ESSAYS ON EFFICIENCY ANALYSIS

A Dissertation

by

NORABAJRA ASAVA-VALLOBH

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2009

Major Subject: Economics

ESSAYS ON EFFICIENCY ANALYSIS

A Dissertation

by

NORABAJRA ASAVA-VALLOBH

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Dennis W. Jansen
Committee Members,	Timothy J. Gronberg
	Qi Li
	Lori L. Taylor
Head of Department,	Larry Oliver

May 2009

Major Subject: Economics

ABSTRACT

Essays on Efficiency Analysis. (May 2009)

Norabajra Asava-vallobh, B.A., Thammasat University;

M.A., Thammasat University

Chair of Advisory Committee: Dr. Dennis W. Jansen

This dissertation consists of four essays which investigate efficiency analysis, especially when non-discretionary inputs exist. A new approach of the multi-stage Data Envelopment Analysis (DEA) for non-discretionary inputs, statistical inference discussions, and applications are provided. In the first essay, I propose a multi-stage DEA model to address the non-discretionary input issue, and provide a simulation analysis that illustrates the implementation and potential advantages of the new approach relative to the leading existing multi-stage models of non-discretionary inputs, such as Ruggiero's 1998 model and Fried, Lovell, Schmidt, and Yaisawarng's 2002 model. Furthermore, the simulation results also suggest that the constant returns to scale assumption seems to be preferred when observations have similar sizes, but variable returns to scale may be more appropriate when their scales are different. In the second essay, I make comments on Simar and Wilson work of 2007. My simulation evidence shows that traditional statistical inference does not underperform the bootstrap process proposed by Simar and Wilson. Moreover, my results also show that the truncated model recommended by Simar and Wilson does not outperform the tobit model in terms of statistical inference. Therefore, the traditional method, t-test, and the tobit model should continue to be considered applicable tools for a multi-stage DEA model with non-discretionary inputs, despite contrary claims by Simar and Wilson. The third essay raises an example of applying my new approach to data from Texas school districts. The results suggest that a lagged variable (e.g. students' performance in the previous year), a variable which has been used in the literature, may not play an important role in determining efficiency scores. This implies that one may not need access to panel data on individual scores to study school efficiency. My final essay applies a standard DEA model and the Malmquist productivity index to commercial banks in Thailand in order to compare their efficiency and productivity before and after Thailand's Financial Sector Master Plan (FSMP) that was implemented in 2004.

This dissertation is dedicated to the people who have truly believed in me over the years:

Surakarnchana, Putpannee, and Vongsvijya Asava-vallobh and Pawinee Chitmongkolsamur.

ACKNOWLEDGEMENTS

I am deeply indebted to Dr. Dennis Jansen for his valuable advice on conducting this entire dissertation and for his precious time spent with me over the years. I also would like to thank Dr. Timothy Gronberg for his kind help in shaping Chapter II, Dr. Qi Li for his indispensible guidance on simulation designs in Chapter III, and Dr. Lori Taylor for her comments in doing empirical research in Chapter IV and also in large parts of the dissertation. Most importantly, I am grateful to the Fiscal Policy Office, Ministry of Finance, Thailand for their financial support without which I could not have come to this prestigious university in the first place.

To my colleagues, who have been with me since the beginning, and without whom I would never have made it so far, many thanks to Yongok Choi, Hyosung Yeo, Bongju Song and every wonderful friend I cannot mention all here. Last but definitely not least, thanks to my wife Pawinee for her patience and love.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

TABLE Page

CHAPTER I

INTRODUCTION

Data Envelopment Analysis (DEA) has been a standard tool in the area of efficiency and productivity analysis. A number of literatures apply the DEA technique to empirical studies in a variety of fields, such as bank, insurance, business, industrial production, education, health service, public service, etc. When it comes to a situation where some inputs cannot be controlled by a manager or a firm, the standard DEA technique would need to be modified. The non-controllable inputs (environmental factors or non-discretionary inputs) have direct impacts on firms' performance, but since firms have to take these factors as given we shouldn't let them affect the measurement of firms' efficiency. For example, school's performance could be affected by students' characteristics which schools cannot control, such as students' family income, students' illness, etc. Recently, there have been many economists proposing several multi-stage DEA approaches to deal with such a situation. However, practitioners may find it hard to choose the most appropriate approach to their problems. Quite a few literatures only provide basic information regarding strengths and weaknesses of each approach. Therefore, this dissertation is intended not only to provide more information about the performance of the existing approaches but also introduce a new approach in order to be an alternative choice for practitioners.

Chapter II is the most important chapter in this dissertation. It reviews most of the leading existing multi-stage DEA approaches for non-discretionary inputs, and identifies potential

---

This dissertation follows the style of Journal of Econometrics.

problems that might occur with the models. Then, it proposes a new approach that could solve the problems; and finally, compare performances of each approach using Monte Carlo simulations.

Chapter III discusses a specific topic in statistical inference. Simar and Wilson (2007) states that a traditional approach (t-test) does not suit multi-stage DEA models for non-discretionary inputs. They propose a new bootstrap method to cope with the problem. However, it has some features that are potentially problematic. Therefore, I design two sets of simulations to re-compare performances of the new approach to see whether their recommended tools are superior.

Since Chapter II introduces a new approach, Chapter IV applies it to Texas school districts as an application. Similarly, Chapter V is also an empirical application but only employs a standard DEA technique and the Malmquist productivity index to analyze banks' performance in Thailand.

CHAPTER II

AN ALTERNATIVE DEA METHODOLOGY FOR NON-DISCRETIONARY INPUTS

2.1 Introduction

DEA is a useful nonparametric modeling approach for estimating technical efficiency. The DEA technique has been applied to analyzing firm efficiency across a variety of economic sectors, including education, health, and financial services. Traditional input-oriented DEA efficiency scores calculate the proportional reduction in inputs that is possible while maintaining observed output levels. In many applications the nature of the possible inefficiency is an important issue, and there is a desire to separate the component of inefficiency that is under the control of management from the component that is outside management's control, at least in the time span considered in the analysis. To address such issues the standard DEA model has been adapted to deal with so-called 'non-controllable' or environmental factors. For example, education production depends not only upon the levels of discretionary inputs like teacher labor but also upon non-discretionary factors such as family inputs. When some inputs are uncontrollable, the operational question of interest is often whether a proportional reduction in controllable inputs is possible, within a given environment, while maintaining observed output levels.

I illustrate the basic environmental variables problem in Figure 2.1, with six efficient firms plotted in output-input space. Firms A, B, and C are in environment z_0, while firms D, E, and

F are in environment z_1. Here z_1 is the more favorable environment. Ignoring the environment would lead to the incorrect conclusion that firms A, B, and C are inefficient.

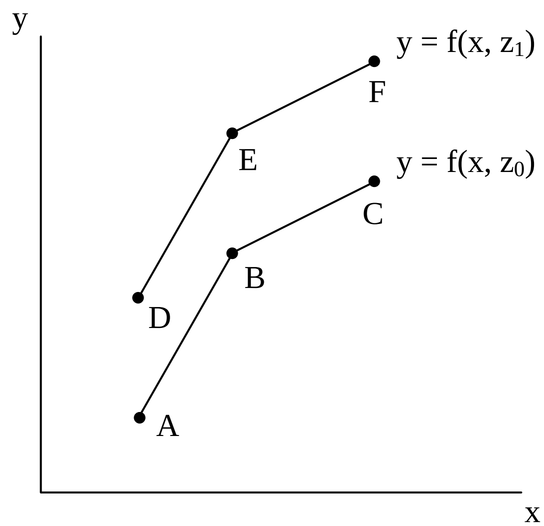


Figure 2.1 Efficient firms in different environments

2.2 Alternative Approaches with Environmental Effects

1) Standard DEA treats environmental variables (z) symmetrically with ordinary inputs (x). Assume J firms, inputs $x = (x_1,\dots,x_m)$, outputs $y = (y_1,\dots,y_n)$, and environmental variables $z = (z_1,\dots,z_r)$. The input-oriented DEA score of each firm can be calculated as:

$$
\begin{aligned}
&\min_{\lambda,\theta} \quad \theta_j && j = 1,\dots,J \\
&\text{s.t.} \quad \sum_{h=1}^{J} \lambda_h y_{kh} \geq y_{kj} && k = 1,\dots,n \\
&\qquad \sum_{h=1}^{J} \lambda_h x_{ih} \leq \theta_j x_{ij} && i = 1,\dots,m \\
&\qquad \sum_{h=1}^{J} \lambda_h z_{lh} \leq \theta_j z_{lj} && l = 1,\dots,r \\
&\qquad \lambda_h \geq 0 && h = 1,\dots,J \\
&\qquad \sum_{h=1}^{J} \lambda_h = 1 \quad \text{if} \quad \text{VRS}
\end{aligned}
\tag{2.1}
$$

Here $\{\lambda_h\}$ indicate linear combinations of firms to compare to the representative firm j and θ_j indicates radial efficiency, the minimum distance to the frontier. If the summation of lambdas is not restricted to one, it means the calculation assumes constant returns to scale (CRS). However, if it is restricted to one, variable returns to scale (VRS) are assumed.

Banker and Morey (1986) were the first to modify the standard model so that it could more properly deal with non-discretionary inputs. Their modification was to drop θ_j out of the third constraint in equation (2.1). Therefore, their model disallows non-discretionary input reduction, but the non-discretionary inputs still need to have a convexity property.

2) Ruggiero (1998) relaxed the convexity assumption for non-discretionary inputs and proposed a 3-stage model. The first stage calculates standard DEA ignoring environmental variables. Then the DEA scores (contaminated by environmental effects) from the first stage are regressed against all environmental variables in the second stage.[1]

$$DEA_{1st,j} = \alpha + \beta_1 z_{1j} + , \dots , + \beta_r z_{rj} + e_j \tag{2.2}$$

From equation (2.2) he creates an index Z_j representing the impact of environmental effects on the individual firm where $Z_j = \beta_1 z_{1j} + , \dots , + \beta_r z_{rj}$

Ruggiero's third stage DEA calculates the weights $\{\lambda_h\}$ conditioned on the index (Z_j), so that any firm having a more favorable environment than that of firm j will not be included in the frontier. The third stage problem is:

$$\begin{aligned}
&\underset{\lambda,\theta}{\text{Min}} \quad \theta_j && j = 1,\dots,J \\
&\text{s.t.} \quad \sum_{h=1}^{J} \lambda_h y_{kh} \geq y_{kj} && k = 1,\dots,n \\
&\qquad \sum_{h=1}^{J} \lambda_h x_{ih} \leq \theta_j x_{ij} && i = 1,\dots,m \\
&\qquad \lambda_h = 0 \quad \text{if} \quad Z_h > Z_j && \\
&\qquad \lambda_h \geq 0 && h = 1,\dots,J \\
&\qquad \sum_{h=1}^{J} \lambda_j = 1 \quad \text{if} \quad \text{VRS} &&
\end{aligned} \tag{2.3}$$

[1] The original model used OLS to create the index Z_j, but Ruggiero noted that researchers could also adopt other techniques, such as tobit, etc.

3) Fried et al. (2002) proposed a multi-stage approach using both DEA and stochastic frontier techniques. The first stage is to calculate the DEA score in the same fashion as in Ruggiero's which ignores all environmental inputs. Then calculate total input slacks[2] from the first stage's information as in equation (2.4):

$$\text{total input slack}_{ij} = x_{ij} - \sum_{h=1}^{J} \lambda_h x_{ih} \tag{2.4}$$

In stage two, they find a relationship between the slacks and all environmental inputs by using the stochastic frontier technique. Equation (2.5), the stochastic frontier regression, is estimated once for each slack to extract the impact of environmental variables on each discretionary input:

$$\text{total input slack}_{ij} = \beta_0 + \beta_1 z_{1j} + \ldots + \beta_r z_{rj} + v_{ij} + u_{ij}, \qquad i = 1,\ldots,m \tag{2.5}$$

where $v_{ij} \sim N(0, \sigma_{vi}^2)$ and $u_{ij} \sim |N(0,\sigma_{ui}^2)|$

Equation (2.6) illustrates how to use the information from the stochastic frontier model to adjust all discretionary inputs so that all firms have the same least favorable environmental effects:[3]

$$x_{ij,adj} = x_{ij} + \left[\max_j (z_j \hat{\beta}^i) - z_j \hat{\beta}^i\right] + \left[\max_j (\hat{v}_{ij}) - \hat{v}_{ij}\right], \qquad i = 1,\ldots,m, \quad j = 1,\ldots,J \tag{2.6}$$

---

[2] Total slack is composed of a radial part derived from the DEA score and a non-radial part which is a difference in a constraint's inequality, if it exists.

[3] The reason that Fried *et al*. (2002) chose the least favorable environment is that the adjustment only adds some positive number to the input. Therefore, the adjusted x always remains positive. Suppose we put this positive/negative issue aside; then theoretically, we should be able to level the playing field at any state of the environment, e.g. the best environment, the worst environment, or somewhere in the middle.

where $\hat{\beta}^i$ is a parameter vector for discretionary input i

Please note that output slacks are supposed to contain an additional piece of information on top of the input slacks. However, it is common in input-oriented problems that most output slacks are zero. Therefore, Fried *et al*. opt to ignore the output slacks.[4]

Finally, the standard DEA is applied to the above modified data $x_{i,adj}$ and outputs y_{kj} in order to calculate each firm's efficiency score

4) Muñiz (2002) introduced a pure multi-stage DEA which is a variant of the Fried *et al*. (2002) model. After calculating the DEA score ignoring all environmental variables in the first stage, Muñiz prefers to utilize the information from both input slack and output slack[5]

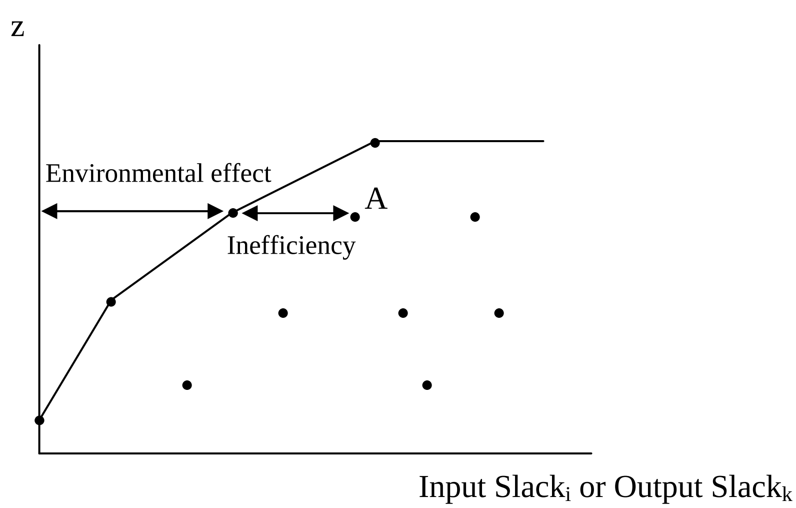


Figure 2.2 Environmental effect

[4] Fried *et al*. (1999) stated that "An input oriented model takes output as given and measures inefficiency by the potential reduction in inputs. Output surplus exists in empirical applications because the data set is sparse for some output vectors. Where it does exist, it is likely to be composed mostly of zeros and have insufficient variation to be useful in the estimation."
[5] The input slack means the amount of input used in excess, while the output slack means the amount of output insufficiently produced.

in the second stage. The idea of the second stage is to estimate the relationship between each slack and all environmental inputs as in Fried *et al*.'s. The difference is that Muñiz uses the DEA technique as shown in equation (2.7) instead of the econometric technique. Figure 2.2 illustrates how the DEA technique works. The horizontal distance from vertical axis to the frontier at each level of z represents an environmental effect due to a particular input i. Therefore, all firms sharing the same level of z will have the same environmental impact on the input i:

$$\begin{aligned}
&\min_{\lambda,\theta} \quad \beta_A \\
&\text{s.t.} \quad \sum_{j=1}^{J} \lambda_j z_{lj} \geq z_{lA} && l = 1,\ldots,r \\
&\qquad \sum_{j=1}^{J} \lambda_j \text{input slack}_{ij} \leq \beta_A^I \text{input slack}_{iA}\ * && \qquad (2.7) \\
&\qquad \lambda_j \geq 0 && j = 1,\ldots,J \\
&\qquad \sum_{j=1}^{J} \lambda_j = 1 && \text{if} \quad \text{VRS}
\end{aligned}$$

* If output slack is considered, we can just simply resolve equation (2.7) where replacing

$\sum_{j=1}^{J} \lambda_j \text{input slack}_{ij} \leq \beta_A^I \text{input slack}_{iA}$ with $\sum_{j=1}^{J} \lambda_j \text{output slack}_{kj} \leq \beta_A^O \text{output slack}_{kA}$.

By the same reasoning as Fried *et al*., Muñiz not only adjusts discretionary inputs but also outputs using equation (2.8). This is to level the playing field where all firms operate under the most favorable environment.

$$x_{i,adj} = x_i - \beta^I \cdot \text{Input Slack}_i \qquad i = 1, \ldots, m \qquad (2.8.1)$$

$$y_{k,adj} = y_k + \text{Output Slack}_k \qquad k = 1, ..., n \qquad (2.8.2)$$

where β is a DEA score calculated from stage two.[6]

Finally, standard DEA is applied to the above modified data $x_{i,adj}$ and $y_{k,adj}$ to calculate each firm's efficiency score.

