}(x^{t+1},y^{t+1})}{D_O^{t+1}(x^t,y^t)}$

In order to avoid choosing an arbitrary benchmark, Färe *et al*. (1994) specified the output-based Malmquist productivity index as the geometric mean of the two CCD-type Malmquist productivity index and as the follows.

(3-24) $M_O(x^{t+1},y^{t+1},x^t,y^t)=\left[\dfrac{D_O^t(x^{t+1},y^{t+!})}{D_O^t(x^t,y^t)}\times\dfrac{D_O^{t+1}(x^{t+1},y^{t+!})}{D_O^{t+1}(x^t,y^t)}\right]^{\frac{1}{2}}$

Following Färe *et al*. (1989, 1992), model (3-26) can be rewrite as

(3-25) $M_O(x^{t+1},y^{t+1},x^t,y^t)=\dfrac{D_O^{t+1}(x^{t+1},y^{t+!})}{D_O^t(x^t,y^t)}\left[\dfrac{D_O^t(x^{t+1},y^{t+!})}{D_O^{t+1}(x^{t+1},y^{t+!})}\times\dfrac{D_O^t(x^t,y^t)}{D_O^{t+1}(x^t,y^t)}\right]^{\frac{1}{2}}$

Model (3-25) can be decomposed into two terms, technical change (TC) and efficiency change (EC) as (3-26) and (3-27), respectively.

(3-26) $TC=\left[\dfrac{D_O^t(x^{t+1},y^{t+!})}{D_O^{t+1}(x^{t+1},y^{t+!})}\times\dfrac{D_O^t(x^t,y^t)}{D_O^{t+1}(x^t,y^t)}\right]^{\frac{1}{2}}$

(3-27) $EC=\dfrac{D_O^{t+1}(x^{t+1},y^{t+!})}{D_O^t(x^t,y^t)}$

Similar to (3-25), (3-26) and (3-27), one can defines the input-based Malmquist productivity index as (3-28) and decomposes it into (3-29) and (3-30)

(3-28) $$M_I(x^{t+1}, y^{t+1}, x^t, y^t) = \frac{D_I^{t+1}(x^{t+1}, y^{t+!})}{D_I^t(x^t, y^t)}\left[\frac{D_I^t(x^{t+1}, y^{t+!})}{D_I^{t+1}(x^{t+1}, y^{t+!})} \times \frac{D_I^t(x^t, y^t)}{D_I^{t+1}(x^t, y^t)}\right]^{\frac{1}{2}}$$

(3-29) $$TC = \left[\frac{D_I^t(x^{t+1}, y^{t+!})}{D_I^{t+1}(x^{t+1}, y^{t+!})} \times \frac{D_I^t(x^t, y^t)}{D_I^{t+1}(x^t, y^t)}\right]^{\frac{1}{2}}$$

(3-30) $$EC = \frac{D_I^{t+1}(x^{t+1}, y^{t+!})}{D_I^t(x^t, y^t)}$$

The four distance functions can be calculated by solving four linear programming as described in the previous chapter.

3.3 The Research Framework

3.3.1 The Three-stage Model

Although DEA has become increasingly widespread in the efficiency measurement in the past two decades, however, existence of input excesses and outputs slacks, neglecting the affects of environmental factors, and without taking statistical errors into account are its common shortcomings. The slack problems arise because we use the sections of piece-wise linear to stand for isoquant. As shown in Figure 3-1, assume that there are five DMUs, both C and E located on the frontier, therefore, they are efficient. DMUA and DMUB, on the other hand, are inefficient which can be moved to C and D, respectively, by proportionally reducing its inputs. However, it is questionable as to whether the point D is an efficient point because D could reduce the amount of input x_I used (by the amount of S_2) and still produce the same amount of output. This is known as non-radial input slack in the literature. In this research, two kinds of slacks are defined, one is radial slack, S_1, and the other is non-radial slack, S_2.

