# Ranking the Efficiency of Small-size Secondary Schools: A Case Study of Amnatcharoen Province

***Jiradate Thasayaphan****

**Abstract**

*The performance of secondary education in this study was ranked by computing technical efficiency scores. Efficiency was predicted by including discretionary inputs (capitation grants, fundamentally-needed funds, average teacher salary) compared to the output (Ordinary National Educational Test results) from 27 small-size schools out of a total of 54 located in Amnatcharoen province in Thailand. The semi-parametric models for computing the technical efficiency scores included three-stage procedures and a bootstrap algorithm. The parametric models included stochastic frontier analysis and a Bayesian stochastic frontier analysis model. All models took the selected institutional variables (politician involvement, class size, board school meeting, and inspection) into account within the service delivery framework, in particular, by regressing the initial data envelopment analysis (DEA) output scores on these institutional variables in the semi-parametric model. The results were compared to the results of the non-statistical assessment procedures conducted by the Office of National Education Standards and Quality (Public Organization) or ONESQA. The paper shows that the schools' ranking among the models were different. However, none of these methods provided consistent ranking results for school accreditation using the ONESQA procedures. It is suggested from the results that it would be interesting to explore the possibility to assess schools using statistical methods and to create assessment criteria to fail some of the schools that have low technical efficiency.*

***Keywords:*** *Education Production, Technical Efficiency, Service Delivery Framework*

---

*Ph.D. (Econ) , National Institute of Development Administration, Policy and Planning Analyst, The Office for National Education Standards and Quality Assessment (Public Organization), 24th Floor, Phaya Thai Plaza Building, 128 Phaya Thai Road, Rajthevee, Bangkok, THAILAND 10400.

# การจัดลำดับประสิทธิภาพการจัดการศึกษาของโรงเรียนชั้นมัธยมศึกษาในจังหวัดอำนาจเจริญ

***จิระเดช ทัศยาพันธุ์****

## บทคัดย่อ

*บทความนี้ศึกษาผลการจัดการศึกษาของโรงเรียนมัธยมศึกษาขนาดเล็กในจังหวัดอำนาจเจริญ โดยจัดลำดับตามคะแนนประสิทธิภาพทางเทคนิค ปัจจัยนำเข้าที่ควบคุมได้คือ ค่าใช้จ่ายรายหัว เงินอุดหนุนปัจจัยพื้นฐาน และเงินเดือนเฉลี่ยของครู ในขณะที่ผลผลิตคือ คะแนนผลสัมฤทธิ์จากการทดสอบของนักเรียนในโรงเรียนขนาดเล็ก 27 โรงเรียน จากประชากร 54 โรงเรียน การวิเคราะห์ด้วยแบบจำลองกึ่งสถิติคือและสถิติ และอิทธิพลของปัจจัยเชิงสถาบัน (การมีส่วนร่วมของนักการเมือง ขนาดของชั้นเรียน การประชุมของคณะกรรมการสถานศึกษา และการตรวจเยี่ยมจากส่วนกลาง) ภายใต้กรอบแนวคิดการให้บริการภาครัฐ ผลการจัดลำดับประสิทธิภาพทางเทคนิคจะนำไปเปรียบเทียบกับการจัดลำดับโดยใช้คะแนนจากการประเมินโรงเรียนโดยสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) หรือ สมศ. การศึกษาชี้ว่าในแต่ละแบบจำลองให้ผลการจัดลำดับที่มีความแตกต่างกัน อย่างไรก็ตาม ไม่มีวิธีใดให้ผลการจัดลำดับโรงเรียนใกล้เคียงกับการเรียงลำดับคะแนนด้วยวิธีการประเมินของ สมศ. ข้อเสนอแนะเชิงนโยบายคือ ควรจะมีการพัฒนาวิธีการรับรองคุณภาพโรงเรียนโดยใช้แบบจำลองทางสถิติ และกำหนดให้โรงเรียนที่มีประสิทธิภาพในลำดับท้ายจำนวนหนึ่งเป็นโรงเรียนที่ไม่ผ่านมาตรฐาน*

***คำสำคัญ*** *ผลผลิตทางการศึกษา ประสิทธิภาพทางเทคนิค กรอบแนวคิดการให้บริการภาครัฐ*

* Ph.D. (Econ) สถาบันบัณฑิตพัฒนบริหารศาสตร์ นักวิชาการวิเคราะห์นโยบายและแผน สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) ชั้น 24 อาคารพญาไทพลาซ่า เลขที่ 128 ถนนพญาไท แขวงทุ่งพญาไท เขตราชเทวี กรุงเทพฯ 10400