2.3 Potential Problems with the Above Models

The standard efficiency model inappropriately treats environmental factors as discretionary inputs. As a result, interpretation of the DEA score in terms of how much *both* types of inputs could be equiproportionally reduced contradicts with the fact that firms can only control discretionary inputs. Even for the Banker and Morey model it has been repeatedly shown in the literature that their model doesn't work as well as others, such as the Ruggiero and the Muñiz models.[7]

The Ruggiero model calculates the weights $\{\lambda_j\}$ conditioned on the index (Z_j), so that all firms having a more favorable environment can never be benchmarks for any firm with a

[6] Equation 2.8.2 is quoted from Muñiz *et al*. (2006), but it could be a typo that the second-stage score as a coefficient of output slack$_k$ is missing. Otherwise, there would be no point to resolving equation (2.7) using output slacks instead of input slacks. However, I have tried performing a variety of Monte Carlo simulations in the same spirit as in the following section, and found that including output slack either with or without the second-stage score as a coefficient of output slack$_k$ (equation 2.8.2) does not enhance the model's performance. This is because most of output slacks are zero which is consistent with a statement in Fried *et al.* (1999), (see footnote 4). Therefore, for the rest of this dissertation, I opt to calculate the Muñiz model under the condition that output slacks are ignored.

[7] See simulation results from Muñiz *et al*. (2006) and Ruggiero J. (2007)

less favorable environment. This may lead to upwardly-biased DEA scores. In other words, individual firms' performance may be alleged to be better than what they actually are. Besides, the bias tends to be more pronounced for lower numbers of observations.

Unlike the Ruggiero model, the Muñiz and the Fried *et al*. models create new set of benchmarks by incorporating all information. Therefore, the upward DEA score bias, if it exists, should not be as serious as in the Ruggiero model. Nevertheless, due to the resemblance of the Muñiz and the Fried *et al*. models, they could still share a problem. Both models intend to adjust the amount of data so that it seems like they are in the same environment. However, these two models may not take good care of the difference in firms' scale.

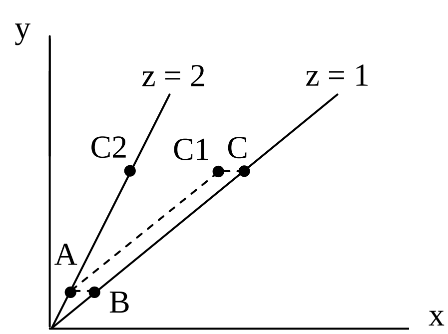


Figure 2.3 Relocating problem

For example, Figure 2.3 shows how the Muñiz model works. Suppose there are three efficient firms, A B and C, but firm A operates in a better environment. Ideally, the adjustment process in the second stage should be able to move firm B and C to the outer frontier, so that every firm lying on the frontier is considered as efficient. Although the

Muñiz model manages to move firm B to the same position as firm A, firm C, sharing the same environment as firm B, can only move as far as C1 instead of to C2. Therefore firm C is labeled as inefficient. Moreover, applying the VRS version of DEA may amplify this problem (i.e., the frontier's curved shape widens the distance between the outer frontier and the location of the firm C1). Slightly different from the Muñiz model, Fried *et al*. employ a stochastic frontier model in the second stage, so there are two components to the adjustment process as reflected in equation (2.6). The deterministic component, $\max_j(z_j\hat{\beta}^i) - z_j\hat{\beta}^i$, would treat all firms having the same level of environment equally regardless of their size. Therefore, this latter model would possibly share the same issue as depicted in Figure 2.3. Furthermore, the stochastic component, $\max_j(\hat{v}_{ij}) - \hat{v}_{ij}$, could treat the firms' scale issue improperly as well, because the largest v_{ij} is likely to be derived from a large firm. Consequently, especially for the small firms, the difference between the max (v_{ij}) and an individual firm's v_{ij} could be oversized.

2.4 An Alternative Model with Non-discretionary Inputs

The alternative model still preserves the advantage of the Muñiz and the Fried *et al*. models by utilizing all observations to avoid the type of upward bias as in the Ruggiero model. However, I try to handle the scale issue via use of a modification procedure described as follows. The first stage again calculates the DEA score (θ_j) while ignoring all environmental variables. In the second stage, the relationship between all environmental variables and (radial) efficiency scores[8] is estimated in order to extract impacts of environmental factors to the use of inputs. However, instead of using total input slacks measured in levels as in the Muñiz's and Fried *et al*.'s models, I give up the non-radial part and employ only the radial part, θ_j, which is measured in ratios in order to help relieve the scale problem. Because the first-stage efficiency estimates (θ_j) indicate the calculated equiproportional reduction in all inputs that are possible to achieve the efficiency frontier, use of theta expresses how much the impacts of environmental factors affect on all discretionary inputs measured as a percentage. Then a level playing field could be created differently from the Muñiz and Fried *et al*. models as it reflects proportional adjustment. The impact of environmental factors for each firm would be represented by a distance from its level of theta to the efficient level of theta given a certain environment. The third and final step consists of applying equation (2.9) to obtain a final DEA score for each firm.

$$DEA_j = \frac{\theta_j}{\theta_j + \text{distance}} \tag{2.9}$$

Note these firms' final DEA scores will be one (indicating full relative efficiency) if the distance is zero. Otherwise, the score is reduced by the percentage explained by environmental impacts in the second stage, the distance. With the adjustment in equation (2.9), the final (input-oriented) DEA score would always range from zero to one.

One of the advantages of DEA approach over other methods is that DEA can easily handle multiple left-hand side and right-hand side variable problems while a problem solved by

[8] Information from efficiency score (θ) is equivalent to information from radial input slack because radial input slack / efficient level of input = 1 - θ. However, total input slack, as used in the Muñiz model and the Fried *et al*. model, contains more information as total input slack = radial input slack + non-radial input slack. Yet, non-radial input slack is mostly either insignificant or even zero in many cases.

parametric approaches is restricted to a single left-hand side variable problem only. However, due to the use of theta which is a single variable, both DEA and any parametric approach (e.g. deterministic frontier model (DF), stochastic frontier model (SF), tobit model (TO), and truncated model (TR)) could be candidates for estimating the relationship between slack measure and environmental factors in stage two. According to Ruggiero (1999) and Jensen (2005), their simulation studies show that the deterministic frontier model seems to be more attractive than the stochastic frontier model. Simar and Wilson (2007) recommended using the truncated model while most of the literatures employs the tobit model. However, I will employ all options and perform simulation exercises comparing results to other existing approaches as will be described in the next section.

Some intuition is provided by the following Figure 2.4. Suppose firms A and B have the same level of environmental effect. Firm B is assumed to be efficient but firm A is not. In the first stage they both lie inside the frontier constructed by other efficient firms (e.g., C) having a more favorable environment. If DEA is chosen in the second stage, the frontier would exist only if it is generated in the z-(1-θ) space where every element of z is assumed to be unfavorable to production.[9]

[9] If z_l is favorable, we could replace it with a series of $\mathrm{Max}_j\,(z_{lj}) - z_{lj}$ before solving the DEA optimization problem in this stage.

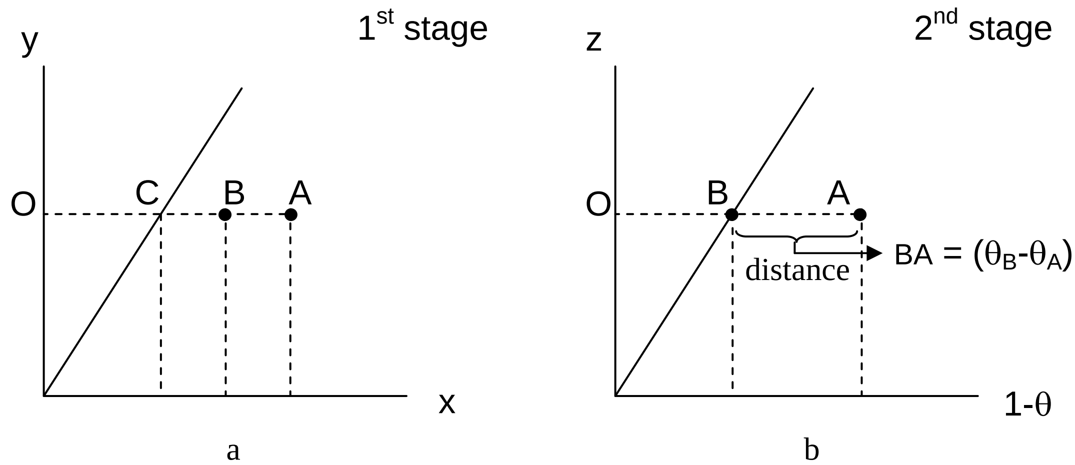


Figure 2.4 A new multi-stage DEA model

To calculate a DEA score for firm A, firm B serves as a benchmark. Therefore, the goal is to find the ratio OB/OA from Figure 2.4a.

$$DEA_A = \frac{OB}{OA} = \frac{\theta_A}{\theta_B} = \frac{\theta_A}{\theta_A + (\theta_B - \theta_A)} = \frac{\theta_A}{\theta_A + \text{distance}} \qquad (2.10)$$

where distance = $(1-\beta_A)(1-\theta_A)$ if DEA is chosen, and

β_A is a DEA score for firm A in the second stage

Generally, no matter what technique is employed in the second stage, it can generate a frontier line used as a benchmark. Therefore, a benchmark B can always be approximated even if firm B is not observed. Compared to the Ruggiero method, which always uses the firm having less favorable environment as a benchmark, the DEA score from my model is based on a larger set of potential comparators and, hence, should be more precise.

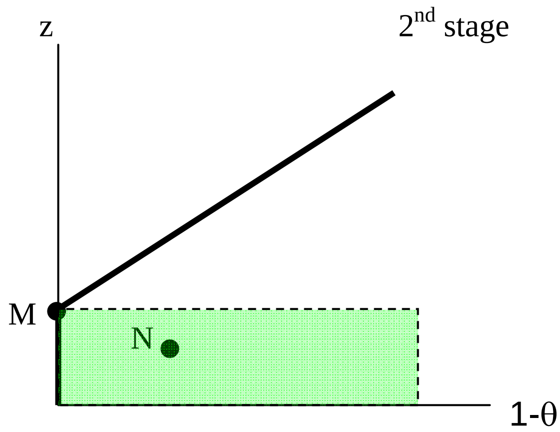


Figure 2.5 Using DEA approach in the second stage

However, when DEA is chosen, a potential problem could surface when there exists a firm, such as M in Figure 2.5, which is efficient in the first stage. If a CRS frontier is assumed in the second stage, the frontier will become just a straight vertical line making the second stage pointless. Therefore, VRS restriction must be assumed. However, the problem remains as the frontier below point M is still vertical. Consequently, all firms locating inside the shaded area will not have appropriate benchmarks for comparison.

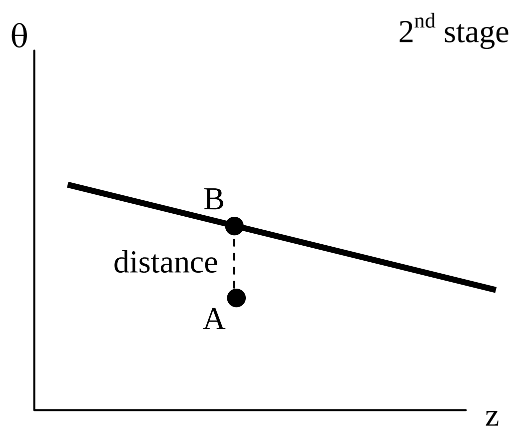


Figure 2.6 Vertical distance in the second stage

As discussed above, DEA is not the only option to use in the second stage. This vertical frontier problem could be solved by using parametric methods.

2.4.1 Parametric Method in the Second Stage

The aim is to find the distance BA in term of theta as in Figure 2.4 or Figure 2.6. Because the frontier is not restricted to have a convex positive slope as in the DEA approach, I can simply regress the first stage efficiency estimates (θ_j) on environmental variables (z) without any presumption of the environmental effect's direction. If the frontier has a negative (positive) slope, it implies that the environmental variable is unfavorable (favorable) to production.

1) The deterministic frontier model assumes there is no noise in the model and presumes a one-side error term to represent distance from an observation to the frontier. The well-known technique is the corrected OLS (COLS) method introduced by Winsten (1957). First, a simple OLS method is employed to find a relationship between theta and z as in equation (2.11):

$$\theta_j = \alpha' + \sum_{l=1}^{r} \beta_l z_{jl} + u'_j \tag{2.11}$$

Then, the constant term and the error term will be corrected as in equation (2.12) and (2.13), so that the estimation line is shifted up to cover data cloud from above and becomes the deterministic frontier.

$$\alpha = \hat{\alpha}' + \max_{j=1,\ldots,J}(\hat{u}'_j) \tag{2.12}$$

$$u_j = \left| \hat{u}'_j - \max_{j=1,\ldots,J}(\hat{u}'_j) \right| \tag{2.13}$$

$$\theta_j = \alpha + \sum_{l=1}^{r} \beta_l z_{jl} - u_j, \qquad u_j \geq 0 \tag{2.14}$$

Equation (2.14) is the deterministic frontier model with the one-side error term (u_j) representing the distance desired.

2) The stochastic frontier model was first proposed by Aigner *et al*. (1977). Then, Meeusen and van den Broeck (1977) added a noise term (v_j) to the model as in equation (2.15). Therefore, the frontier is stochastic depending on the value of v_j.

$$\theta_j = \alpha + \sum_{l=1}^{r} \beta_l z_{jl} + v_j - u_j, \text{where } v_j \sim N(0, \sigma_v^2) \text{ and } u_j \sim \left| N(0, \sigma_u^2) \right| \tag{2.15}$$

Parameters can be estimated by the maximum likelihood method and the log-likelihood function as follows:

$$l(\alpha, \beta, \sigma, \lambda) = -n \ln(\sigma) + \frac{n}{2} \ln(\frac{2}{\pi}) + \sum_{j=1}^{J} \left[\ln \Phi(\frac{-e_j \lambda}{\sigma}) - \frac{1}{2} (\frac{e_j}{\sigma})^2 \right] \tag{2.16}$$

where $e_j = v_j - u_j, \quad \lambda = \frac{\sigma_u}{\sigma_v}, \quad \sigma^2 = \sigma_v{}^2 + \sigma_u{}^2,$

$\Phi(.)$ and $\phi(.)$ represent the standard normal cumulative distribution and density function.

However, my aim is to extract the distance (u_j) from e_j. Jondrow *et al*. (1982) proposed equation (2.17) to estimate the value of u_j as follows:

$$E(u_j | e_j) = \frac{\sigma\lambda}{1+\lambda^2}\left[\frac{\phi(e_j\lambda/\sigma)}{\Phi(-e_j\lambda/\sigma)} - \frac{e_j\lambda}{\sigma}\right] \tag{2.17}$$

3) The tobit model is probably the most widely used model in the DEA literature due to the fact that the DEA score is censored at one. The tobit model utilizes all observations including both censored data (efficient firms) and ordinary data (inefficient firms). It assumes that both error term and probability to observe ordinary data are based on the normal distribution. The log-likelihood function is as follows:

$$l(\alpha,\beta,\sigma) = \sum_{j|\theta_j<1} \ln\phi\left(\frac{\theta_j - \alpha - \sum_{l=1}^{r}\beta_l z_{jl}}{\sigma}\right) + \sum_{j|\theta_j=1} \ln\left(1 - \Phi\left(\frac{1 - \alpha - \sum_{l=1}^{r}\beta_l z_{jl}}{\sigma}\right)\right) \tag{2.18}$$

Equation (2.18) represents a right-censored model at one. Since the tobit model is linear as in equation (2.11), the distance would be calculated by the same method as the deterministic model using equation (2.13).

Recently, McDonald (2009) argued that the tobit model is not appropriate in second stage DEA analyses. He suggested OLS instead of tobit because the first stage DEA score is fractional data rather than censored data. It is true that DEA scores range from zero to one and therefore seem to be fractional data. However, it is not always true. If we assume that efficient firms are barely observed, one might observe a kinked frontier in the second stage as illustrated in Figure 2.7.

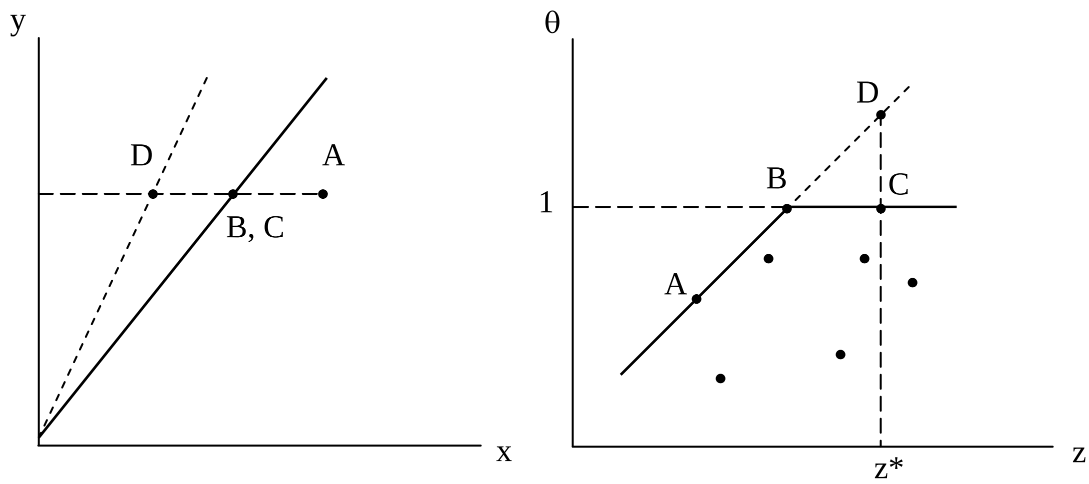


Figure 2.7 Empirical data and kinked frontier

Practically, one is likely to observe more than one best-practice firm (unit score firms), like firms B and C, lying on the horizontal part of the frontier (Figure 2.7, right panel) especially when z is a continuous variable. Firms A and B form an upward slope frontier implying that the environmental variable is favorable to the production. It would make a little sense if we admit that firm C, being in a more favorable environment, has the same efficiency score as firm B which locates in a less favorable environment level, and both of them lie on the frontier. It implies that we do not observe an efficient firm in environment z*, firm D. However, the left panel of Figure 2.7 helps explain that if firm D is observed, it would be the only firm that obtains a unit score in the first stage and helps get rid of the illogically horizontal part of the frontier. In this sense, firm C is censored because we observe its z level but θ_C is bounded at one. Therefore, the tobit model should be a fine tool in this context.