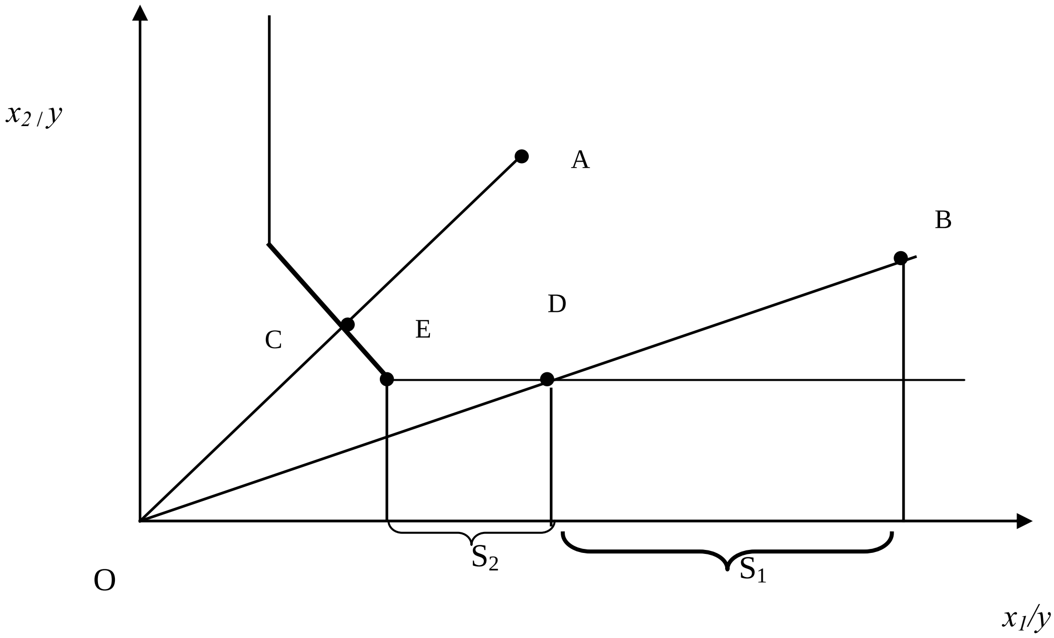

Figure. 3-1 Radial and Non-radial Input Slacks

Similar to input slacks, the output slacks can be portrayed by using Figure 3-2. In which, each DMU utilize one input to produce two outputs. C and E are output-oriented efficient, and B and A are inefficient. S_2 (that is DE) is the non-radial slack.