## Introduction

The concept of the function of education production assumes that the individual school maximizes the educational outcome of its students given their budget constraints. In a general sense, education provision is efficient if its producers make the best possible use of available inputs, and the sole fact that education inputs weigh heavily on public expenditures would call for a careful efficiency analysis. It is believed that education is the an important source of human capital formation, and productivity, as suggested by economic theory (Krueger and Lindahl, 2001). In this paper, the output from 27 small-sized lower secondary schools (so-called expand-opportunity schools) with resource utilization was systematically compared. However, most of the schools were found to perform below the frontier, and estimation of the distance that each school was from that line provided a so-called "technical efficiency score."

However, there is no unique measure of educational outputs that is perfect. Schools use instructional and non-instructional inputs that sometimes the school manager cannot control. School outputs that are associated with achievement scores are generally measured in terms of the student-teacher ratio, teaching experience, and various expenditures. Non-school inputs are environmental or exogenous factors, including the socioeconomic status, which include family income and parental education. These institutional factors are often measured by geographical location, the heterogeneity of students, and the schools' operating environments. For example, Barro and Lee (2001) found that student performance was positively correlated with the level of school resources, such as pupil-teacher ratios, and also with family background (income and education of parents). Additionally, Hanushek and Kimko (2000) and Hanushek and Luque (2003) found little or no evidence of a positive link from more resources allocated to the education system and test performance. However, they found that adult schooling levels have a positive and significant effect on student performance.

Fried et al. (2002) introduced a three-stage procedure. In the first-stage, data envelopment analysis (DEA) was applied to obtain the technical efficiency scores. In the second stage, stochastic frontier analysis (SFA) was used to regress the first stage performance measures against a set of environmental variables. This provides, for each input or output, a three-way decomposition of the variation in performance into a part attributable to environmental effects, a part attributable to managerial inefficiency, and a part attributable to statistical noise. During the third stage, either inputs or outputs were adjusted to account for the impact of the environmental effects and the statistical noise uncovered in the second stage, and DEA was used to re-evaluate producer performance. The analysis emphasis was placed on slacks as appropriate measures of producer performance. However, the fact that the DEA output scores were likely to be biased (Simar and Wilson, 2000), and that the environmental variables were correlated with output and input variables, the use bootstrapping techniques was recommended. Simar and Wilson (2004) have suggested that multi-stage estimation procedures, where nonparametric estimates of productive efficiency are obtained in the first stage and then regressed on environmental variables in a subsequent stage, may have unknown serial correlations among the estimated efficiencies. The data-generating processes (DGP), wherein firms' efficiencies were reinfluenced by environmental variables. The double bootstrap procedure improved statistical efficiency during the second-stage regression.

There are also arrays of techniques available for estimating efficiency scores, including parametric SFA and Bayesian Stochastic Frontier Analysis (BSFA). The BSFA is a relatively recent methodological development, with a limited number of applications in the literature to date. The fact that there are many alternative methods has meant that applied researchers across a vast range of different problem settings have sought guidance from the literature in order to find the appropriate methodology to employ. Numerous papers have compared the results generated by various frontier methodologies (e.g., Ahmad and Bravo-Ureta, 1996; Borger and Kerstens, 1996; Sharma et al., 1997; Cummins and Zi, 1998; Bauer et al., 1998; Kim and Schmidt, 2000; and Gocht and Balcombe, 2006). These comparisons add to the extensive literature that has compared the relative advantage and disadvantage of these techniques.

Therefore, there has been an effort to help with the transition of the education sector and frontier methods to provide a suitable methodology. Monte Carlo studies (Giannakas et al., 2003) can cast light, for example, on the performance of different methods under alternative DGPs. Albeit subject to different assumptions regarding the DGP, disagreement across methods may lead to more tentative conclusions. Finally, Sickles (2005) has suggested that a form of model averaging used to interpret efficiency estimates generated by multiple methods should be employed.