4) The truncated model was recently recommended by Simar and Wilson (2007).[10] One of the assumptions they make prior to estimation is that the probability to observe efficient firms is zero, so any observed efficient firm is spurious and should be removed from the sample set. As a result, the truncated model is the only appropriate model that corresponds to this assumption. Equation (2.19) shows the truncated model's log-likelihood function where the truncation points are at zero and one. Again, suppose equation (2.11) is a linear truncated model. The distance can be determined by equation (2.13).

$$l(\alpha,\beta,\sigma) = \sum_{j|\theta_j<1} \ln\phi\left(\frac{\theta_j - \alpha - \sum_{l=1}^{r}\beta_l z_{jl}}{\sigma}\right) - \ln\Phi\left(\frac{1-\alpha-\sum_{l=1}^{r}\beta_l z_{jl}}{\sigma}\right) - \ln\Phi\left(\frac{-\alpha-\sum_{l=1}^{r}\beta_l z_{jl}}{\sigma}\right) \quad (2.19)$$

2.5 Simulation Evidence

I illustrate the potential advantages of my approach through a simulation analysis. In particular, I compare the performance of my approaches to four existing models: (1) a standard one-stage DEA that treats environmental and discretionary inputs symmetrically; (2) Muñiz's (2002) three-stage model; (3) Fried *et al*. three-stage model: and (4) Ruggiero's (1998) three-stage model. Muñiz *et al*. (2006) and Ruggiero (1998) present simulation evidence that indicates Ruggiero's model and Muñiz's model are the first two best performers of a large set of alternatives, so I focus on comparing my model to theirs. I also focus on the case of multiple outputs (and multiple inputs), since DEA is a particularly

[10] See Chapter III for more discussion about the use of tobit and truncated models in multi-stage DEA analysis.

attractive methodology for assessing production efficiency in such environments. In order to compare all models in different aspects, I set up many experiments as described below.

2.5.1 First case: Base case

Similar to Muñiz *et al*. (2006), each x and z are generated independently from the following distributions: $x_i \sim \text{Uniform}(30,50)\ \forall i$, $z_l \sim \text{Uniform}(1,2)\ \forall l$. Efficiency is generated as $\gamma = e^{-|u|}$, where $u \sim N(0, 0.09)$. After calculating y from the production function and the random draws of x and z, the observed x, x*, is scaled by $1/\gamma$, so that the observed x is inefficiently large relative to the observed output. In my simulation, analysts are assumed to observe the vectors y, x*, and z. Following Jensen (2005), all calculations were performed with 200 replications for each experiment.

Unlike Muñiz *et al*. (2006), I do not impose the requirement that there is a set of fully efficient firms in the simulation. Compared to the stochastic frontier model the probability to observe efficient firms (u equal to zero) is close to zero.[11] Besides, it would be more challenging for all approaches to be compared under such a circumstance. In addition, there are quite a few studies in the literature that incorporate multiple outputs into the production function. For instance, multiple outputs in Simar and Wilson (2007) are perfect substitutes (linear relationship). To my knowledge, this dissertation is the first work that employs multiple outputs where they are imperfect substitutes (nonlinear relationship). When there is more than one output in equation (2.20), each y for each firm is uniformly generated so that

[11] By similar reasoning, Simar and Wilson (2007) also assume that we do not observe any efficient firm.

their Euclidean distance is equivalent to the right hand side of the Cobb Douglas production function. As shown in all tables, I name the models by the number of outputs, discretionary inputs, and non-discretionary inputs respectively. The total number of variables is six for all models where the specifications are as follows:

$$\begin{aligned}
&\text{Model 222:} \quad (y_1^2 + y_2^2)^{.5} = x_1^{.3}x_2^{.3}z_1^{.2}z_2^{.2}\\
&\text{Model 114:} \quad y = x_1^{.6}z_1^{.1}z_2^{.1}z_3^{.1}z_4^{.1} \qquad (2.20)\\
&\text{Model 141:} \quad y = x_1^{.15}x_2^{.15}x_3^{.15}x_4^{.15}z_1^{.4}\\
&\text{Model 411:} \quad (y_1^2 + y_2^2 + y_3^2 + y_4^2)^{.5} = x_1^{.6}z_1^{.4}
\end{aligned}$$

Note that these specifications are repeatedly employed in Tables on pages 24-32.

The two criteria I use to assess model performance are those used by Muñiz *et al*. (2006) and by Ruggiero (1998), the Spearman rank correlation of DEA efficiency scores, and the MAD (mean absolute deviation) of estimated DEA scores from true values. The ideal model would be the one having the highest rank correlation and the lowest MAD. However, the rank correlation is my top priority since ranking might be of greatest interest to those who want to compare individual firms' performance.

In Table 2.1, though every approach is calculated by assuming a constant returns to scale frontier in the first stage,[12] the second stage of Muñiz's model and my model with the DEA approach would assume variable returns to scale to avoid any technical problem that might occur.[13] Table 2.2 assumes variable returns to scale in the first and third stage, if they exist.

---

[12] All models that need DEA in the third stage, e.g. the Ruggiero model, the Fried *et al*. model, and the Muñiz model, are assumed to have the same type of returns to scale as in the first stage.

[13] If a constant-returns-to-scale DEA is employed in the second stage, it could result in a vertical frontier, causing the second stage to collapse. Therefore, we always assume variable returns to scale in the second stage as illustrated in Figure 5.

Table 2.1
Simulation results: base case (constant returns to scale)

Number of firm		50		100		150		500	
Model*	Approach	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD
222	Standard	0.6293	0.1705	0.6944	0.1582	0.7217	0.1519	0.7924	0.1331
	Muñiz	0.7606	0.0759	0.8132	0.0587	0.8377	0.0530	0.8845	0.0492
	Fried et al.	0.8550	0.0608	0.8707	0.0523	0.8844	0.0489	0.9247	**0.0393**
	Ruggiero	0.7334	0.1087	0.8135	0.0864	0.8387	0.0755	0.8941	0.0499
	New (DEA)	0.7636	0.0756	0.8149	0.0585	0.8391	0.0525	0.8866	0.0472
	New (TR)	0.8448	0.0577	0.8855	0.0543	0.9021	0.0541	0.9253	0.0633
	New (TO)	0.8615	0.0518	**0.8949**	0.0519	**0.9071**	0.0524	0.9260	0.0626
	New (DF)	**0.8649**	**0.0501**	0.8938	**0.0500**	0.9059	0.0509	0.9250	0.0618
	New (SF)	0.8528	0.0610	0.8780	0.0515	0.8981	**0.0469**	**0.9276**	0.0401
114	Standard	0.5961	0.1587	0.6444	0.1474	0.6695	0.1410	0.7358	0.1255
	Muñiz	0.5501	0.0925	0.6469	0.0775	0.6856	0.0734	0.7939	0.0689
	Fried et al.	0.8802	**0.0469**	**0.9185**	**0.0406**	**0.9300**	**0.0390**	**0.9402**	**0.0348**
	Ruggiero	0.7761	0.0616	0.8460	0.0543	0.8747	0.0533	0.9099	0.0568
	New (DEA)	0.5570	0.0935	0.6516	0.0756	0.6903	0.0691	0.7965	0.0581
	New (TR)	0.8863	0.0693	0.9107	0.0698	0.9191	0.0716	0.9276	0.0774
	New (TO)	0.8955	0.0600	0.9135	0.0651	0.9205	0.0686	0.9275	0.0769
	New (DF)	**0.8960**	0.0605	0.9135	0.0655	0.9205	0.0688	0.9274	0.0770
	New (SF)	0.8817	0.0738	0.9112	0.0626	0.9210	0.0556	0.9288	0.0472
141	Standard	0.8498	0.1384	0.8767	0.1278	0.8941	0.1215	0.9147	0.1085
	Muñiz	0.8510	0.0584	0.8819	0.0476	0.8989	0.0435	0.9207	0.0434
	Fried et al.	0.8769	0.0619	0.9052	0.0514	0.9152	0.0477	0.9420	0.0384
	Ruggiero	0.8429	0.1067	0.8772	0.0888	0.9026	0.0787	0.9339	0.0529
	New (DEA)	0.8511	0.0573	0.8831	0.0466	0.9008	0.0423	0.9266	0.0394
	New (TR)	0.8986	0.0456	0.9239	0.0421	0.9322	0.0433	0.9485	0.0499
	New (TO)	**0.9114**	0.0422	**0.9298**	0.0412	**0.9361**	0.0427	0.9483	0.0497
	New (DF)	0.9102	**0.0415**	0.9268	**0.0404**	0.9330	0.0422	0.9465	0.0496
	New (SF)	0.8841	0.0539	0.9192	0.0426	0.9328	**0.0377**	**0.9522**	**0.0320**
411	Standard	0.6197	0.1761	0.7135	0.1635	0.7504	0.1558	0.8284	0.1351
	Muñiz	0.7751	0.0816	0.8073	0.0664	0.8242	0.0614	0.8533	0.0609
	Fried et al.	0.7884	0.0783	0.8085	0.0671	0.8160	0.0642	0.8718	0.0529
	Ruggiero	0.7016	0.1403	0.7735	0.1162	0.8007	0.1038	0.8626	0.0686
	New (DEA)	0.7758	0.0818	0.8094	0.0660	0.8274	0.0604	0.8607	0.0575
	New (TR)	0.8041	0.0641	0.8532	0.0567	0.8682	0.0582	0.8985	0.0710
	New (TO)	0.8175	0.0583	**0.8626**	0.0570	**0.8749**	0.0601	0.9008	0.0722
	New (DF)	**0.8213**	**0.0570**	0.8601	**0.0541**	0.8709	**0.0573**	0.8984	0.0706
	New (SF)	0.7873	0.0764	0.8164	0.0647	0.8389	0.0583	**0.9022**	**0.0439**

* See equation (2.20).

Although the true relationships in the first stage are grounded in decreasing returns to scale in discretionary inputs, performance in Table 2.1 is apparently better than performance in Table 2.2 for models 222, 141, and 411. However, model 114 in Table 2.2 performs slightly better than in Table 2.1 for most cases when the number of observation is greater than 50. Because the constant returns to scale assumption is preferred in most cases, the following discussion would mainly focus on Table 2.1 only.

In Table 2.1, the standard model is the worst by any means due to the deficient treatment of environmental factors. The diversity among the rest of the approaches is due to not only to the fact that the techniques (e.g. DEA, TR, TO, OLS, SF) employed in the second stage are different but the major concern is how to appropriately deal with the non-discretionary inputs. Therefore, it would be interesting to compare the approaches that use the same technique.

First, the Fried *et al*. model employs the stochastic frontier model, so it would be comparable with my approach using the stochastic frontier model, New (SF). For models 222, 141, and 411 where constant returns to scale is preferred, my approach outperforms the Fried *et al*. model in most cases except for models 222 and 411 when the number of observation is 50. In the case of model 114 where variable returns to scale is preferred, my approach is still superior measured by rank correlation.

Table 2.2
Simulation results: base case (variable returns to scale)

Number of firm		50		100		150		500	
Model*	Approach	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD
222	Fried et al.	0.7019	0.0899	0.7622	0.0708	0.7865	0.0637	0.8463	0.0545
	Ruggiero	0.4533	0.1681	0.5786	0.1495	0.6269	0.1385	0.7572	0.1037
	New (TR)	0.7095	0.0733	0.8016	0.0610	0.8368	0.0584	0.8938	0.0710
	New (TO)	0.7462	0.0680	**0.8251**	0.0658	**0.8518**	0.0660	**0.8993**	0.0797
	New (DF)	**0.7601**	**0.0622**	0.8225	**0.0563**	0.8475	**0.0568**	0.8961	0.0726
	New (SF)	0.7027	0.0876	0.7640	0.0693	0.7905	0.0626	0.8869	**0.0438**
114	Fried et al.	0.8657	0.0533	0.9062	**0.0409**	0.9192	**0.0393**	0.9418	**0.0342**
	Ruggiero	0.6450	0.0944	0.7460	0.0708	0.7968	0.0599	0.8852	0.0369
	New (TR)	0.8572	0.0644	0.9145	0.0607	0.9289	0.0655	0.9500	0.0872
	New (TO)	0.8843	0.0514	**0.9254**	0.0554	0.9345	0.0627	0.9500	0.0850
	New (DF)	**0.8845**	**0.0494**	0.9241	0.0539	0.9333	0.0612	0.9494	0.0841
	New (SF)	0.8717	0.0626	0.9210	0.0455	**0.9352**	0.0407	**0.9524**	0.0373
141	Fried et al.	0.8202	0.0732	0.8493	0.0600	0.8619	0.0548	0.8847	0.0469
	Ruggiero	0.6351	0.1550	0.7139	0.1369	0.7676	0.1254	0.8640	0.0927
	New (TR)	0.8230	0.0612	0.8715	0.0506	0.8851	0.0475	0.9100	0.0543
	New (TO)	0.8432	0.0548	0.8795	0.0499	**0.8911**	0.0481	**0.9114**	0.0569
	New (DF)	**0.8496**	**0.0530**	**0.8801**	**0.0470**	0.8904	**0.0450**	0.9104	0.0532
	New (SF)	0.8210	0.0722	0.8508	0.0606	0.8631	0.0544	0.8922	**0.0454**
411	Fried et al.	0.4786	0.1280	0.5447	0.1077	0.5760	0.0985	0.6812	0.0809
	Ruggiero	0.2803	0.1926	0.3675	0.1850	0.4207	0.1791	0.5520	0.1551
	New (TR)	0.4907	0.1005	0.6374	0.0780	0.6867	0.0711	0.7944	0.0799
	New (TO)	0.5129	0.0944	0.6581	0.0875	**0.7157**	0.0875	**0.8203**	0.1023
	New (DF)	**0.5472**	**0.0857**	**0.6596**	**0.0702**	0.7034	**0.0687**	0.8047	0.0859
	New (SF)	0.4781	0.1275	0.5461	0.1068	0.5775	0.0978	0.7611	**0.0616**

* See equation (2.20).

Second, the Ruggiero model using OLS in the second stage should be compared with my approach with the deterministic frontier, New (DF). My approach is better for all cases as measured by rank correlation. The Ruggiero model is good in MAD only in models 114 and 222 where substantial numbers of observations are present.

Third, though my approach (DEA) is slightly better than the Muñiz model in most cases, both approaches perform relatively poorly. Therefore, it is not particularly interesting to consider these models further, including the standard model, so I have removed them from the remainder of the tables.

Lastly, among all approaches, the best approach by certain criterion is the one marked with bold numbers. In Table 2.1, my approach is mostly best by both criteria in models 222, 141, and 411, while the Fried *et al*. model is the best under model 114. However, if the right assumption is chosen for model 114, my approach is still the best by the rank correlation measure as shown in Table 2.2.

Additionally, it is worth noting that the stochastic frontier technique has the most complicated likelihood function compared to other maximum likelihood techniques I have used in these exercises. Consequently, although the stochastic frontier seems to be preferred in some situations, it sometimes suffers from technical problems when sigma u_j (σ_j) is equal to zero or all estimators are zero. This problem would make the second stage of the Fried *et al*. model and my "New (SF)" worthless. Therefore, if this happens with any application, one should switch to another method.

Table 2.3
Simulation results: firms with different scale (constant returns to scale)

Number of firm		50		100		150		500	
Model*	Approach	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD
222	Fried et al.	**0.6566**	**0.1009**	0.6784	0.1058	0.6830	0.1089	0.7042	0.1175
	Ruggiero	0.5375	0.1123	0.6093	**0.1033**	0.6305	**0.1014**	0.6820	**0.1023**
	New (TR)	0.6436	0.1106	0.6780	0.1172	0.6859	0.1220	0.7065	0.1374
	New (TO)	0.6476	0.1080	**0.6818**	0.1166	**0.6882**	0.1230	**0.7079**	0.1383
	New (DF)	0.6494	0.1064	0.6817	0.1154	0.6879	0.1218	0.7075	0.1378
	New (SF)	0.6427	0.1157	0.6671	0.1290	0.6656	0.1397	0.6359	0.1859
114	Fried et al.	**0.7128**	**0.0861**	**0.7346**	**0.0965**	**0.7349**	**0.1006**	**0.7379**	**0.1085**
	Ruggiero	0.5677	0.1157	0.6412	0.1203	0.6582	0.1239	0.6978	0.1355
	New (TR)	0.6724	0.1415	0.6973	0.1471	0.7007	0.1503	0.7118	0.1612
	New (TO)	0.6732	0.1315	0.6985	0.1421	0.7011	0.1469	0.7119	0.1600
	New (DF)	0.6741	0.1322	0.6986	0.1425	0.7011	0.1472	0.7119	0.1601
	New (SF)	0.6621	0.1471	0.6707	0.1592	0.6632	0.1648	0.5353	0.1971
141	Fried et al.	**0.6792**	**0.0979**	0.6969	0.1012	0.6981	0.1017	**0.7307**	**0.0953**
	Ruggiero	0.6163	0.1016	0.6506	**0.0989**	0.6700	**0.0987**	0.7050	0.0998
	New (TR)	0.6728	0.1079	0.6949	0.1147	0.6959	0.1190	0.7146	0.1316
	New (TO)	0.6777	0.1091	**0.6994**	0.1156	**0.6994**	0.1198	0.7167	0.1318
	New (DF)	0.6769	0.1078	0.6984	0.1148	0.6985	0.1192	0.7161	0.1317
	New (SF)	0.6573	0.1223	0.6808	0.1258	0.6816	0.1409	0.6224	0.1797
411	Fried et al.	0.6132	0.1037	0.6498	**0.1063**	0.6445	0.1122	0.6705	0.1271
	Ruggiero	0.5262	0.1256	0.5883	0.1081	0.6098	**0.1035**	0.6645	**0.0993**
	New (TR)	0.6254	0.1020	0.6658	0.1096	0.6667	0.1176	0.6927	0.1399
	New (TO)	0.6275	0.1033	**0.6696**	0.1131	**0.6709**	0.1213	**0.6955**	0.1429
	New (DF)	**0.6276**	**0.1006**	0.6689	0.1105	0.6695	0.1192	0.6947	0.1417
	New (SF)	0.6123	0.1066	0.6534	0.1079	0.6574	0.1175	0.6565	0.1691

* See equation (2.20).