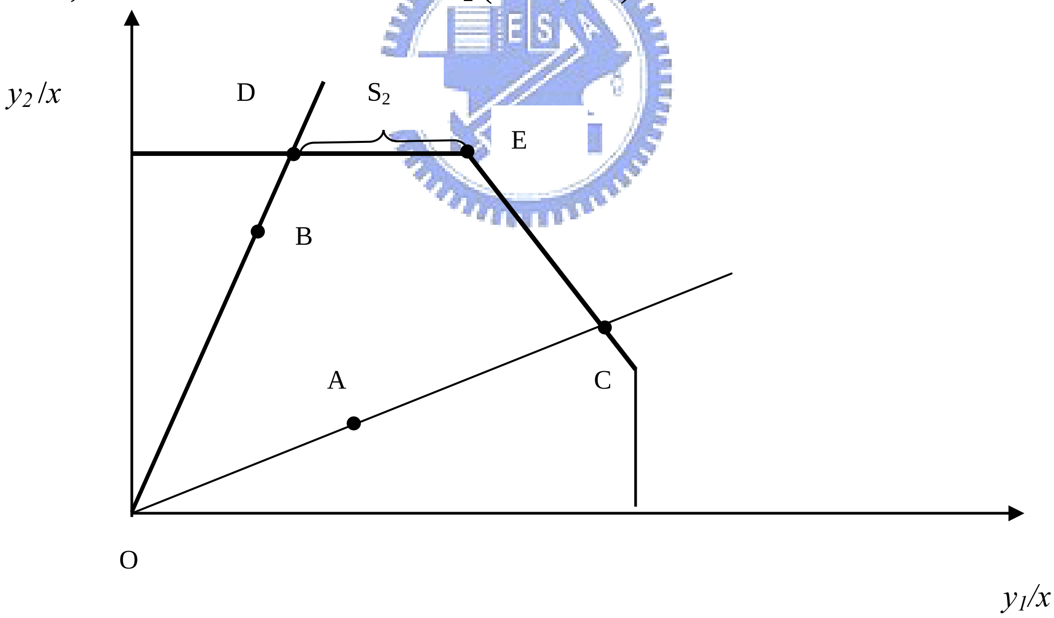

Figure. 3-2 Radial and Non-radial Output Slacks

Some researchers are aware that these shortcomings may influence the measure results and thus propose several revised models. Based on literatures, the revised models can be classified into three categories. The first, to take the non-discretionary environmental factors into account, for example, Banker and Morey (1986a) propose an exogenously fixed inputs and outputs DEA model. Banker and Morey (1986b) further introduce a categorical DEA model, in which benchmark is classified into several reference sets based on operating environments. Then a particular DMU only be compared to other DMUs with the same rating of operating environment. The second category, to consider the external operating environment and slacks simultaneously, Fried *et al*. (1993) adopt traditional DEA model to evaluate performance of U.S. credit unions in the first stage, and then regress the sum of radial and non-radial slacks on some explaining variables by using seemingly unrelated regression (SUR) technique in the second stage. Some researchers consider that the ability of a DMU to transform inputs into outputs is influenced by its technical efficiency and external operating environment. Fried *et al*. (1999) thus introduced a procedure for obtaining a measure of managerial efficiency that controls for exogenous features of operating environment. Since the slacks are always generated in conventional DEA models, which are assumed influenced by the environmental factors, statistical noises, and managerial efficiency, Fried *et al*. (2002) thus proposed a three-stage technique. The method proposed by Fried *et al*. (2002) can be attributed to the third category, and are described as follows.

In the first stage, conventional DEA model (CCR and BCC model, that is, (3-1) and (3-2)) are applied to measure the preliminary efficiency score for each DMU using input and output quantity data only. The results thus are used to compare and to test which model is more suitable for rail industry. The optimal one provides initial performance evaluation for each DMU. In addition to efficiency scores, the solutions also contain nonnegative slacks as shown in Figure 3-1. However, actual performances are likely to be under-evaluate since environmental factors and statistical noises, as well as slacks are neglected.

The total slacks estimated from the first stage (radial and non-radial slacks) then regress on environmental factors by using stochastic frontier analysis (SFA) approach in the second stage. The N separate SFA regressions take the general form

(3-31) $S_{ni} = f^n(z_i;\beta^n) + v_{ni} + u_{ni}, n = 1,..., N, i = 1,..., I$

Where the $f^n(z_i;\beta^n)$ are deterministic slack frontier with composed error terms $v_{ni} + u_{ni}$, β^n are the parameters to be estimated. Consistent with a stochastic cost frontier formulation, Fried *et al*. (2002) assumed that the $v_{ni} \sim N(0,\sigma_{vn}^2)$ reflect statistical noises and the $u_{ni} \sim N^+(\mu^n,\sigma_{un}^2)$ reflect the managerial inefficiency. This allowed the slacks to be decomposed into three parts, a part attributable to environmental effects, a part attributable to managerial efficiency, and a part attributable to statistical noise. Each of the N regressions may be estimated by maximum likelihood technique described in 3.1.2 section.