The rationale for examining lower secondary schools in Amnatcharoen is that the education system is decentralized and administrative control is deconcentrated. As a result, there is not a great deal of pressure on schools to be efficient. As stipulated in Section 81 of the 1997 Constitution of the Kingdom of Thailand, a national education law was required; hence the drafting of the 1999 National Education Act, which became effective on August 20, 1999. Chapter 6 of the Act on Education Standards and Quality Assurance mandates establishment of the Office for National Education Standards and Quality Assessment (ONESQA), enjoying the status of a public organization. It has been well recognized that evaluation is indeed and essential steps are required for feedback information, which will provide the basis for assessing the extent of the achievement of educational standards. ONESQA also identified weaknesses or problems for which remedial measures were needed so as to facilitate subsequent planning and actions required to achieve the goals effectively and efficiently. The results of the application examples using survey data applied to these techniques will be compared to those arising from the available results of the ONESQA assessment reports. In short, the main objective of this paper is to test the model using parametric and non-parametric methods for distinguishing the poor performers from the good performers. It also attempts to identify the causes of school inefficiency.

The structure of this paper is as follows. In section 2, the conceptual framework is described. In Section 3 the method used for estimation is discussed. In Section 4 the data set and variables used in the model are explained and in Section, the results of the efficiency analysis are portrayed. Finally, Section 6 contains the conclusions.

## Conceptual Framework

The accountability relationship, which is a set of relationships among service delivery actors, is comprised of five features (World Bank, 2003): *delegating* consists of the explicit or implicit understanding that a service (or goods embodying the service) will be supplied; *financing* consists of the providing of the resources to enable the service to be provided or paid for; *performing* consists of the supplying of the actual service; *having information about performance* consists of the ability to obtain relevant information and to evaluate performance against expectations, regarding either formal or informal norms; and *enforcing* consists of the ability to impose sanctions for inappropriate performance or to provide rewards when performance is appropriate. For example, a typical employment relationship forming an accountability relationship is as follow: a person is given a set of tasks (*delegation*) and is paid a salary (*finance)*, and the employee works (*performance*). The contribution of the employee is assessed (*information*) and based on that information, the employer acts to reinforce good or to discourage bad performance (*enforceability*). In public service, in the case of *finance*, however, the actors may be controlled by another who is not actually paid a salary. This situation will not create an accountability relationship; hence, to be a “stakeholder” it is necessary to put up a stake.

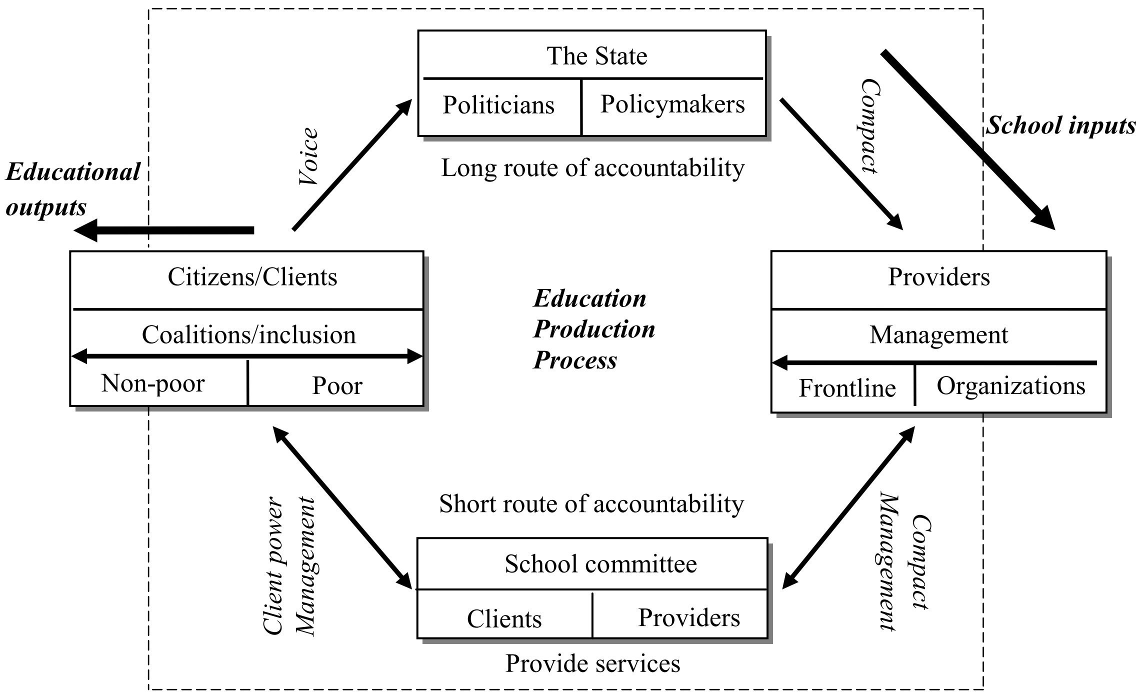


**Figure 1** School-Based Management and Accountability Relationships
**Source:** Adapted from Barrera-Osorio et al. (2009)