2.5.2 Second case: Firms with different scale

It is not clear which returns to scale assumption is more appropriate under certain circumstances. At least I have found the case where the variable returns to scale assumption is to be preferred. I alter the data generating process (DGP) slightly by generating the first half of x_{ij} from Uniform(30,50) and the second half from Uniform(60,100) instead of Uniform(30,50) for all observation j. As mentioned earlier, the true relationship in the first stage should fit with the variable returns to scale assumption, so I introduce size differences to make this point clearer. Note that the constant returns to scale frontier and the variable returns to scale frontier could be somewhat similar when observations are concentrated, but they would definitely diverge when observations are spread out.

Table 2.3 and Table 2.4 show results when firms have different sizes according to the above DGP. The only difference is that Table 2.3 assumes constant returns to scale but Table 2.4 assumes variable returns to scale. Clearly, results in Table 2.4 for the models 222, 114, and 141 overcome results in Table 2.3 for every approach except that the Ruggiero model's MAD in Table 2.3 is preferred in some cases. For model 411, the variable returns to scale assumption is good only when we observe a large number of observations (e.g. 500 firms). However, the Ruggiero model is the only one that doesn't work out well with the variable returns to scale assumption. Actually, the Ruggiero model is the most DEA-oriented model of all those considered in this study. The variable-returns-to-scale DEA may not be able to function well if the number of observations is relatively limited.

Table 2.4 shows that my approach with the stochastic frontier technique again outperforms the Fried *et al*. model for the model 222, 114, 141, and 411 (only for the 500 observations case). However, if constant returns to scale are chosen for the model 411 when the number of observations is less than 500, my approach and the Fried *et al*. model appear to be indifferent in terms of results.

Table 2.4
Simulation results: firms with different scale (variable returns to scale)

Number of firm		50		100		150		500	
Model*	Approach	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD
222	Fried et al.	0.6882	0.0990	0.7428	0.0783	0.7613	0.0713	0.8194	0.0614
	Ruggiero	0.5086	0.1675	0.6186	0.1491	0.6716	0.1386	0.7860	0.1055
	New (TR)	0.6789	0.0784	0.7830	0.0644	0.8222	0.0623	0.8890	0.0817
	New (TO)	0.7091	0.0720	**0.8018**	0.0669	**0.8355**	0.0679	**0.8949**	0.0892
	New (DF)	**0.7254**	**0.0684**	0.7998	**0.0588**	0.8298	**0.0591**	0.8899	0.0814
	New (SF)	0.6887	0.0975	0.7444	0.0772	0.7662	0.0698	0.8843	**0.0442**
114	Fried et al.	0.8506	0.0551	0.8964	0.0425	0.9071	0.0413	0.9250	0.0387
	Ruggiero	0.6598	0.0994	0.7735	0.0746	0.8201	0.0638	0.9038	0.0394
	New (TR)	0.8500	0.0637	0.9148	0.0566	0.9328	0.0628	0.9591	0.0843
	New (TO)	0.8754	0.0512	0.9260	0.0507	0.9385	0.0593	0.9588	0.0820
	New (DF)	**0.8776**	**0.0490**	0.9247	0.0494	0.9371	0.0581	0.9580	0.0811
	New (SF)	0.8680	0.0598	**0.9265**	**0.0396**	**0.9390**	**0.0374**	**0.9623**	**0.0319**
141	Fried et al.	0.7232	0.089	0.7519	0.0784	0.7569	0.0757	0.7845	0.0715
	Ruggiero	0.6473	0.1565	0.7259	0.1391	0.7765	0.1286	0.8674	0.0972
	New (TR)	0.7647	0.0767	0.8110	0.0750	0.8233	0.0813	**0.8558**	0.1059
	New (TO)	**0.7737**	0.0742	**0.8164**	0.0785	**0.8269**	0.0842	0.8557	0.1052
	New (DF)	0.7734	**0.0668**	0.8127	**0.0698**	0.8218	0.0759	0.8524	0.0995
	New (SF)	0.7262	0.0872	0.7650	0.0759	0.7772	**0.0729**	0.8530	**0.0593**
411	Fried et al.	0.5161	0.1335	0.5711	0.1116	0.5986	0.1025	0.6779	0.0858
	Ruggiero	0.3717	0.1895	0.4752	0.1804	0.5360	0.1737	0.6608	0.1493
	New (TR)	0.4917	0.0999	0.6396	0.0781	0.6931	0.0727	0.8069	0.0839
	New (TO)	0.5067	0.0941	0.6478	0.0830	**0.7107**	0.0837	**0.8276**	0.1015
	New (DF)	**0.5406**	**0.0923**	**0.6508**	**0.0717**	0.6983	**0.0689**	0.8109	0.0858
	New (SF)	0.5158	0.1335	0.5716	0.1115	0.6007	0.1020	0.7754	**0.0616**

* See equation (2.20).

The Ruggiero model is inferior to my approach under the deterministic frontier in Table 2.4 in most cases except for models 114 and 141 with 500 observations. If constant returns to scale is assumed my approach is still carries a favorable rank correlation but not at MAD when observations total more than 50. The gap between the Ruggiero's performances and my approach's performance in most cases becomes closer as the number of observations increases. A larger array of sample sizes may be needed to see if this trend is persistent or not.

The best performer in each model in Table 2.4 is always one of my approaches by any criterion. However, for model 411 in Table 2.3, my approach is only best when measured by rank correlation, but the Ruggiero model is good at MAD when observations are equal to or greater than 150.[14]

2.5.3 Third case: Change in the true efficiency's variance

This experiment should be able to shed some light on how sensitive the models are to the distribution of true efficiency. In the base case, efficiency is generated as $\gamma = e^{-|u|}$, where u ~ N(0, 0.09). Here I introduce the diversity of firms' competence into the experiment by generating u from N(0, 0.25). Table 2.5 shows that all approaches for all models improve relative to Table 2.1 in term of rank correlation. Intuitively, I never experienced any model having unit rank correlation or zero MAD, which means that every method always incurred some error. When firms' efficiency is spread out, even if the models make errors, it should be easier to rank them in an accurate order than when each firm's score is concentrated. However, once scores spread out, the magnitude of error could be larger as well. We can see this phenomenon from the MAD of the Ruggiero model, the Fried *et al*. model, and my approach with the stochastic frontier technique. Surprisingly, my approaches (e.g. TR, TO, and DF) have performed fairly well as indicated by a lower MAD in most cases, especially when observations total more than 50.

[14] See Appendix A for additional simulation exercises to supplement the base case and the second case.

Table 2.5
Simulation results: expand the distribution of the true efficiency (constant returns to scale)

Number of firm		50		100		150		500	
Model*	Approach	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD
222	Fried et al.	**0.9251**	0.0718	0.9255	0.0573	0.9430	0.0491	0.9557	**0.0428**
	Ruggiero	0.7906	0.1361	0.8690	0.1047	0.8924	0.0908	0.9407	0.0561
	New (TR)	0.9096	0.0591	0.9376	0.0499	0.9490	0.0472	0.9648	0.0501
	New (TO)	0.9208	0.0548	**0.9450**	0.0465	**0.9541**	0.0447	0.9661	0.0496
	New (DF)	0.9238	**0.0545**	0.9446	**0.0459**	0.9534	**0.0441**	0.9655	0.0491
	New (SF)	0.9238	0.0767	0.9381	0.0664	0.9485	0.0620	**0.9672**	0.0581
114	Fried et al.	0.9318	0.0571	**0.9586**	**0.0411**	**0.9658**	**0.0369**	**0.9735**	**0.0351**
	Ruggiero	0.8239	0.0712	0.8962	0.0562	0.9198	0.0518	0.9545	0.0497
	New (TR)	0.9346	0.0652	0.9549	0.0623	0.9623	0.0628	0.9692	0.0677
	New (TO)	0.9433	**0.0540**	0.9582	0.0562	0.9640	0.0585	0.9693	0.0669
	New (DF)	**0.9436**	0.0544	0.9582	0.0566	0.9640	0.0588	0.9693	0.0670
	New (SF)	0.9320	0.1111	0.9501	0.1025	0.9635	0.0920	0.9700	0.0771
141	Fried et al.	0.9373	0.0687	0.9469	0.0554	0.9535	0.0505	0.9703	**0.0393**
	Ruggiero	0.8904	0.1270	0.9205	0.1037	0.9421	0.0907	0.9647	0.0593
	New (TR)	0.9441	0.0477	0.9585	0.0391	0.9646	0.0381	0.9747	0.0408
	New (TO)	**0.9538**	**0.0430**	**0.9647**	**0.0366**	**0.9691**	**0.0362**	0.9756	0.0405
	New (DF)	0.9531	0.0436	0.9628	0.0369	0.9673	0.0365	0.9746	0.0406
	New (SF)	0.9394	0.0684	0.9571	0.0547	0.9661	0.0507	**0.9787**	0.0432
411	Fried et al.	0.8890	0.0921	0.9038	0.0719	0.9090	0.0646	0.9251	0.0563
	Ruggiero	0.7829	0.1780	0.8462	0.1442	0.8719	0.1274	0.9228	0.0812
	New (TR)	0.8906	0.0703	0.9206	0.0558	0.9304	0.0518	0.9512	**0.0547**
	New (TO)	0.9019	**0.0647**	**0.9298**	**0.0519**	**0.9378**	0.0493	0.9537	0.0561
	New (DF)	**0.9040**	0.0671	0.9279	0.0522	0.9356	**0.0492**	0.9524	0.0550
	New (SF)	0.8895	0.0913	0.9056	0.0729	0.9156	0.0671	**0.9549**	0.0560

* See equation (2.20).

The comparison between each approach still displays a similar pattern as the base case where my approaches are preferred in most cases. The variable returns to scale version of this experiment performs worse than the constant returns to scale models in most cases, including model 114. Because of the similarity in results, I only show results from the constant returns to scale version.

Table 2.6
Simulation results: Cobb Douglas production function with uneven Power (constant returns to scale)

Number of firm		50		100		150	
Model*	Approach	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD
222	Fried et al.	0.8497	0.0617	0.8651	0.0535	0.8811	0.0499
	Ruggiero	0.7485	0.1086	0.8178	0.0865	0.8402	0.0761
	New (TR)	0.8471	0.0573	0.8851	0.0564	0.9011	0.0564
	New (TO)	0.8620	0.0519	**0.8942**	0.0542	**0.9058**	0.0550
	New (DF)	**0.8647**	**0.0501**	0.8926	**0.0525**	0.9041	0.0535
	New (SF)	0.8475	0.0617	0.8777	**0.0525**	0.8972	**0.0475**
114	Fried et al.	0.8800	**0.0469**	**0.9185**	**0.0409**	**0.9303**	**0.0392**
	Ruggiero	0.7767	0.0615	0.8483	0.0541	0.8714	0.0532
	New (TR)	0.8870	0.0685	0.9108	0.0704	0.9186	0.0718
	New (TO)	0.8958	0.0600	0.9134	0.0656	0.9201	0.0688
	New (DF)	**0.8963**	0.0605	0.9134	0.0660	0.9200	0.0691
	New (SF)	0.8802	0.0709	0.9126	0.0609	0.9209	0.0554
141	Fried et al.	0.8765	0.0615	0.9053	0.0506	0.9161	0.0469
	Ruggiero	0.8407	0.1055	0.8748	0.0875	0.9011	0.0774
	New (TR)	0.8968	0.0466	0.9241	0.0431	0.9313	0.0445
	New (TO)	**0.9107**	0.0432	**0.9298**	0.0422	**0.9350**	0.0441
	New (DF)	0.9091	**0.0424**	0.9269	**0.0413**	0.9320	0.0435
	New (SF)	0.8843	0.0541	0.9189	0.0434	0.9320	**0.0382**

* See equation (2.21).

2.5.4 Fourth case: Cobb Douglas production with uneven power

Because all the above exercises assume equal marginal effects within each type of inputs relative to output, I make them different in this last experiment. The new true production functions are as follows:

$$\begin{aligned} &\text{Model 222:} && (y_1^2 + y_2^2)^{.5} = x^{.5}x^{.1}z^{.3}z^{.1} \\ &\text{Model 114:} && y = x^{.6}z^{.17}z^{.1}z^{.08}z^{.05} \\ &\text{Model 141:} && y = x^{.3}x^{.15}x^{.1}x^{.05}z^{.4} \end{aligned} \tag{2.21}$$

As shown in Table 2.6, all results are surprisingly similar to the base case. The comparisons among all approaches would be same as the base case. In sum, at least in this exercise, all approaches are not sensitive to the marginal effect of the Cobb Douglas production function.

2.6 Conclusions

I offer a new method of DEA efficiency measurement to deal with the presence of non-discretionary inputs where nonparametric techniques (DEA) or parametric techniques (e.g. TR, TO, DF, SF) can be adopted in the second stage. The potential advantage of my approaches in a multi-output, multi-input setting relative to the most prominent alternative method is demonstrated through simulation studies. In particular, my approaches display superior performances in terms of both rank correlation and MAD in most situations. DEA practitioners, particularly those in education and health, will find my framework to be of value.

It is quite surprising that though firms are generated from decreasing returns to scale (in discretionary inputs) technology, Table 2.1 and 2.2 show that assuming CRS returns better results than assuming VRS. However, Table 2.3 and 2.4 show the opposite results where VRS assumption is preferred when firms have more scale differences. Furthermore, my approaches using the TR, TO, and DF techniques improve their rank correlations and MADs when firms are heterogeneous in efficiency score level, while other approaches improve only their rank correlation. Finally, none of the approaches is sensitive to changes in marginal effect from the predetermined Cobb Douglas production function.

2.7 Future Research

In the standard DEA model literature, where non-discretionary inputs do not exist, the Malmquist productivity index is a commonly used tool to analyze productivity changes overtime. However, it has not been widely used in the DEA with environmental factors yet. Therefore, future research could study how to apply the Malmquist index to my alternative model. Conceptually, the Malmquist index cannot be directly applied to any of the above multi-stage DEA models, because the environment in two periods could be theoretically different. In the context of my alternative model, if one wants to level the environment between two periods, a parametric approach in the second stage may be needed to quantify the impact of environmental factors. Then the discretionary inputs could be adjusted by the same reasoning as in the Fried *et al*. model except that the adjustment is based on theta, not on the level of slacks. However, all of the above parametric techniques are basically linear in theta-z space, so a problem could arise for a certain range of differences in z when the adjustment in theta is larger than one.

CHAPTER III

A NOTE ON THE STATISTICAL INFERENCE OF A MULTI-STAGE DEA MODEL WITH NON-DISCRETIONARY INPUTS

3.1 Simar and Wilson (2007) Review

Simar and Wilson (SW) criticize dozens of papers in the area of multi-stage DEA with non-discretionary inputs because their statistical inference is invalid mainly due to the serial correlation problem. The problem arises because all firms' efficiency estimates (DEA score in the first stage) depend on the same frontier. Also, Simar and Wilson create a statistical model defined by assumptions in which a truncated regression, but not a tobit regression, is sensible in the second stage. Simar and Wilson suggest that the truncated model should be the only model one should use because the probability to observe efficient firms in finite sample is zero. Therefore, any observation which has a DEA score equal to one is spurious and should be removed.

Simar and Wilson concentrate on the serial correlation problem and claim that there is an endogeneity problem as well. However, the problems will disappear asymptotically, and then the estimator is consistent. Nonetheless, the convergence rate is quite slow.

Therefore, to improve the statistical inference, Simar and Wilson suggest two bootstrap algorithms. They first describe a DGP where efficiency estimates do not yield a probability mass at one. Therefore, a truncated model is more appropriate than a tobit model. Bootstrap

Algorithm I is similar to a traditional bootstrap except that the error term is drawn from an independent truncated normal distribution. Algorithm II incorporates a two-step procedure. First, Simar and Wilson generate a sample set of x and y in order to recalculate the DEA score to obtain a bias-corrected DEA score. Then, the second step is just the same as Algorithm I, except that it uses the bias-corrected estimate to calculate original estimators.

Finally, Simar and Wilson provide simulation exercises to examine their proposed bootstrap procedures and to compare the truncated model with the tobit model using coverage of confidence intervals and root-mean-square-errors (RMSE) of the estimates as criteria.