In the third stage, input or output data, (depending on the orientation used in the first stage) are adjusted by using the results of SFA regressions in the second stage. The equation for adjustment is as

(3-32) $x_{ni}^{A} = x_{ni} + \left[\max_i (z_i \beta^n) - z_i \beta^n\right] + \left[\max_i (v_{ni}) - v_{ni}\right], n = 1,\ldots, N, i = 1,\ldots, I$

Where x_{ni}^{A} and x_{ni} are adjusted and observed input quantities, respectively. The first adjustment on the right hand side of the equation puts all DMU into a common operating environment, the least favorable environment observed in the sample. The second adjustment puts all DMU into a common state of nature, the unluckiest situation encountered in the sample. After adjusting the data, then re-evaluate producer performance by using DEA model again.

3.3.2 The Four-stage Model

The analytical procedure proposed by Fried *et al*. (2002) take environmental effects and statistical noise into account when measuring efficiency, however, there is no guarantee that such measurement can always completely eliminate the slacks. In order to measure the efficiency of DMUs by taking the residual slacks into account, this research extends Fried's *et al.* (2002) three-stage DEA to four-stage DEA and further expands it to effectiveness, productivity, and sales force measurement. The analytical framework can be described as follows.

3.3.2.1 The efficiency measurement

The first-stage

In the first stage, we use input-orientation DEA (measuring the maximum possible proportional reduction in all inputs, keeping all outputs fixed) to measure the technical efficiency and productivity 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 inputs and passenger-train-kilometer per kilometer of lines and freight-train-kilometer per kilometer of lines as inputs and passenger-train-kilometer per kilometer of lines and freight-train-kilometer per kilometer of lines as outputs. Assume that there are *n* firms, each of them produces *k* products by utilizing *m* input factors; the input-orientation BCC DEA model can then be mathematically described as follows (Banker *et al*., 1984).

$$
(3\text{-}33)\quad \begin{aligned}
&Minimize_{\theta,\lambda}\theta \\
&s.t.\ \ -y_i + Y\cdot\lambda \ge 0, \\
&\qquad \theta\cdot x_i - X\cdot\lambda \ge 0, \\
&\qquad \sum\lambda = 1,\quad \lambda \ge 0
\end{aligned}
$$

Where, X and Y are the $m\times n$ input matrix and the $k\times n$ output matrix (for the ith firm these are represented by the vector x_i and y_i), respectively. λ is a $n\times1$ vector of constant and θ is a scalar, which stands for efficiency of ith firm. Solve this LP for each of the n firms; one obtains the efficiency score for each firm.

The Second-stage

The variables involved in the first-stage are observed input and output data. To incorporate the environmental effects into the models, factors affecting the technical efficiency are further investigated by Tobit regression analysis in some previous studies (see for example, Oum and Yu, 1994). One shortcoming is that, however, previous studies only consider radial inefficiency and ignore the input and output slacks. Thus, in this research, the input slacks estimated from the first-stage are decomposed into environmental influences, managerial inefficiency and statistical error terms at the third-stage by using the following stochastic cost frontier function model.

$$(3\text{-}34)\quad S_{ni} = f(z_i;\beta) + v_{ni} + u_{ni},\ n = 1,\ldots,N,\ i = 1,\ldots,I,$$

Where, dependent variables S_{ni} are the sum of radial and non-radial input excesses estimated from first-stage, z_i are environmental factors, β are parameters to be estimated, $f(z_i;\beta)$ are deterministic slack frontiers. Consistent with stochastic cost frontier function, v_{ni} is assumed to be a statistical noise and follow normal distribution with zero mean and variance σ_v^2, u_{ni} represents managerial inefficiency which was assumed that the $u_{ni}\sim N^+(\mu, \sigma_u^2)$ and distributed independently with v_{ni}.