The four actors in the chain of public service delivery (education production process) are defined as follows: (*i*) *citizens and clients*: as citizens, they participate both as individuals and through coalitions in the political process and they define collective objectives; they also strive for control and direct public action in accomplishing those objectives. As direct clients of service providers, individuals and households hope to obtain quality public service; *ii*) *politicians and policymakers*: politicians derive and control state power and discharge fundamental responsibilities. The other actors that exercise the power of the state are policymakers. In general, politicians set general directions, but policymakers set the fundamental "rules of the game" for public service providers to operate; *iii*) *organizational providers*: a provider organization can be a ministry, department, or agency. It can be large (ministries with tens of thousands of teachers) or small (a single community-run primary school). The policymakers set and enforce the rules of the game of organization providers and the head of the provider makes "internal policies" specific to the organization; (*iv*) *frontline professionals*: all services require a provider that comes into direct contact with clients, including teachers, doctors, nurses, and so on.

The four actors have complex accountability relationships among them. These accountability relationships are explained as follows: (*i*) *voice* used to express the complex accountability relationships among citizens and politicians. Voice is about politics, but it covers the relationship of formal political mechanisms and informal ones. Delegation and finance between citizens and the state concern decisions about pursuing collective objectives and mobilizing public resources to meet those objectives. Citizens need information in order to understand which actions on the part of the state can promote their welfare. At the same time, if politicians and policymakers do not pursue objectives effectively, citizens will need some mechanisms with which to make them accountable; (*ii*) a *compact* expresses the relationships among policymakers and service providers. This does not mean legally enforceable as a contract. It is a broad agreement about a long-term relationship. Policymakers provide resources and delegate powers to the service providers and receive reports on the organization's performance in return. When the compact specifies rewards and penalties that depend on the actions and outputs of the service providers, enforceability may come into play; (*iii*) *management* provides frontline professionals with assignments and delineated areas of responsibility. In public agencies, this management function is not clear in comparison with the private sector, because providers are employees of "the government;" however, general management functions, including selecting, training, and motivating, still apply. All types of organizations have to create an accountability relationship with their frontline professionals; (*iv*) *client power* is the services that citizens demand of the service providers. Citizens also monitor the provider's services provision. Clients and organizational providers, such as frontline professionals and workers, interact through the individuals that provide services.

Weakness in any accountability relationships will lead to the weak institutional capacity of the actors, which can result in service failures. In our case, the school-based management (SBM) framework has the potential to hold school-level decision makers accountable for their actions. At the same time, it may be necessary to build the institutional capacity of the community members, teachers, and principals in order to create a culture of accountability. Theoretically, for any individual service transaction to be successful, there need to be frontline professionals that are capable, that have access to adequate resources and inputs, and that are motivated to pursue achievable goals. However, there is an important question: what institutional conditions support the emergence of capable, motivated frontline professionals with clear objectives and resources? The answer could be: successful public services for poor people emerge from institutional relationships in which the actors are accountable to each other (World Bank, 2003; p.46). In order to test the institutional relationships in the SBM framework that would affect the school outputs, efficiency measurement techniques were introduced to employ this tasks.

## Methods

The methods used for estimation included parametric and non-parametric methods. The derivation of each model is described in the appendix.

### *Data envelopment analysis framework*

The DEA approach was constructed using linear programming methods. Following Banker et al. (1984), the DGP of variable return to scale (VRS) output-orientated measure of technical efficiency (*TE*) can be express as:

$$\max_{\phi,\lambda} \phi, \tag{1}$$
$$\begin{aligned} \text{st} \quad & -\phi y_i + Y\lambda \geq 0, \\ & x_i - X\lambda \geq 0, \\ & I1'\lambda = 1 \\ & \lambda \geq 0, \end{aligned}$$

where $\phi$ is a scalar,

$y_i$ and $x_i$ are column vectors of outputs and inputs for the *i*-th school respectively.

$\lambda$ is a $N \times 1$ vector of constants,

*Y* is an $M \times N$ output matrix,

*X* is a $K \times N$ matrix in which $1 \leq \phi < \infty$,

$\phi - 1$ is the proportional increase in outputs that could be achieved by the *i*-th firm, with the fixed inputs,

and $I1$ is a $I \times 1$ vector of ones.