3.2 Comments on the Simar and Wilson (2007)

If the convergence rate is slow, then the problems remain. The serial correlation problem in generating efficiency estimates, θ, causes the error terms to be correlated with each other which means that the independence assumption of maximum likelihood (e.g. the tobit model or the truncated model) is violated. In order to perform the maximum likelihood estimation, one needs to know the serial correlation structure. Unfortunately, the form of the serial correlation is unknown, so it is probably impossible to improve the estimators themselves. In the aspect of statistical inference, once the tobit model or the truncated model is estimated, it is possible that bootstrapping could help improve the statistical inference. However, in principle, it is not clear that the SW bootstrap processes are able to cope with the problem. Since the problem is serial correlation, bootstrap process should somewhat preserve the dependence structure of the error term (by the same reasoning as block bootstrap). The SW bootstrap processes draw errors from independent truncated normal distributions instead of using the empirical error, as the traditional bootstrap does.

Besides, although Simar and Wilson's DGP looks sensible, it could make the truncated model better fit with their generated sample than other models because their DGP employs truncated normal distributions. Consequently, Simar and Wilson use their simulation results to show that the truncated model outperforms the tobit model and explain that the tobit model has a misspecification problem. However, in my view, the truncated model might outperform the tobit model in this particular circumstance because of the way the data was generated.

Unlike Simar and Wilson's view, I believe that the tobit model is one of the most appropriate models as well. I agree with the assumption that the probability of observing efficient firms is close to zero. However, in a finite sample, the DEA mechanism in the first stage always inflates *all* firms' scores and, in particular, assigns one to all best-practice firms which are actually not truly efficient. It is hard to determine a pattern of the bias that inflates the scores. Therefore, a likelihood function cannot be modified in order to properly deal with the bias. In this sense, all maximum likelihood models including the truncated model would suffer from a misspecification problem. However, maximum likelihood technique is still attractive because of its advantages and might be considered as a second best technique. *All* firms' score are contaminated by the unpredictable bias, but Simar and Wilson solve the problem by removing *only* the best-practice firms which they treat as spurious and employ the truncated model to estimate a regression. To me, all observations are spurious, but we cannot help it. Once all observations are contaminated due to the DEA technique in the first stage, compiling all observations including the unit-score firms by employing the tobit model should not conflict with the assumption that there is no efficient firm in a sample set. It is true that we do not have probability mass on efficient firms, but after we assign scores in the first stage we do have probability mass at one, and the tobit model should be able to work fine in this circumstance (as illustrated in Figure 2.7). To examine these issues more closely, Monte Carlo simulations are set up in the next section.

3.3 Simulation Experiment

In fact, Simar and Wilson provide several simulation exercises to compare between their bootstrap and conventional method and to compare between the truncated model and the tobit model. However, the simulation in this section is intended to make the comparison more clear and complete. First of all, size and power should be the common criteria to measure the effectiveness of the hypothesis testing method, so they should be expressed explicitly. Yet, Simar and Wilson's criteria somewhat infer size and power as well. They predetermine the true value of the estimator and account for the coverage of confidence interval. It implies that their null hypothesis (H_0: β = true value) is true. Therefore, the coverage of confidence interval is the value of one minus size. The second criterion, RMSE, implies a range of the confidence interval which inversely relates with power. Second of all, the conventional method and the tobit model are compared only by the first criterion but not the second one; hence the comparison is not complete. In addition, when they compare the tobit model with the truncated model, their DGP fits better with the truncated model, and results in the superior performance of the truncated model. For these reasons, the following sets of simulation experiments are provided to explicitly show the magnitude of size and power and to carefully incorporate both the conventional method and the tobit model into the analysis.

3.3.1 Simulation I: Does the SW bootstrap outperform the traditional t-test in the presence of serial correlation?

The first set of the simulation exercises, Simulation I, is to compare between the Algorithm I bootstrap and t-tests. (There are two sets of t-tests; one uses traditional standard errors and the other uses robust standard errors.[15]) The DGP is analogous to Simar and Wilson's where number of x and number of y equal to two. However, the set of criteria in this case is size and power rather than the coverage of the confidence interval and RMSE.

If we are concerned about power, equation 3.1 demonstrates a linear relationship between theta and non-discretionary inputs, z_1 and z_2. However, if we consider size, z_2 will be left out of equation 3.1. Each of the nondiscretionary inputs is independently generated from a N(2,4). Then ε is drawn from a N(0,1) left-truncated $1-\beta_0-\beta_1 z_1-\beta_2 z_2$ if power is measured,

[15] The Huber/White robust standard error is not specifically designed to solve a problem in this situation. However, it tends to be routine in the literature to compute this estimator. Therefore, it's also interesting to see how much it helps improve the statistical inference comparing to others.

but $1-\beta_0-\beta_1 z_1$ if size is measured, where $\beta_0 = \beta_1 = 0.5$ and $\beta_2 = 0.1$.[16] Subsequently, a series of the true thetas is obtained.

There are two discretionary inputs, x_1 and x_2, and two outputs, y_1 and y_2. Each x is independently drawn from Uniform[6,16]. Then, $y_1 = \theta^{-1}(x_1^{0.75} + x_2^{0.75})(\alpha)$ and $y_2 = \theta^{-1} (x_1^{0.75} + x_2^{0.75})(1-\alpha)$, where $\alpha \sim$ Uniform[0,1].

After a series of thetas is estimated from the discretionary inputs and the outputs using an output-oriented DEA technique, the estimated theta is regressed on z_1 and z_2 to obtain the estimators of β_0, β_1, and β_2 as in equation 3.1. This simulation exercise focuses on the hypothesis testing of β_2 only. When size is considered, the null hypothesis, $\beta_2=0$, is true because z_2 is left out of the DGP. On the other hand, when power is considered, the null hypothesis, $\beta_2=0$, is false because z_2 is incorporated in the DGP.[17]

$$\theta = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \varepsilon \tag{3.1}$$

In each experiment, the simulation is repeated 500 times, while the Algorithm I bootstrap[18] is conducted with 1,000 bootstrap replications.

---

[16] The original β_2 Simar and Wilson use in their paper was 0.5. When I use $\beta_2=0.5$ in the following experiment, every approach's power becomes one and cannot be distinguished. Therefore, to decrease their power, I change β_2 to 0.1, so that the null hypothesis (H_0: $\beta_2=0$) can be easily rejected.

[17] Size and power shown in tables are the percentage of how many times its null hypothesis is rejected.

[18] For example, suppose the level of significance (α) is 0.05, all bootstrap estimators will be sorted by ascending order and use 2.5 percentile and 97.5 percentile estimators to obtain the 95% confidence interval.

As shown in Table 3.1, every approach of hypothesis testing works fine and seems to have similar performance. However, the traditional t-test performs a bit better than the SW bootstrap because it has higher power and size that is closer in number to the level of significance.

Table 3.1
Result of the simulation I

Number of observation	Level of significance	Truncated Regression					
		Bootstrap (Algo. I)		t-test		t-test (Robust se)	
		size	power	size	power	size	power
150	$\alpha = 0.1$	0.132	0.442	0.118	0.456	0.140	0.462
	$\alpha = 0.05$	0.068	0.300	0.056	0.320	0.060	0.346
400	$\alpha = 0.1$	0.118	0.814	0.086	0.836	0.088	0.830
	$\alpha = 0.05$	0.062	0.742	0.040	0.746	0.040	0.738

3.3.2 Simulation II: Does the truncated model outperform the tobit model in the aspect of hypothesis testing?

The second set of simulation exercises, Simulation II, is used to compare the truncated model with the tobit model. The simulation design is similar to the Simulation I except for the DGP. Simar and Wilson’s DGP tends to prefer the truncated model in this setting. However, in my view, removing unit-DEA firms could not confirm that the truncated model is the true model either. Not only are the best-practice firms contaminated from the DEA in the first stage, but so are the rest of the observations. Since observations in the second stage must contain at least one unit-score firm, the tobit model could be also a potential candidate to estimate a regression. Since both models are not correctly specified, it will be interesting to see which one performs better under specific circumstances. With the following DGP where the true model is unknown, I preserve the assumption that the probability to observe efficient firms is zero. Unlike Simar and Wilson's DGP, we cannot guess whether the tobit model or the truncated model fits better with the DGP. Yet, both the truncated model and the tobit model may not only suffer from serial correlation but also from misspecification problems.

Similar to Muñiz *et al*. (2006), each x and z is generated independently from the following distributions: first half of x_i ~ Uniform[30,50], second half of x_i ~ Uniform[60,100][19] for i=1,2, z_k ~ Uniform[1,2] for k =1,2. Then, $g(y_1,y_2) = x_1^{.3}x_2^{.3}z_1^{.2}z_2^{.2}$ if power is considered, and $g(y_1,y_2) = x_1^{.3}x_2^{.3}z_1^{.4}$ if size is considered, where $g(a,b) = (a^2+ b^2)^{.5}$. Each output (y_1 and y_2) is randomly drawn to satisfy the function g(.). Moreover, true efficiency is generated as $\gamma = e^{-|u|}$, where u ~ N(0, 0.09). Each efficient input (x) is scaled by 1/ γ to obtain observed inputs (x*). There are 150 firms in each simulation experiment.

In the first stage, theta is estimated from y and x* using an input-oriented DEA technique. Similar to the Simulation I, the estimated theta is regressed on z_1 and z_2. However, the relationship is estimated twice by employing two different models, a tobit model and a truncated model. Again, the hypothesis testing of z_2's coefficient is the only concern in this simulation.

All calculations are performed with 500 replications for each experiment. In each trial, 500 iterations are run to create critical values from bootstrap confidence intervals.

[19] This is to create size difference among firms.

Table 3.2
Result of the simulation II

Level of significance	Bootstrap (Algorithm I)				t-test (conventional se)				t-test (Robust se, Huber/White)			
	Size		Power		Size		Power		Size		Power	
	TR	TO	TR	TO	TR	TO	TR	TO	TR	TO	TR	TO
$\alpha = 0.1$	0.112	0.284	0.642	1.000	0.100	0.086	0.654	0.980	0.102	0.084	0.658	0.978
$\alpha = 0.05$	0.050	0.202	0.556	1.000	0.036	0.038	0.538	0.968	0.046	0.044	0.552	0.966

Table 3.2 shows that the SW bootstrap for the tobit model is oversized, but it has good power. The truncated model has good size and reasonable power. However, the performance of the SW bootstrap for a truncated model is not much different from traditional t-tests. In the case of the tobit model, the t-test has good size and even better power, there seems to be no difference between using the usual standard error or the robust standard error. In general, size and power should move together in the same direction. The size of the t-test for the tobit regression is a bit lower than that of the truncated regression, but its power is a lot higher than the power of the truncated regression. Therefore, estimating a tobit model and applying a t-test seems to be a good choice in conducting statistical inference.

3.4 Conclusions

From Simulation I, the traditional t-test works fine and does not underperform the SW bootstrap in term of size and power. From Simulation II where the true model is unknown, the tobit model seems to be preferred; its size is similar to the size of the truncated model, but its power is greater, especially when t-test is applied. Therefore, unlike Simar and Wilson, I conclude that the tobit model and the traditional t-test could be applicable tools for practitioners.

CHAPTER IV

APPLICATION TO TEXAS SCHOOL DISTRICTS

4.1 Introduction

The new approach was proposed in Chapter II and the hypothesis testing was analyzed in Chapter III. This chapter will provide an application of those tools. An education sector is selected here since it has the desirable feature that schools cannot control some factors, such as students' socio-economic or some schools' characteristics (at least in the short run). However, they have impacts on schools' performance. These factors are non-discretionary inputs, and therefore suit the new multi-stage DEA model. Besides, it has been researchers' theoretical concern regarding the importance of student level data which is usually not publicly available. In general, the individual data is needed even for the analysis at the aggregate level. Therefore, it becomes an issue that in most cases prevents researchers from conducting an analysis in this area. In the following sections, I will describe the importance of the individual data in theory and use the new tools to help quantify impact of the data on DEA scores.

4.2 Literature Review

1) Ruggiero (1996) applies his model to measure technical efficiency of 636 school districts in New York State for the 1990-1991 school year. His outputs are average scores of Pupil Evaluation Program (PEP) in reading mathematics and social studies. Discretionary inputs are number of teacher aides and teacher assistants per student, number of computers and classrooms per student, and proportion of teachers with at least a certain amount of training. There is only one non-discretionary input in this study which is the poverty rate. The aim of the paper is to measure Koopmans efficiency in New York State school districts by using his two-stage procedure based on the multivariate technique of a canonical regression.

2) Muñiz (2002) tests his model through an evaluation of 62 public high schools located in the Spanish region of Asturias during the 1996-1997 academic year. He analyzes the data by using a production approach. Two outputs he considers are (1) passing rate and (2) average grade obtained from students who pass. Two discretionary inputs he considers are (1) expenditure per student and (2) number of teachers per 100 students. A survey is conducted and summarized into five non-discretionary inputs which are (1) percentage of diligent students (2) percentage of students who believe that their parents have high prospects for them (3) percentage of rich family (4) percentage of students who did not transfer their school in that year and (5) percentage of students who are only children. He compares his approach with a one-stage model by Banker and Morey (1986) and finds that many units labeled as efficient in the one-stage model are actually highly inefficient in his multi-stage model.

4.3 Data

Due to data availability, I have information for 560 school districts in Texas from 2004 to 2006. Following Muñiz (2002) and Ruggiero (1996), I adopt a similar set of variables in order to calculate technical efficiency but use a different multi-stage DEA technique. A production approach needs three groups of quantity variables to analyze technical efficiency: outputs, discretionary inputs, and non-discretionary inputs.

Table 4.1
Descriptive statistics of Texas school district data

	2004			2005			2006		
	mean	min	max	mean	min	max	mean	min	max
Outputs									
MAT	78.08	45.29	98.13	73.61	42.57	93.90	76.60	45.31	95.89
REA	88.09	67.85	45.29	86.22	63.75	97.97	90.11	69.29	100.00
CR	93.84	40.00	100.00	93.80	33.30	100.00	92.83	36.20	100.00
ADV	17.87	0.00	47.20	18.27	0.00	66.10	18.38	0.00	68.30
SAT	13.13	0.00	50.05	13.67	0.00	53.28	13.33	0.00	53.98
Discretionary Inputs									
TEA	0.08	0.06	0.14	0.08	0.06	0.14	0.08	0.06	0.15
AID	0.07	0.01	0.20	0.07	0.02	0.20	0.07	0.02	0.17
ADM	0.01	0.00	0.03	0.01	0.00	0.03	0.01	0.00	0.03
Non-discretionary Inputs									
HGH	29.86	16.63	60.78	29.93	13.72	58.10	29.89	14.64	53.92
ECON	48.91	0.10	97.80	50.80	3.20	98.10	51.80	4.20	97.90
SPEC	13.73	6.00	28.30	13.78	5.00	25.90	13.07	4.30	23.60
SIZ	4,748.78	121.00	211,157	4,814.13	112.00	208,454	4,941.91	105.00	209,879
LPR	86.14	60.21	97.43	84.80	60.43	97.26	80.92	56.29	96.08

Table 4.1 displays statistics for the raw data of 560 school districts in Texas. Five outputs are specified: 1) MAT, a mathematics passing rate; 2) REA, an English (reading) passing rate;[20] 3) CR the completion rate; 4) ADV, the percentage of students taking at least one advanced course; and 5) SAT, the percentage of students having SAT scores above a certain criterion. Three discretionary inputs are specified: 1) TEA, teacher per student ratio; 2) AID, aide per student ratio; and 3) ADM, administrator (including both central and campus) per

[20] Both passing rates measure share of students who met "standard". However, the standard was increasing over time from 2003 through 2006.

student ratio. Five non-discretionary inputs or environmental factors are specified: 1) HIGH, the percentage of high school students; 2) ECON, the percentage of economic disadvantage student; 3) SPEC, the percentage of students having special education; 4) SIZ, the number of enrolled students in a district; and 5) LPR, the one year lagged average passing rate (mathematics and reading). Instead of using mathematics and reading passing rates separately, I average them in order to avoid a multicolinearity problem in the second stage regression due to its correlation of about 0.8 in each year.

4.4 Model and Result

To help select the most appropriate approach for Texas school district data, I design a tailor-made experiment whose DGP is similar to the real data's distributions. Except for the DGP, this simulation exercise is conducted by the same manner as the base case in Chapter II. There are four outputs, three discretionary inputs, and four non-discretionary inputs. Each discretionary and non-discretionary input is assumed to follow a uniform distribution where its range is the same as the real data's range in 2004 as shown below. The pre-determined production function is a constant returns to scale Cobb-Douglas production function, $(y_1^2 + y_2^2 + y_3^2 + y_4^2)^{.5} = x_1^{.2}x_2^{.2}x_3^{.2}z_1^{.1}z_2^{.1}z_3^{.1}z_4^{.1}$.

DGP:	$x_1 \sim$ Uniform[0.06,0.14]	$z_1 \sim$ Uniform[17,61]
	$x_2 \sim$ Uniform[0.01,0.2]	$z_2 \sim$ Uniform[0.1,97.8]
	$x_3 \sim$ Uniform[0,0.03]	$z_3 \sim$ Uniform[6,28.3]
		$z_4 \sim$ Uniform[121,211157]

Table 4.2 suggests that my approach to the tobit model assuming constant returns to scale is the best at rank correlation and therefore should be an appropriate model for Texas school district data.