The Third-stage

At the third-stage, producers' adjusted inputs are constructed from the estimated results of (3-34) by using

$$(3\text{-}35)\qquad x_{ni}^A = x_{ni} + \left[max_i(z_i\hat{\beta}^n) - z_i\hat{\beta}^n\right] + \left[max_i(\hat{v}_{ni}) - \hat{v}_{ni}\right],\ n = 1,\ldots,N,\ i = 1,\ldots,I,$$

Where, x_{ni}^A and x_{ni} are adjusted and observed input quantities, respectively. The adjustment thus puts all DMUs into a common operating environment and a common state of nature (Fried *et al.*, 2002). DEA-based efficiencies are re-estimated by substituting the adjusted data into the model (3-33), which are incorporated with the environmental and noise effects.

The Fourth-stage

Although we re-run the efficiency measurement model based on the adjusted data, however, there is no guarantee that such measurement can always completely eliminate the slacks, at fifth-stage, we thus employ the Slack-adjusted DEA model proposed by Sueyoshi (1999). The SA-DEA model can be expressed as follows, for more detail, refers Sueyoshi (1999), Sueyoshi *et al*. (1999), Hibiki and Sueyoshi (1999), and Sueyoshi and Goto (2001).

$$\begin{aligned}
& Minimize\ \ \theta - \frac{1}{m+k}\left[\left(\sum_{i=1}^{m}(s_i^- / R_i^-)\right)+\left(\sum_{r=1}^{k}(s_r^+ / R_r^+)\right)\right] \\
& subject\ to\ \ \ -y_{io}+\sum_{j\in J} y_{ij}\lambda_j - s_r^+ = 0, \quad r = 1,\dots,k. \\
(3\text{-}36) \qquad & \qquad \theta\, x_{io} - \sum_{j\in J} x_{ij}\lambda_j - s_i^- = 0, i = 1,\dots,m, \\
& \qquad \sum_{j\in J}\lambda_j = 1, \quad \lambda_j \geq 0, \qquad j = 1,\dots,n, and\ \ \theta\ :\ free \\
& Where\ , R_i^- = \max_j\ x_{ij}(i = 1,\dots,m), and\ \ R_r^+ = \max_j\ y_{ij}(r = 1,\dots,k)
\end{aligned}$$

The model (3-36), proposed by Sueyoshi (1999), counts the slacks in one model. However, the results are likely to be biased if slacks occur in two or more dimensions. In order to measure slacks more precisely, Coelli (1998) suggests using a multi-stage DEA model to avoid the problems inherent in the model. In this research, Coelli's (1998) model is adopted to estimate the efficiency and slacks, then, substitute the efficiencies and slacks into the objective function of the model (3-36) to get the slack-adjusted efficiencies.

3.3.2.2 The effectiveness measurement

In the effectiveness measurement, similar procedures (that is four-stage model) are adopted with exceptions of input and output data selected and model adopted. More specifically, this research employs output-orientation DEA model (measuring the maximum possible proportional expansion in all outputs while all inputs remaining unchanged) by selecting passenger-train-kilometer and freight-train-kilometer as input factors, and passenger-kilometer and ton-kilometer and output variables. The model for estimating service effectiveness then becomes:

$$\begin{aligned}
& Maximize\ _{\phi,\lambda}\phi \\
(3\text{-}37) \qquad & s.t.\ \ -\phi\cdot y_i + Y\cdot\lambda \geq 0 \\
& \qquad x_i - X\cdot\lambda \geq 0 \\
& \qquad \sum \lambda = 1, \lambda \geq 0
\end{aligned}$$

Where Y, X, x_i, y_i and λ are defined as previous; ϕ denotes proportional increase in output, which ranges from one to infinity; $1/\phi$ defines the service effectiveness of firm, which varies between zero and one.