Table 4.2
Simulation results: An experiment for Texas school district data

560 observations		CRS		VRS	
Model	Approach	Rank Correlation	MAD	Rank Correlation	MAD
434	Fried et al.	0.6080	**0.1041**	0.4945	0.1229
	Ruggiero	0.6337	0.1280	0.3814	0.1832
	New (DF)	0.6706	0.1184	0.5825	**0.1100**
	New (TO)	**0.6776**	0.1319	**0.5969**	0.1563

To calculate efficiency score, the process is the same as in Chapter II.[21] In the first stage, only outputs and discretionary inputs are needed to evaluate preliminary efficiency scores (theta). Then, the impact of non-discretionary inputs or environmental factors on the preliminary score would be quantified in the second stage using the tobit model. Finally, the information from the second stage would be used to adjust the preliminary score to see how much each discretionary input should be efficiently reduced given its environment.

In the second stage, equation 4.1 expresses a relationship between the preliminary constant-returns-to-scale DEA and the environmental factors:

[21] Note that all approaches in Chapter II are cross section models (not panel). Therefore, I estimate DEA score for Texas school districts in each year separately.

$$\theta_{jt} = \alpha_t + \beta_{1t}HGH_{jt} + \beta_{2t}ECON_{jt} + \beta_{3t}SPEC_{jt} + \beta_{4t}\ln(SIZ_{jt}) + \beta_{5t}LPR_{jt} + \varepsilon_{jt} \qquad (4.1)$$

where t is year 2004, 2005, and 2006

HGH, ECON, SPEC, and SIZ are factors in the current environment that schools cannot control but that have an impact on current preliminary performance (θ_{jt}). LPR is the only lagged variable in the model. It represents students' ability in the past, as a stock of human capital, which affects their current performance. It also incorporates impacts from other past environment factors (HGH, ECON, SPEC, and SIZ) that could influence the current performance. Therefore, LPR seems to be theoretically important and should not be omitted. Although the multi-stage model is a cross-sectional model at the district (aggregate) level, one still needs student level data in order to construct the variable LPR. This is because both current and lagged variables for a particular year must be collected from the same coverage of students even when students move from one district to another. For example, suppose A and B are the only two students in district X in 2003. At the beginning of 2004, C move from another state to district X, so three students are observed in 2004. To create 2004 variables, the current passing rate (2004 records) and lagged passing rate (2003 record) must be collected from all students including C no matter where he was located in the previous year. However, if we cannot track C's previous-year record, C would be removed from 2004 data set. In 2005, suppose B moves out of state, so that only A and C are the only two students whose records are available for two consecutive years (2004-2005). Then, current passing rate and lagged passing rate variables for 2005 would be based on A and C's records. In conclusion, if researchers do not have access to individual data, it is impossible to correctly create a consistent lagged variable. Unfortunately, it becomes a common problem for most researchers because the student level data is not publicly available.

In Chapter II, we see that, in terms of rank correlation and MAD, both the tobit and truncated models are quite similar though the tobit model performs somewhat better at rank correlation. Furthermore, Chapter III also compares between the two models in terms of statistical inference. Therefore, in this Chapter, I apply both models to the same application in order not only to examine the importance of the variable LPR but also to support the results in Chapter II.[22, 23] Each model is estimated twice; the first estimation includes LPR as in equation (4.1) but the second estimation does not include LPR. Tables 4.3 to 4.6 show the coefficient and p-value of each non-discretionary variable. Although the regression could suffer from the serial correlation as mentioned in Simar and Wilson (2007), it has been shown in Chapter III that t-test could be an applicable tool to perform the hypothesis testing.

Table4.3
Tobit estimation results (with LPR)

Regressand	2004		2005		2006	
	β	P > |t|	β	P > |t|	β	P > |t|
HGH	-0.0035	0.0006	-0.0032	0.0027	-0.0016	0.1462
ECON	-0.0014	0	-0.0015	0	-0.0016	0
SPEC	-0.0038	0.0003	-0.0055	0	-0.0052	0
ln(SIZ)	0.0477	0	0.0461	0	0.0531	0
LPR	0.0014	0.0617	0.0008	0.3186	0.0022	0.0008

[22] This application may not support the result in Chapter II well. Although it is true that results from both models are expected to be very similar as stated in Chapter II, the reasoning might be different. Because there are only 2%-3% of observations have unit score in the first stage, it is common that both models would perform toward OLS and therefore return similar results. (Note that when truncated model is estimated, all unit score observations would be removed.)

[23] Regardless of my comment on McDonald (2009) in Chapter II, McDonald (2009) suggested that employing either the tobit model or the OLS model would be indifferent due to the small number of unit score observations.

Table 4.4
Tobit estimation results (without LPR)

Regressand	2004		2005		2006	
	β	P > |t|	β	P > |t|	β	P > |t|
HGH	-0.0036	0.0005	-0.0033	0.0019	-0.0020	0.0732
ECON	-0.0017	0	-0.0017	0	-0.0022	0
SPEC	-0.0038	0.0004	-0.0054	0	-0.0049	0
ln(SIZ)	0.0474	0	0.0458	0	0.0526	0

Table 4.5
Truncated estimation results (with LPR)

Regressand	2004		2005		2006	
	β	P > |t|	β	P > |t|	β	P > |t|
HGH	-0.0033	0.0006	-0.0027	0.0096	-0.0015	0.1622
ECON	-0.0013	0	-0.0013	0	-0.0015	0
SPEC	-0.0039	0.0001	-0.0048	0	-0.0051	0
ln(SIZ)	0.0504	0	0.0522	0	0.0586	0
LPR	0.0011	0.1309	0.0008	0.2839	0.0025	0.0001

Table 4.6
Truncated estimation results (without LPR)

Regressand	2004		2005		2006	
	β	P > |t|	β	P > |t|	β	P > |t|
HGH	-0.0034	0.0005	-0.0029	0.0067	-0.0019	0.0700
ECON	-0.0015	0	-0.0015	0	-0.0021	0
SPEC	-0.0038	0.0001	-0.0047	0	-0.0047	0
ln(SIZ)	0.0501	0	0.0518	0	0.0578	0

All of the above tables are quite similar in sign, magnitude, and significance level. HGH has a negative impact on the first stage efficiency scores. It implies when share of the high school student is lower, a low-performance group of students is the first who leaves while other high-performance students remain. Therefore, the remaining students perform better. ECON refers to students having financial difficulties, and SPEC is students with either mental or physical disabilities. Therefore, they tend to have negative impacts on overall performance. Given all things being equal, SIZ implies economies of scale and helps improve performance. Last but not least, students' abilities in the past, LPR (in Tables 4.3 and 4.6), positively determine their current abilities. As stated in footnote 20, the passing standard was changing over time, so panel analysis would not be appropriate and the relationship between DEA scores and LPR could be changed over time. In other words, they are different proxies of students' ability. Therefore, the estimators are not comparable.

The final step is to compute DEA scores as in equation (2.9) where the distance can be calculated by equation (2.13). The input-oriented scores will always range from zero to one with at least one district having unit score. Figures 4.1 and 4.2 show distributions of final DEA scores in each year where Figure 4.1 is derived from the tobit model and Figure 4.2 is derived from the truncated model. Between two figures, the shapes of the histograms are quite similar which implies that results from the tobit and the truncated models are alike. Within each figure, the variable LPR does not play an important role due to the resemblance between the left and the right columns in each year.

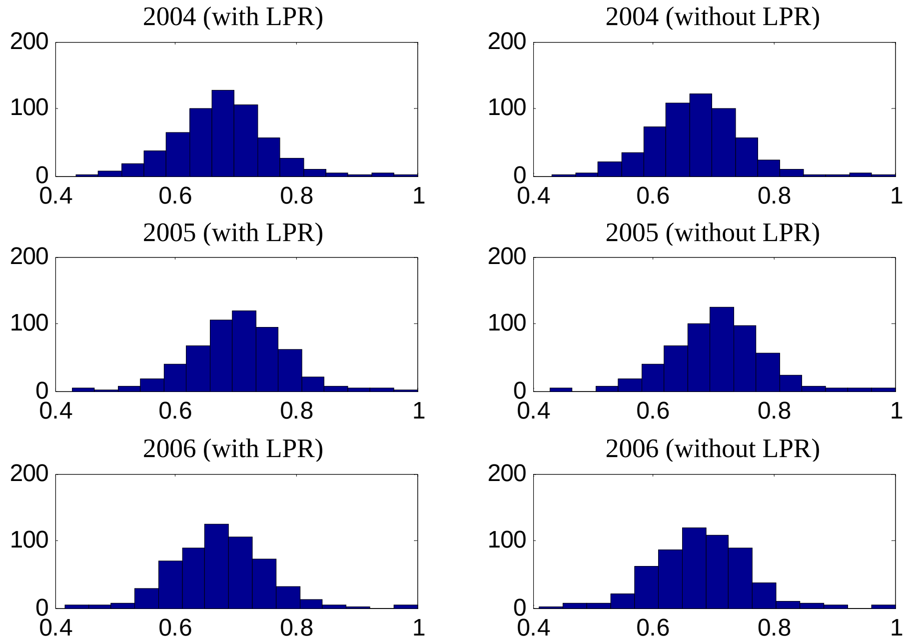


Figure 4.1 Distributions of final DEA scores from tobit model with and without LPR

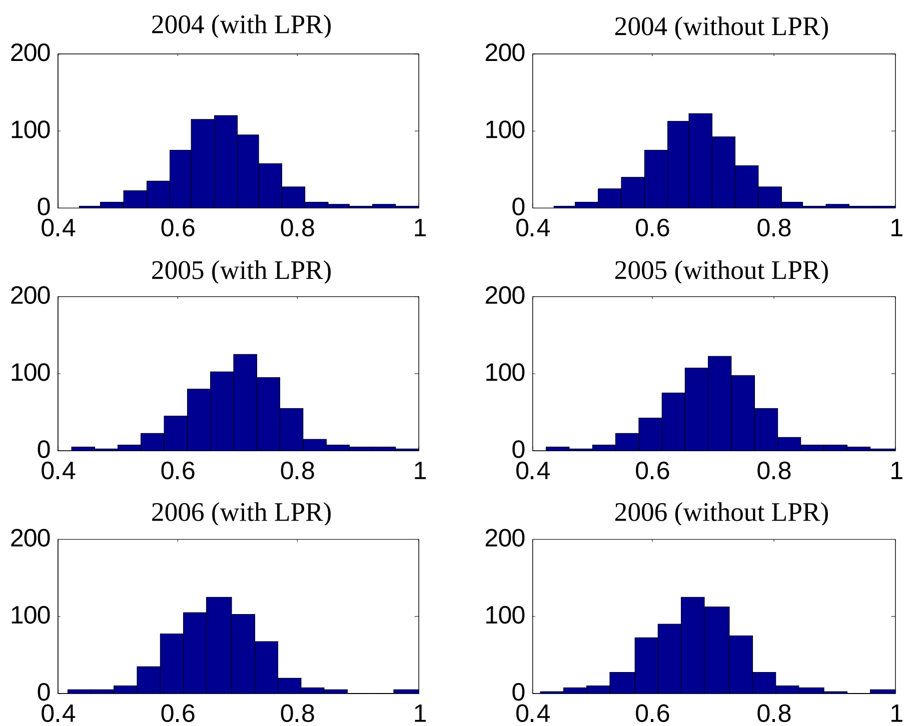


Figure 4.2 Distributions of final DEA scores from truncated model with and without LPR

Furthermore, I also report summary statistics of the final DEA scores. Due to the similarity among all the above cases, only the statistics from the tobit model with LPR is reported in Table 4.7. Table 4.8 illustrates the rank correlation between years. It shows how volatile in ranking school districts behave over time. Additionally, I also test all results in table 4.8 and find that all rank correlations are significantly different from zero at the 1% level of confidence.

Table 4.7
Summary statistics, final DEA score (tobit model with LPR)

	2004	2005	2006
Mean	0.7537	0.6731	0.7103
Median	0.7556	0.6734	0.7136
Maximum	1	1	1
Minimum	0.4175	0.3917	0.4060
Std. Dev.	0.0943	0.0880	0.0938

Table 4.8
Rank correlation matrix (tobit model with LPR)

	2004	2005	2006
2004	1	0.5933	0.4899
2005	0.5933	1	0.7160
2006	0.4899	0.7160	1

Lastly, Tables 4.9 and 4.10 display rank correlation and MAD among every possible combination between the two approaches (the one with LPR and the one without LPR) and two techniques (the tobit and the truncated models). High rank correlations and low MADs are robust across years. Therefore, the tobit and the truncated models return almost the same results which is consistent with the results in Chapter II. In addition, the variable LPR does not change the DEA scores much. Actually, the correlations between the current passing rate and the lagged passing rate (LPR) are fairly high: 0.91 in 2004, 0.92 in 2005, and 0.94 in 2006. Therefore, the additional information gained from including LPR might be minor.

Table 4.9
Comparisons of final DEA score between approaches with and without LPR

	Tobit model			Truncated model		
	2004	2005	2006	2004	2005	2006
Rank Correlation	0.9983	0.9994	0.9951	0.9990	0.9993	0.9937
MAD	0.0035	0.0018	0.0070	0.0027	0.0020	0.0084

Table 4.10
Comparisons of final DEA score between tobit model and truncated model

	Models with LPR			Models without LPR		
	2004	2005	2006	2004	2005	2006
Rank Correlation	0.9990	0.9976	0.9986	0.9991	0.9976	0.9987
MAD	0.0028	0.0071	0.0080	0.0023	0.0070	0.0065

4.5 Conclusions

Comparing to other leading existing multi-stage model of non-discretionary inputs, my approach is not only a powerful tool as shown in Chapter II, but also less demanding in terms of calculation. It only requires a simple adjustment as in equation (2.9) after a standard regression model is estimated. Therefore, it could be a handy tool in measuring efficiency analysis.

After applying a new approach to the Texas school districts, results between the case including LPR and not including LPR are compared using rank correlation and MAD. The conclusion is that LPR in the context of Texas school districts contribute only a little to the model and therefore might be excluded if the student level data is not available. Besides, employing either the tobit or the truncated models may return similar results.

CHAPTER V

BANK EFFICIENCY IN THAILAND: PRE AND POST FINANCIAL SECTOR MASTER PLAN

5.1 Introduction

After the recovery from the economic crisis in 1997, the Bank of Thailand tried to strengthen the Thai financial institution system by implementing the Financial Sector Master Plan (FSMP)[24]. The major target is to improve efficiency of the system. It restructures the sector by allowing only two types of Thai financial institutions to operate: full-service banks and retail banks. Full-service banks (commercial banks) can carry out all types of financial transactions, except insurance underwriting as well as brokering, trading and underwriting of equity securities. Although retail banks can provide a similar set of financial transactions, they are not permitted to conduct business related to foreign exchange and derivatives products. Also, their services are subject to conditions set by the Bank of Thailand, such as lending limit per customer. Foreign financial institutions can also have only two types of licenses: subsidiaries and single branch.[25] Both types of foreign financial institutions are able to undertake some activities as those of full-service banks. The only difference is that a subsidiary can open up to four offices, one in Bangkok and three outside, while single branches can only have one.[26] Moreover, as part of the measures to increase efficiency of the

[24] Bank of Thailand (2006) and Watanagase and Financial Institutions Policy Group, Bank of Thailand (2006)
[25] Instead of the word “single branch”, the Bank of Thailand calls it “full branch”.
[26] The minimum capital requirement for full-service banks, retail banks, subsidiaries, and single branches are 5 billion baht, 250 million baht, 4 billion baht and 3 billion baht respectively.

sector, the “one presence” policy is introduced. Instead of allowing many deposit-taking financial institutions within their financial group to reflect the artificial regulatory distinction between commercial banks, financial companies, and credit foncier companies, one presence policy will remove distinction in the scope of business by permitting them to have only one license of deposit-taking institution for a financial conglomerate, e.g., Thai bank,[27] subsidiary, or single branch. The FSMP started implementing at the beginning of 2004. By merging and upgrading the finance companies and credit foncier companies[28] to full-service or retail bank status, discontinuation of International Banking Facilities (IBFs) and consolidation under the “one presence” policy, the number of financial institutions[29] in Thailand have been declining from 83 at the end of 2003 to 44 in 2007. Basically, the FSMP encourages all small financial institutions to merge each other and expect that they will operate more efficiently. In this paper, therefore, I will focus on Thai banks to see whether they succeed in improving their efficiency and productivity by comparing their performances before and after the FSMP.

5.2 Data

Due to the availability of data and avoiding any complexity during the crisis period originated in Thailand (1997) and the US (2007), seventeen banks in the period of 2001[30] to 2006 are analyzed in this paper. They are categorized by size using market share of total

---

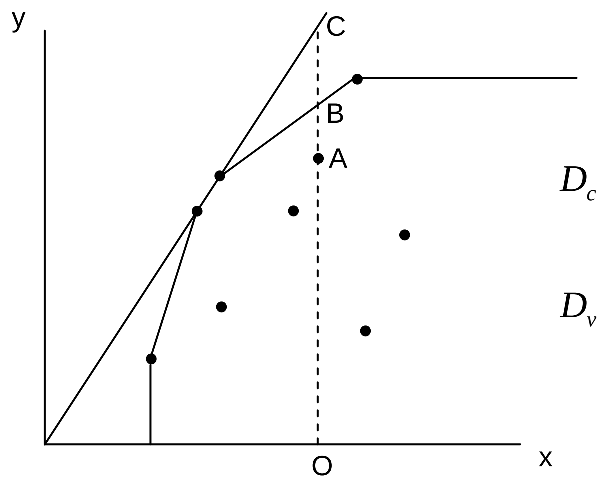
[27] By merging between foreign and Thai companies, foreign companies could transform to a Thai bank.
[28] Business of credit foncier is mainly to lend money on the security of mortgage of immovable property and to buy immovable property under contract of sale with a right of redemption.
[29] Include only financial institutions that are under the supervision of the Bank of Thailand
[30] For the Malmquist index as will be described below, the data in year 2000 is used to calculate the index for the year 2001.

assets as shown in Table 5.1. Apparently, the target group of the policy seems to be small companies rather than large companies. We can see three small banks disappear (uobr merged with uobt, dtdb merged with tmb and bmb merged with scib). Also, there are four small newborn banks, namely tbank, tisco, kk, and acl. Moreover, under the one presence policy, uobt and scbt merged their smaller companies into one company in 2005.

5.3 Methodology

To measure efficiency, DEA has been widely used in many areas, and has become a standard tool especially when the number of observations is limited. Since it is a nonparametric approach, there is no need to pre-assume any functional form. To show the degree of efficiency, an output-oriented DEA would measure how much output quantities can be proportionally expanded without changing the input quantities used. The DEA score will range from zero to one. For example, if a firm's score is equal to 0.8, the firm should be able to increase the level of output by 20 percent.

Conceptually, DEA is a mathematical program which envelopes the observed data from above to determine a best-practice frontier, and measures efficiency level by calculating distant ratio from the frontier. For instance, output-oriented DEA score for firm A can be illustrated in Figure 5.1.



$$D_c(x_A, y_A) = \frac{OA}{OC}$$

$$D_v(x_A, y_A) = \frac{OA}{OB}$$

Figure 5.1 Output oriented DEA

In output-input space, there could be two types of frontiers, constant returns to scale (CRS) and variable returns to scale (VRS). The constant returns to scale frontier is a ray from the origin whereas the variable returns to scale frontier is a convex piecewise linear covering data cloud. D(.) is a distant ratio comparing a level of output produced to the frontier. In this paper, variable returns to scale frontier is chosen to relax the assumption of constant returns to scale. The distant ratio can be calculated as in equation (5.1).

$$
\begin{aligned}
&\mathrm{Max}_{\phi,\lambda}\phi_j = D(X_j, Y_j)^{-1} \\
&\text{st.} \quad -\phi_j y_{kj} + \sum_{j=1}^{J} y_{kj}\lambda_j \geq 0 \qquad && k = 1,\ldots,n \\
&\qquad x_{ij} - \sum_{j=1}^{J} x_{ij}\lambda_j \geq 0 && i = 1,\ldots,m \qquad\qquad (5.1) \\
&\qquad \lambda_j \geq 0 && j = 1,\ldots,J \\
&\qquad \sum_{j=1}^{J}\lambda_j = 1 && \text{if} \quad \text{VRS}
\end{aligned}
$$

where there are j firms which use input set $X = \{x_1, x_2, \ldots, x_i\}$to produce output set $Y = \{y_1, y_2, \ldots, y_k\}$.

Table 5.1
Market share of total assets

	2001	2002	2003	2004	2005	2006
LARGE BANK						
bbl	22.0	21.6	22.2	21.8	20.1	19.8
ktb	17.2	18.3	18.5	17.8	16.6	16.0
kbank	13.7	13.1	13.4	12.8	12.0	12.4
scb	12.6	11.7	12.2	11.9	11.7	13.7
MEDIUM BANK						
bay	7.7	8.1	8.5	8.9	9.3	8.8
tmb	6.5	6.8	6.2	10.4	10.3	9.9
scib	5.3	8.4	7.6	7.3	6.5	5.5
SMALL BANK						
bt	4.8	4.7	4.2	3.6	3.9	2.9
uobt	2.8	2.8	2.8	2.5	2.8	2.5
scbt	1.2	1.1	1.0	1.0	2.0	2.5
uobr	1.0	0.9	0.9	0.8		
dtdb	1.7	1.7	1.6			
bmb	3.4					
tbank			0.9	1.1	2.8	3.4
tisco						1.1
kk						1.0
acl						0.5

Note: Abbreviations stand for the following: bbl- Bangkok bank public company limited, ktb- Krung thai bank public company limited, kbank- Kasikorn bank public company limited, scb- Siam commercial bank public company limited, bay- Bank of ayudhya public company limited, tmb- TMB bank public company limited, scib- The siam city bank public company limited, bt- Bankthai public company limited, uobt- United overseas bank (Thai) public company limited, scbt- Standard chartered bank (Thai) public company limited, uobr- UOB radanasin bank public company limited, dtdb- DBS thai danu bank public company limited, bmb- Bangkok metropolitan bank public company limited, tbank- Thanachart bank public company limited, tisco- Tisco bank public company limited, kk- Kiatnakin bank public company limited, acl- ACL bank public company limited.

However, DEA is used for analyzing cross-section data only. Comparing DEA scores across time might be misleading due to the different benchmark overtime. Therefore, the well known Malmquist productivity index will be employed to capture the dynamic of productivity change. Let t equal time period. The Malmquist productivity index can be calculated as follows:

$$M_j(t_1, t_2) = \left(\frac{D_c^{t_1}(X_j^{t_2}, Y_j^{t_2})}{D_c^{t_1}(X_j^{t_1}, Y_j^{t_1})} \times \frac{D_c^{t_2}(X_j^{t_2}, Y_j^{t_2})}{D_c^{t_2}(X_j^{t_1}, Y_j^{t_1})} \right)^{\frac{1}{2}} \qquad (5.2)$$

As in equation (5.2), the index is just a geometric mean of two terms. Each term represents efficiency ratio comparing to different frontiers, one uses a frontier in period one as a benchmark but the other uses a frontier in period two. If the index is equal to one, there is no productivity change for that firm. The value greater (smaller) than one shows productivity improvement (decline). Note that the productivity index is equivalent to the ratio of the CRS distance functions even if the technology was not characterized by the constant returns to scale.[31]

5.4 Models

In general, there are two approaches to analyzing banking operations, intermediary approach and production approach. The major difference between the two is how to treat deposit money. The intermediary approach treats it as an input while it is an output in the production approach. In this paper, the intermediary approach is preferred due to the nature of banking business. Unfortunately, the number of observations is fairly limited. Therefore, only one output and two inputs are chosen in each model. There will be two models to compare in two different aspects. Model A employs quantity variables to show technical efficiency from production, whereas model B uses value variables to show X-efficiency or revenue

[31] Fare and Grosskopf (1994)

efficiency from allocation. While efficiency represents how much the output can be produced given a level of inputs, revenue efficiency indicates how much revenue can be generated given a level of expense. In other words, revenue efficiency shows profitability of a firm. In model A, two inputs are specified: 1) deposit and borrowing money; and 2) non-interest expense.[32] The only output in model A is loans and investment. In Model B, two inputs are specified: 1) interest expense; and 2) non-interest expense. One output is specified: interest income and non-interest income.

5.5 Result

As shown in Table 5.2, in terms of technical efficiency, the large bank outperforms the other two groups. Even though the small bank is ranked the third for most years, it catches up and becomes more efficient than the medium group in 2006. In general, firms having scores equal to one are efficient. However, in case of bold numbers, they not only have scores equal to one, but they also have lambda equal to zero in every firm's problem. In other words, they were not benchmarks for any firm.[33] Therefore, it is not clear whether they are efficient or not. Subsequently, they should be removed from average values in the lower part of Table 5.2.

[32] Due to a lack of quantity variable, such as number of labor, non-interest expense is used as a proxy.

[33] Because the VRS frontier is composed of many piecewise linear pieces, some pieces of them may never be compared with others, especially when they locate away from the data cloud. Unlike the case of CRS frontier or the Malmquist index, the frontier is a single linear line. Thus, all firms below the frontier will be compared to the same benchmark.

Table 5.2
DEA score (technical efficiency)

	2001	2002	2003	2004	2005	2006
LARGE BANK						
bbl	**1.000**	**1.000**	**1.000**	1.000	1.000	**1.000**
ktb	1.000	1.000	1.000	0.958	**1.000**	1.000
kbank	0.996	0.961	0.955	0.973	0.981	0.975
scb	1.000	0.979	0.980	1.000	1.000	1.000
MEDIUM BANK						
bay	0.958	0.928	0.947	0.903	0.914	**1.000**
tmb	1.000	0.932	0.947	1.000	1.000	0.898
scib	0.986	1.000	**1.000**	1.000	0.993	0.964
SMALL BANK						
bt	0.857	0.694	0.679	0.752	0.787	0.896
scbt	1.000	1.000	0.878	0.980	1.000	1.000
tbank			1.000	1.000	1.000	0.970
tisco						0.956
kk						0.990
acl						1.000
uobt	0.944	0.935	0.882	0.889	0.971	0.943
uobr	1.000	1.000	0.840	1.000		
dtdb	0.955	0.901	0.838			
bmb	0.881					
AVERAGE VALUE						
Large	0.999	0.980	0.978	0.983	0.994	0.992
Medium	0.981	0.953	0.947	0.968	0.969	0.931
Small	0.938	0.898	0.847	0.919	0.935	0.964

To examine the productivity change before and after the FSMP, the Malmquist index for both models A and B are shown in Tables 5.3 and 5.4. Since the plan mainly contributes to small banks, their productivity increases after the plan by six percent on average. However, it seems like both large and medium banks cannot adjust themselves well under the new circumstance. As a result, their productivities decline by four and fourteen percent, respectively. For model B, the index looks a bit different due to price effect.

The price effect as shown in Figure 5.2 is effective interest rate spread. It is calculated from the gap between effective loan rate and effective deposit rate. In general, all banks' deposit rates are highly correlated, above 0.9 for each pair of groups. However, due to their reputation[34] and significantly different number of branches,[35] large banks offer relatively low deposit rates compared to those of medium and small banks. Nonetheless, large and medium banks have similar types of loans, such as manufacturers and public utilities. Therefore, patterns of large and medium banks' spreads are quite similar, while the gap between the two is derived from their difference in deposit rates. Differently, small banks' spread is diverse, because they lend money to various types of loans, especially on personal loans and housing loans. However, one of the large banks recently moved its target to housing loans, so increasing competition causes drops in loan rates of small banks and spreads. See Appendix B for more details about the spread.

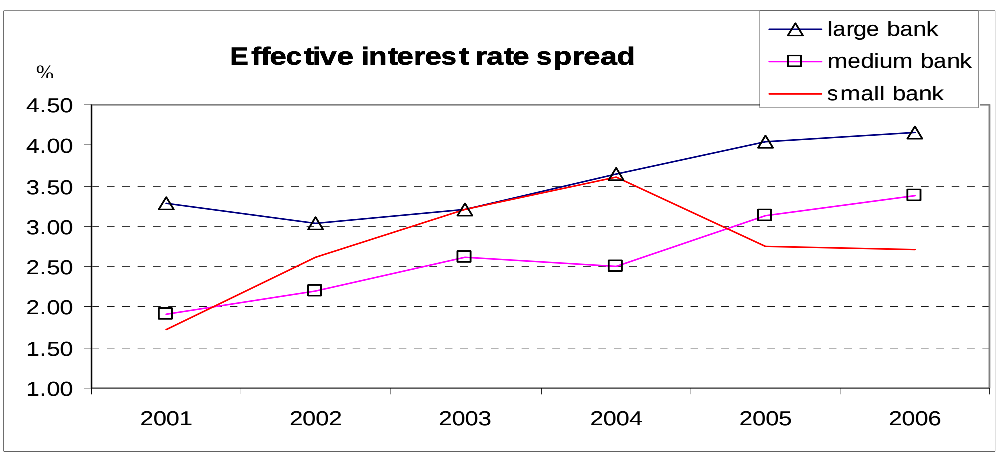


Figure 5.2 Effective interest rate spread

[34] None of the large banks have ever experienced a loss or the merging with lost companies as many of the medium and small banks have been through.
[35] On average, large, medium, and small banks have 591, 397, and 84 branches per bank, respectively.

Table 5.3
The Malmquist productivity index for model A

	2001	2002	2003	2004	2005	2006	2001-2003	2003-2006	2001-2006
LARGE BANK									
bbl	1.02	1.01	0.98	1.04	1.03	0.97	0.99	1.05	0.98
ktb	1.00	1.14	0.99	0.91	1.03	0.98	1.12	0.84	0.96
kbank	0.99	1.01	0.99	1.03	1.04	0.98	0.99	0.99	1.04
scb	0.99	1.00	1.00	1.04	1.02	0.99	1.00	0.99	0.98
MEDIUM BANK									
bay	1.04	1.07	1.01	0.94	1.02	0.98	1.04	0.96	1.00
tmb		0.99			1.05	0.82			0.89
scib		1.39	0.85	0.97	1.00	0.90	1.17	0.77	1.02
SMALL BANK									
bt	0.82	0.97	0.97	1.19	1.07	0.97	0.83	1.31	1.10
scbt	0.96	1.02	0.96	1.05	1.04	1.01	0.98	1.12	1.09
tbank				1.09	0.94	0.75		0.78	
uobt	0.95	1.01	0.97	1.00	1.10	0.98	0.98	1.10	1.07
uobr	0.98	1.03	0.92	1.11			0.95		
dtdb		1.00	0.98				0.98		
AVERAGE VALUE									
Large	1.00	1.04	0.99	1.00	1.03	0.98	1.02	0.96	0.99
Medium	1.04	1.14	0.93	0.96	1.02	0.90	1.10	0.86	0.97
Small	0.93	1.01	0.96	1.09	1.04	0.92	0.94	1.06	1.09[36]

Source: Author's calculation

From the fact that value is equal to quantity times price, if the interest rate spread increases (decreases), we can expect higher (lower) levels of the Malmquist index in model B. Although the small banks succeed in improving their productivity after the FSMP, they don't perform well in terms of profitability because of the price effect. Some problems seem to occur with large and medium banks. Yet, they have advantages from price effect, their indexes are still below one after the FSMP.

[36] The number of the overall period (2001-2006), 1.09, exceeds the range, 0.94 – 1.06, of pre and post period because of different coverages.

Table 5.4
The Malmquist productivity index for model B

	2001	2002	2003	2004	2005	2006	2001-2003	2003-2006	2001-2006
LARGE BANK									
bbl	1.38	1.03	1.04	1.11	1.04	0.90	1.01	1.12	1.18
ktb	2.36	1.35	0.98	0.96	1.11	0.84	1.39	1.01	1.49
kbank	1.06	1.15	1.22	1.06	1.03	0.75	1.44	0.93	1.42
scb	0.96	0.56	2.91	1.24	0.98	0.69	1.53	0.87	1.40
MEDIUM BANK									
bay	1.21	1.24	0.96	0.96	1.03	0.95	1.20	0.93	1.13
tmb[37]		0.94			1.35	0.55			0.63
scib		2.07	1.00	1.04	0.96	1.07	2.22	1.04	2.25
SMALL BANK									
bt	1.39	1.24	0.55	1.33	1.04	1.00	0.71	1.35	0.92
scbt	1.63	1.37	1.39	1.51	0.80	0.81	1.96	0.81	1.52
tbank				1.10	0.74	1.07		0.86	
uobt	1.08	1.19	2.22	0.92	1.07	0.77	2.64	0.77	2.14
uobr	1.44	1.26	1.21	1.16			1.56		
dtdb		1.04	0.52				0.52		
AVERAGE VALUE									
Large	1.35	0.97	1.38	1.09	1.04	0.79	1.33	0.98	1.37
Medium	1.21	1.34	0.98	1.00	1.10	0.82	1.63	0.99	1.17
Small	1.37	1.22	1.01	1.19	0.90	0.90	1.24	0.92	1.44

5.6 DEA and the Malmquist Index vs. Conventional Measures of Performance

There are many traditional measures of performance that people have been using for a long time. In this paper, Return on Asset ratio (ROA), Return on Equity (ROE), and Cost to Income ratio are picked to compare with DEA and the Malmquist index. Using correlation coefficient, we cannot apply it directly due to different nature of numbers. DEA is not comparable across years. Therefore, correlations between DEA and financial ratios will be calculated for each year and only the average values are reported. Differently, the Malmquist

[37] Because tmb's interest income is negative in 2003, the index for 2003 and 2004 cannot be calculated. Also, for comparison purposes, the index for model A in this period is ignored.

index infers growth which is not comparable with financial ratio. Therefore, the index needs to transform by multiplying the index overtime, so that it becomes level, called chained Malmquist index. Moreover, due to different starting points of the chained Malmquist index, they are not comparable across firms. Therefore, correlations will be calculated for each firm, and only the average values are reported.

Since return and cost are values which imply profitability, we can expect that correlation between those financial ratios and results from model A will be low. As shown in Table 5.5, only the chained Malmquist index from model B highly correlates with financial ratios. Apparently, efficiency and productivity measures from model A provide different information from traditional financial ratios.

Table 5.5
Correlation coefficient

	DEA	Chained Malmquist (A)	Chained Malmquist (B)
ROA	0.34	0.35	0.80
ROE	0.32	0.29	0.76
Cost to Income	-0.39	-0.36	-0.78

5.7 Conclusions

This paper attempts to measure banks' efficiency before and after the FSMP which is the only huge financial institution policy package during the recovery period of Thai economy. However, before interpreting all numbers, we should understand limitations of the methodology and models. First, in model A, the efficiency is loosely defined on only the amount of loans, deposits, and non-interest expense. It assumes constant risk across the board. Consequently, the score could be a misleading message. For instance, increasing in bank's reserve as well as a drop in loans for any good reason would always result in deteriorating of bank's efficiency. Moreover, in model B, price is exogenously determined by market structure. Increasing competition or any change in market structure could have effects to price and profitability of the firms. However, it does not necessarily mean that lower profits always leads to the lower level of efficiency.