Similar to the procedure of efficiency measurement, at the second-stage, stochastic cost frontier function model (i.e. model 3-34) is applied to find out the potential factors which influencing the amount of slacks, and the amount of slacks which were affected by the factors. At the third-stage, data are adjusted by using model (3-35), and the effectiveness of each DMU is re-estimate by adopting model (3-37) based on the adjusted data. Again, there is no guarantee that such measurement can always completely eliminate the slacks, at fourth-stage, I thus employ the output-orientated Slack-adjusted DEA model, which then becomes (Sueyoshi, 1999)

$$\begin{aligned}
&\mathit{Maximize}\ \ \phi + \frac{1}{m+k}\left[\left(\sum_{i=1}^{m}(s_i^- / R_i^-)\right)+\left(\sum_{r=1}^{k}(s_r^+ / R_r^+)\right)\right] \\
&\mathit{subject\ to}\quad -\phi \cdot y_{io} + \sum_{j\in J} y_{ij}\lambda_j - s_r^+ = 0, \quad r = 1,\dots,k. \\
(3\text{-}38)\qquad &\qquad\qquad x_{io} - \sum_{j\in J} x_{ij}\lambda_j - s_i^- = 0, i = 1,\dots,m, \\
&\qquad\qquad \sum_{j\in J}\lambda_j = 1, \quad \lambda_j \geq 0, \qquad j = 1,\dots,n\ \mathit{and}\ \theta\ :\ \mathit{free}
\end{aligned}$$

Where, $R_i^- = \max_j x_{ij}(i = 1,\dots,m)$, and $R_r^+ = \max_j y_{ij}(r = 1,\dots,k)$

3.3.2.3 The productivity measurement

Conventional FGNZ Method

Following Färe, Grosskopf, Norris and Zhang (1994) (hereinafter, FGNZ) assume that the production technology satisfies constant returns to scale and free disposability of inputs and outputs, the input-based Malmquist Productivity Index (MPI) m_I can be expressed as

$$(3\text{-}39)\quad m_I(y_s, x_s, y_t, x_t) = \left[\frac{d_I^s(y_t, x_t)}{d_I^s(y_s, x_s)} \times \frac{d_I^t(y_t, x_t)}{d_I^t(y_s, x_s)}\right]^{\frac{1}{2}}$$

It should be note that we choose input-oriented MPI since we are measuring the productivity rather than sales force, the objective of distance functions is to seek a minimum proportional reduction in input usage while keeping outputs unchanged. Thus the $d_I^t(y_t, x_t)$ in (3-39), stands for input-oriented distance function between observation (y_t, x_t) of period t and frontier under technology of period t. The MPI m_I can be further decomposed to efficiency change (EC) and technology change (TC) as follows.

$$(3\text{-}40)\quad m_I(y_s, x_s, y_t, x_t) = \frac{d_I^t(y_t, x_t)}{d_I^s(y_s, x_s)} \times \left[\frac{d_I^s(y_t, x_t)}{d_I^t(y_t, x_t)} \times \frac{d_I^s(y_s, x_s)}{d_I^t(y_s, x_s)}\right]^{\frac{1}{2}}$$

Where, the first term in right hand side is defined as EC, which captures the catching-up effect, and the second term is defined as TC, which measures the movement of the frontier. To measure the m_I one needs to calculate four distance functions by using linear programming technique. It is worthy to note that, when solving four LPs, we adopt CCR model (i.e. model 3-1), rather than BBC model. The reason can be found from Färe *et al*. (1997). Four distance function models are as follows.

$$\begin{aligned} &d_I^t(x_t, y_t) = Min_{\theta,\lambda}\theta \\ (3\text{-}41)\quad &s.t. \;\; \text{-}\,y_{it} + Y_t \cdot \lambda \geq 0, \\ &\qquad \theta \cdot x_{it} - X_t \cdot \lambda \geq 0, \quad \lambda \geq 0 \end{aligned}$$

$$\begin{aligned} &d_I^s(x_s, y_s) = Min_{\theta,\lambda}\theta \\ (3\text{-}42)\quad &s.t. \;\; \text{-}\,y_{is} + Y_s \cdot \lambda \geq 0, \\ &\qquad \theta \cdot x_{is} - X_s \cdot \lambda \geq 0, \quad \lambda \geq 0 \end{aligned}$$