According to the above limitations, if other things are equal, we may conclude that, given the definition and assumption described in this paper, the FSMP was at least successful in improving small banks' efficiency and productivity. However, because of a drop in the interest rate spread lately, small banks' profitability declines. For large and medium banks, they perform worse after the FSMP in both productivity and profitability aspects. Yet, they are still the most efficient group ranked by the DEA score.

CHAPTER VI

CONCLUSIONS

This dissertation is devoted to developing and understanding tools in efficiency analysis so as to appropriately deal with non-discretionary inputs. The major intended contribution is to propose an alternative multi-stage DEA approach which has a superior performance in term of rank correlation and MAD. Four essays are provided as follows:

My first essay is entitled "An Alternative DEA Methodology for Non-Discretionary Inputs." DEA is a standard methodology for assessing technical efficiency. In many DEA applications, e.g. the case of schools or hospitals, the issue arises of calculating efficiency in the presence of non-discretionary (environmental) inputs. I propose a multi-stage DEA model to address the environmental input issue, and I provide a simulation analysis that illustrates the implementation and potential advantages of my approach relative to the leading existing multi-stage model of non-discretionary inputs. Moreover, one of the most important findings in these experiments is that, for all approaches, CRS assumption seems to be preferred when firms have similar size, though they actually belong to decreasing returns to scale (in discretionary inputs) technology. However, VRS tends to be preferred when firms are different in size.

My second essay is entitled "A Note on the Statistical Inference of a Multi-Stage DEA Model with Non-Controllable Inputs." Simar and Wilson (2007) indicated that most of the multi-stage DEA models with non-discretionary inputs share the same serial correlation problem which causes the hypothesis testing to be invalid. Therefore, they proposed bootstrap methods to create a valid confidence interval. Besides, they suggested practitioners to use the truncated model instead of the tobit model, the most popular model in the literature. However, the bootstrap methods seem to differ from the traditional way of solving the serial correlation problem where the bootstrap process should preserve the structure of the error term dependency. Due to the concern of the methods' effectiveness, two sets of simulation experiments are performed to compare between the bootstrap (Algorithm I) and t-test, and also the truncated model and the tobit model in terms of size and power. Consequently, the simulation experiments show somewhat different results where the t-test is good at both size and power, and so as the tobit model.

My third essay is entitled "Application to Texas School Districts." It illustrates how to apply empirical data to my new approach. I compare results from two different models (one which includes LPR, the lagged average passing rate) using rank correlation and MAD. They both have very high rank correlation and low MAD. Therefore, the conclusion is that LPR, in the context of Texas school districts, contribute only a little to the model. Consequently, it might be excluded from the analysis if the student level data is not available. In addition, I find that the tobit and the truncated models return similar results.

The final essay is entitled "Bank Efficiency in Thailand: Pre and Post Financial Sector Master Plan." This essay attempts to measure banks' efficiency before and after the FSMP which is the only huge financial institution policy package during the recovery period of Thai economy from the crisis in 1997. Mainly, it encourages small financial companies to merge together and upgrade to banks. The FSMP was at least successful in improving small bank’s efficiency and productivity (given the ways I defined efficiency and productivity). However, because of a drop in the interest rate spread lately, small banks’ profitability declines. For large and medium banks, they perform worse after the FSMP in both the productivity and profitability aspect. Yet, they are still the most efficient group ranked by DEA score.

REFERENCES

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

Bank of Thailand, 2006. Thailand's Financial Sector Master Plan Handbook, Bangkok.

Banker, R., Morey, R., 1986. Efficiency analysis for exogeneously fixed inputs and outputs. Operations Research 34, 513-521.

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

Fried, H.O., Lovell, C.A.K., Schmidt, S.S., 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., Yaisawarng, S., 1999. Incorporating the operating environment into a nonparametric measure of technical efficiency. Journal of Productivity Analysis 12, 249-267.

Jensen, U., 2005. Misspecification preferred: the sensitivity of inefficiency rankings. Journal of Productivity Analysis 23, 223-244.

Jondrow, J., Lovell, C. A. K., Materov, I. S., Schmidt, P., 1982. On the estimation of technical inefficiency in the stochastic frontier production function model. Journal of Econometrics 19, 233–238.

McDonald, J., 2009. Using least squares and tobit in second stage DEA efficiency analyses. European Journal of Operational Research 197 (2), 792-798.

Meeusen, W., van den Broeck, J., 1977. Efficiency estimation from Cobb–Douglas production functions with composed error. International Economic Review 18, 435–444.

Muñiz, M.A., 2002. Separating managerial inefficiency and external conditions in data envelopment analysis. European Journal of Operational Research 143, 625-643.

Muñiz, M.A., Paradi, J., Ruggiero, J., Yang, Z.J., 2006. Evaluating alternative DEA models used to control for non-discretionary inputs. Computers & Operations Research 33, 1173-1183.

Ruggiero, J., 1996. Efficiency of educational production: an analysis of New York school districts. The Review of Economics and Statistics 78 (3), 499-509.

Ruggiero, J., 1998. Non-discretionary inputs in data envelopment analysis. European Journal of Operational Research 111, 461-469.

Ruggiero, J., 1999. Efficiency estimation and error decomposition in the stochastic frontier model: a Monte Carlo analysis. European Journal of Operational Research 115, 555-563.

Ruggiero, J., 2007. Non-discretionary inputs. In: Zhu, J., Cook, W.D. (Eds.), Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis. Springer, New York, pp. 85-101 (Chapter V).

Simar, L., Wilson, P.W., 2007. Estimation and inference in two-stage, semi-parametric models of production processes. Journal of Econometrics 136, 31-64.

Watanagase, T. and Financial Institutions Policy Group, Bank of Thailand, 2006. Maintaining financial stability: the experience of Thailand. BIS Papers No28, Basel, Switzerland.

Wheelock, D.C., Wilson, P.W., 1999. Technical progress, inefficiency, and productivity change in U.S. banking, 1984-1993. Journal of Money Credit and Banking 31 (2), 212-234.

Winsten, C. B., 1957. Discussion on Mr. Farrell's paper. Journal of the Royal Statistical Society A120, 282–284.

Supplemental Sources Consulted

Asava-vallobh, N., Gronberg, T.J., Jansen, D.W., 2008. An alternative DEA methodology for non-controllable (environmental) inputs. PERC Working Paper #0810, Texas A&M University, College Station.

Avkiran, N.K., 1999. The evidence on efficiency gains: the role of mergers and the benefits to the public. Journal of Banking & Finance 23, 991-1013.

Coelli, T., 1996. A guide to DEAP version 2.1: a data envelopment analysis (computer) program. CEPA Working Papers No.8/96, Department of Econometrics, University of New England, Armidale, Australia.

Coelli, T., 2005. An Introduction to Efficiency and Productivity Analysis. Springer, New York.

Gronberg, T.J., Jansen, D.W., Naufal, G.S., 2006. Efficiency and performance in Texas public schools. PERC Working Paper #0609, Texas A&M University, College Station.

Gronberg, T.J., Jansen, D.W., Taylor, L.L., 2007. The impact of facilities on the cost of education. Unpublished Working Paper, Texas A&M University, College Station.

Hanushek, E.A., Taylor, L.L., 1990. Alternative assessments of the performance of schools. Journal of Human Resources 25 (2), 179-201.

Ruggiero, J., Vitaliano, D.F., 1999. Assessing the efficiency of public schools using data envelopment analysis and frontier regression. Contemporary Economic Policy 17, 321-331.

Syrjänen, M.J., 2004. Non-discretionary and discretionary factors and scale in data envelopment analysis. European Journal of Operational Research 158, 20-33.

Zhu, J., 2003. Quantitative Models for Performance Evaluation and Benchmarking : Data Envelopment Analysis with Spreadsheets. Kluwer Academic Publishers, Boston.

APPENDIX A

ADDITIONAL SIMULATION EXPERIMENTS

The below additional experiments are to supplement the base case and the second case (firms with different scale) in Chapter II. Most of the results in Chapter II show that all models (222, 114, 141, and 411) perform similarly in each case, so I depict only the first model (222) in this section.

1) Supplement to the base case

All settings in this case are the same as the base case in Chapter II except the distributions of discretionary inputs and non-discretionary inputs: $x_i \sim$ Uniform[1,2] $\forall i$ and $z_l \sim$ Uniform[30,50] $\forall l$. The discretionary inputs' ranges are shortened, while the non-discretionary inputs' ranges are widened compared to the base case in Chapter II. However, results in Table A1 and A2 still have similar patterns as in Table 2.1 and 2.2 where constant returns to scale assumption and my alternative approach are preferred in most cases.

Table A1
Simulation results: supplement to the base case (constant returns to scale)

CRS		Number of firm: 50		Number of firm: 100		Number of firm: 150	
Model	Approach	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD
222	Fried et al.	**0.8323**	0.0632	0.8520	0.0557	0.8603	**0.0539**
	Ruggiero	0.6825	0.1067	0.7681	0.0843	0.7943	0.0745
	New (TR)	0.8132	0.0639	0.8548	0.0603	0.8705	0.0596
	New (TO)	0.8259	0.0575	**0.8630**	0.0571	**0.8750**	0.0582
	New (DF)	0.8308	**0.0563**	0.8627	**0.0555**	0.8745	0.0567
	New (SF)	0.8280	0.0647	0.8536	0.0583	0.8664	0.0545

Table A2
Simulation results: supplement to the base case (variable returns to scale)

VRS		Number of firm: 50		Number of firm: 100		Number of firm: 150	
Model	Approach	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD
222	Fried et al.	0.7068	0.0975	0.7697	0.0742	0.7964	0.0653
	Ruggiero	0.4986	0.1646	0.6173	0.1455	0.6671	0.1341
	New (TR)	0.6786	0.0788	0.7929	0.0601	0.8339	0.0547
	New (TO)	0.7139	0.0687	0.8176	0.0594	**0.8504**	0.0575
	New (DF)	**0.7327**	**0.0679**	**0.8178**	**0.0530**	0.8476	**0.0505**
	New (SF)	0.7062	0.0966	0.7699	0.0739	0.7967	0.0652

2) Supplement to the second case I

In Chapter II, the difference in scale is defined as two separate ranges of discretionary inputs. However, scale difference could be defined differently. Following the above set up but creating scale difference by two overlapping ranges of discretionary inputs, the distributions of discretionary inputs and non-discretionary inputs are the following: x_i ~ [1,2] $\forall$i and [1,4] and z_l ~ Uniform[30,50] $\forall$l.

Table A3 and A4 still support the results in Chapter II which conclude that variable returns to scale assumption is preferred when data is spread out.

Table A3
Simulation results: supplement to the second case I (constant returns to scale)

CRS		Number of firm: 50		Number of firm: 100		Number of firm: 150	
Model	Approach	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD
222	Fried et al.	0.6414	**0.1033**	0.6661	**0.1086**	0.6684	**0.1141**
	Ruggiero	0.5219	0.1148	0.5821	0.1066	0.6075	0.1044
	New (TR)	0.6284	0.1125	0.6650	0.1187	0.6719	0.1241
	New (TO)	0.6304	0.1089	0.6669	0.1177	**0.6732**	0.1245
	New (DF)	**0.6322**	0.1074	**0.6670**	0.1166	0.6731	0.1235
	New (SF)	0.6335	0.1098	0.6584	0.1239	0.6493	0.1392

Table A4
Simulation results: supplement to the second case I (variable returns to scale)

VRS		Number of firm: 50		Number of firm: 100		Number of firm: 150	
Model	Approach	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD
222	Fried et al.	**0.7065**	0.1058	0.7633	0.0808	0.7889	0.0708
	Ruggiero	0.5168	0.1654	0.6361	0.1463	0.6871	0.1358
	New (TR)	0.6582	0.0813	0.7709	0.0632	0.8216	0.0556
	New (TO)	0.6815	**0.0733**	0.7916	0.0594	**0.8359**	0.0559
	New (DF)	0.7076	0.0751	**0.7953**	**0.0557**	0.8340	**0.0507**
	New (SF)	0.7052	0.1051	0.7633	0.0808	0.7889	0.0708

3) Supplement to the second case II

To investigate more in the scale difference case, this section defines scale difference as a wider range of discretionary inputs where x_i ~ U[30,100] $\forall$i and z_l ~ Uniform[1,2] $\forall$l. As a result, Table A5 and A6 still have the similar pattern as before. Though I've defined various types of scale difference, their results are quite consistent. In other words, employing my approach assuming variable returns to scale tends to be the best way to deal with scale-difference firms. However, it is worth noting that constant returns to scale assumption could be preferred when there are not many observations.

Table A5
Simulation results: supplement to the second case II (constant returns to scale)

CRS		Number of firm: 50		Number of firm: 100		Number of firm: 150	
Model	Approach	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD
222	Fried et al.	0.7115	0.0832	0.7373	**0.0834**	0.7439	**0.0862**
	Ruggiero	0.5791	0.1146	0.6556	0.0949	0.6788	0.0879
	New (TR)	0.7091	0.0870	0.7474	0.0933	0.7582	0.0964
	New (TO)	0.7144	0.0842	**0.7508**	0.0923	**0.7606**	0.0976
	New (DF)	**0.7168**	**0.0821**	0.7507	0.0903	0.7600	0.0958
	New (SF)	0.7082	0.0894	0.7396	0.0921	0.7541	0.0931

Table A6
Simulation results: supplement to the second case II (variable returns to scale)

VRS		Number of firm: 50		Number of firm: 100		Number of firm: 150	
Model	Approach	Rank Correlation	MAD	Rank Correlation	MAD	Rank Correlation	MAD
222	Fried et al.	0.6411	0.1072	0.7037	0.0854	0.7301	0.0774
	Ruggiero	0.4716	0.1738	0.5936	0.1567	0.6484	0.1461
	New (TR)	0.6256	0.0843	0.7488	0.0699	0.7921	0.0677
	New (TO)	0.6562	0.0781	**0.7681**	0.0745	**0.8058**	0.0758
	New (DF)	**0.6738**	**0.0744**	0.7643	**0.0644**	0.7989	**0.0648**
	New (SF)	0.6406	0.1060	0.7041	0.0850	0.7325	0.0766

APPENDIX B

INTEREST RATE SPREAD

Table B1
Deposit rate, loan rate, and interest rate spread

	2001	2002	2003	2004	2005	2006	Correlation
Interest Rate Spread							
Large	3.28	3.04	3.20	3.64	4.04	4.16	0.87[1]
Medium	1.91	2.19	2.61	2.51	3.12	3.38	0.37[2]
Small	1.73	2.61	3.20	3.61	2.75	2.71	0.16[3]
Effective Deposit Rate							
Large	2.30	1.91	1.30	0.93	0.93	2.08	0.96[1]
Medium	2.74	2.27	1.76	1.13	1.34	3.02	0.98[2]
Small	2.97	2.49	1.89	1.49	1.42	3.65	0.92[3]
Effective Loan Rate							
Large	5.58	4.95	4.50	4.57	4.97	6.24	0.90[1]
Medium	4.65	4.46	4.37	3.64	4.46	6.40	0.72[2]
Small	4.70	5.11	5.09	5.10	4.17	6.36	0.57[3]
Policy Rate							
Fed (US)	1.75	1.25	1.00	2.25	4.25	5.25	0.97[4]
RP-14 (Thai)	2.46	1.75	1.25	1.90	3.94	5.00	

Remark: 1 represents correlation between large banks and medium banks
2 represents correlation between medium banks and small banks
3 represents correlation between large banks and small banks
4 represents correlation between federal funds rate and the Thai policy rate, RP-14
Source: author's calculation

In Table B1, movements of the Thai interest rates generally follow the Thai policy rate, RP-14, which highly correlates with the federal funds rate. The spread is calculated as follows:

$$\text{Interest Rate Spread}_i = \frac{\sum_{j|i} \text{interest on loans}}{\sum_{j|i} \text{net loans and accrued interest receivable}} - \frac{\sum_{j|i} \text{interest on deposits}}{\sum_{j|i} \text{deposits}}$$

where i is group (large, medium, small) and j is an individual bank in each group

The first term and the second term represent effective loan rate and effective deposit rate respectively. Interest on loans and interest on deposits are from the statements of income for each year, whereas net loans and deposits are from the balance sheets at the end of each year. In particular, the net loans are as follows:

Net loans = total loans - allowance for doubtful accounts - Revaluation allowance for debt restructuring

However, the effective loan rate could have irregular movement sometimes. For instance, in 2004, the medium banks' loan rate dropped while the policy rate increased. This is mainly because one of the medium banks, tmb, merged with a nonlucrative company in that year. As a result, its loans considerably increased, but its interest revenue still remained about the same.

VITA

Name: Norabajra Asava-vallobh

Address: Fiscal Policy Office, Ministry of Finance, Rama 6 Rd., Phayathai, Bangkok, Thailand

Email Address: norabajra@hotmail.com

Education: B.A., Economics, Thammasat University, Bangkok, Thailand, 1999

M.A., Economics, Thammasat University, Bangkok, Thailand, 2003

Ph.D., Economics, Texas A&M University, College Station, TX, 2009

Fields of Specialization:

Efficiency Analysis, International Economics, Financial Economics, and Econometrics

Experience (at Texas A&M University, Fall 2005 – Spring 2009):

Instructor: Microeconomic Theory (Undergraduate), Spring 2009

Research Assistant, Department of Economics, Spring 2008 – Fall 2008

Student Supplemental Instruction Assistant: Fundamental Mathematics for Economists (Graduate), Fall 2007

Reference:

Prof. Dennis Jansen, Dept. of Economics, Texas A&M University, College Station, TX, Tel: (979) 845-7375, Email: d-jansen@tamu.edu