$$\begin{aligned} &d_I^s(x_t, y_t) = Min_{\theta,\lambda}\theta \\ (3\text{-}43)\quad &s.t. \;\; \text{-}\,y_{it} + Y_s \cdot \lambda \geq 0, \\ &\qquad \theta \cdot x_{it} - X_s \cdot \lambda \geq 0, \quad \lambda \geq 0 \end{aligned}$$

$$\begin{aligned} &d_I^t(x_s, y_s) = Min_{\theta,\lambda}\theta \\ (3\text{-}44)\quad &s.t. \;\; \text{-}\,y_{is} + Y_t \cdot \lambda \geq 0, \\ &\qquad \theta \cdot x_{is} - X_t \cdot \lambda \geq 0, \quad \lambda \geq 0 \end{aligned}$$

Proposed Four-stage Method

In the productivity measurement model (3-39) and (3-40), the Malmquist Productivity Index is based on four distance functions (3-41)-(3-44), which can be solved by using linear programming technique. However, the solutions of linear programs frequently contain slacks, which are typically ignored. When slacks are presented, radial efficiency measures may overstate the true efficiency thus affects productivity index in an unknown way. In addition, previous studies do not take environmental factors and statistical noise into account when measuring productivity index. To measure MPI more precisely, we thus solve four distance functions by substituting adjusted data obtained from third-stage in efficiency measurement procedure and adopting SA-DEA model (3-36). The model (hereafter, four-stage method) thus takes environmental factors, statistical noise, and residual slacks into account.

3.3.2.4 The sales force measurement

Measuring Sales Force by FGNZ Method

The model (3-40) and (3-41), measure the Malmquist Productivity Index, which is related to productive efficiency measurement. As aforementioned, sales process differs from production process. Therefore, productivity measurement reflects only the capability of production division; on the other hand, sales force usually to be used to measure a firm's capability in sale or a salesperson's productivity. However, the previous studies related to measurement of railways' performance do not measure sales force. In this research, sales force is defined as the sale capability of railways, which is corresponding to service effectiveness. The Malmquist Sales force Index (MSI) thus can be defined as output-based Malmquist Index as follows.

$$(3\text{-}45)\quad m_o(y_s, x_s, y_t, x_t) = \left[\frac{d_o^s(y_t, x_t)}{d_o^s(y_s, x_s)} \times \frac{d_o^t(y_t, x_t)}{d_o^t(y_s, x_s)}\right]^{\frac{1}{2}},$$

$$(3\text{-}46)\quad m_o(y_s, x_s, y_t, x_t) = \frac{d_o^t(y_t, x_t)}{d_o^s(y_s, x_s)} \times \left[\frac{d_o^s(y_t, x_t)}{d_o^t(y_t, x_t)} \times \frac{d_o^s(y_s, x_s)}{d_o^t(y_s, x_s)}\right]^{\frac{1}{2}}$$

Similar to MPI, MSI can be decomposed into effectiveness change (EC) and sales technology change (TC). In model (3-46), the first term of right hand side stands for EC, and the second term represents for TC. To measure MSI, one needs to calculate four distance functions as shown in chapter two by using linear programming technique. Hereinafter, the method proposed by Färe, Grosskopf, Norris and Zhang (1994) is called FGNZ method.

Measuring Sales Force by Proposed Four-stage Method

Similar to productivity measurement, the FGNZ method does not take environmental factors, statistical noises, as well as slacks into account. When slacks are presented, radial effectiveness measures may overstate the true effectiveness thus affects sales force index in an unknown way. In order to measure MSI more precisely, this research thus proposes to solve four distance functions by substituting adjusted data obtained from third-stage in effectiveness measurement and adopting SA-DEA model (3-38). Hereinafter, the method proposed by this research is called four-stage method. Again, the four-stage method accounts for environmental factors and statistical noise, as well as slacks.