### *Adjusted DEA: A three-stage procedure*

Following Fried et al. (2002), the three-stage technique was used for incorporating environment effects and statistical noise into producer performance evaluation based on the DEA. The adjusted DEA can be calculated using the following equation:

$$y_{ni}^{A} = y_{ni} + [\max_i \{z_i\hat{\beta}^n\} - z_i\hat{\beta}^n] + [\max\{\hat{v}_{ni}\} - \hat{v}_{ni}], n = 1,...,N, i = 1,...,I, \tag{2}$$

where $y_{ni}^{A}$ and $y_{ni}$ are adjusted and observed output quantities, respectively,

$z'_{ni}$ is a transpose vector of institutional factors,

$\beta$ is a $((L+1)\times 1)$ transpose vector of parameters,

$v_{ni}$ is statistical noise, identically and independently distributed with zero mean and constant variance $\sigma^2$.

The first adjustment on the right side of the equation puts all schools into a common operating environment, which is the least favourable environment. The second adjustment puts all schools into a common state of nature, which was the unluckiest situation encountered. In the third stage, outputs were adjusted to account for the impact of the environmental effects and statistical noise uncovered in the second stage, and the DEA was used to re-evaluate producer performance.

***Stochastic Frontier Analysis***

Following Battese and Coelli (1992), $TE_i$ is the output-orientated technical efficiency of producer *i*, and can be expressed as:

$$TE_i = \frac{y_i}{f(\mathrm{x}_i;\beta)\cdot \exp(\nu_i)}, \quad (3)$$

where $y_i$ is output and $x_j$ are inputs,

$\nu_i$ standard white-noise disturbance,

$\beta$ is a vector of unknown parameters.

To explore the sources of inefficiency, Battese and Coelli (1995) extend the model (3) which defines $u_i$ as:

$$u_i = z_i\delta + c_i \quad (4)$$

where $z_i$ is a $(1 \times M)$ vector of institutional variables associated with technical inefficiency,

$\delta$ is a $(M \times 1)$ vector of unknown parameters, and $c_i$ is the non-negative unobserved random variable obtained by truncation of the $c_i \sim N^+(0, \sigma_c^2)$ such that $c_i \geq -z_i\delta$.

The testing of this model indicated that the technical inefficiency effects were not present in the model, expressed as $H_0 : \gamma = 0$, where $\gamma = \sigma_u^2 / (\sigma_\nu^2 + \sigma_u^2)$.

***Bootstrap and Non-discretionary inputs***

Simar and Wilson (2004) proposed two algorithms to achieve two stages, which are presented below.

The first algorithm involves the following steps:

Step 1 The computation of $\hat{\phi}_i$, for all *n* decision units by solving problem (1);

Step 2 The estimation of output surplus by maximum likelihood, considering it was a "truncated" regression (and not a Tobit regression), denoted by $\hat{\beta}$ and $\hat{\sigma}_{\varepsilon}$ which are the maximum likelihood estimates of $\beta$ and $\sigma_{\varepsilon}$;

Step 3 The computation of *L* the bootstrap estimation of $\beta$ and $\sigma_{\varepsilon}$ was made in the following way. For *i* = 1,…,*n* draw $\varepsilon_i$ from a normal distribution with variance $\hat{\sigma}_{\varepsilon}^2$ and right truncation at $1 - z_i\hat{\beta}$ and compute $\phi_i^* = z_i\hat{\beta} + \varepsilon_i$. Estimate the truncated regression of $\phi_i^*$ on $z_i$ by maximum likelihood, yielding a bootstrap estimate $(\hat{\beta}^*, \hat{\sigma}_{\varepsilon}^*)$.

The estimation of $\hat{\phi}_i$ would be biased towards 1 in small samples. Simar and Wilson (2004) proposed a second bootstrap procedure, "the second algorithm," which included a parametric bootstrap in the first stage problem, so that bias-corrected scores are as follows:

Step 1 Compute $\hat{\phi}_i$ for all *n* decision units by solving problem (1);

Step 2 Estimate output surplus by maximum likelihood, considering it was a truncated regression. Let $\hat{\beta}$ and $\hat{\sigma}_{\varepsilon}$ be the maximum likelihood estimates of $\beta$ and $\sigma_{\varepsilon}$;

Step 3 Obtain the $L_1$ bootstrap estimate for each $\phi_i$ (e.g. $L_1$=100) in the following way: for *i* = 1,…, *n* draw $\varepsilon_i$ from a normal distribution with variance $\hat{\sigma}_{\varepsilon}^2$ and right truncation at $1 - z_i\hat{\beta}$ and compute $\phi_i^* = z_i\hat{\beta} + \varepsilon_i$. Let $y_i^* = \frac{\hat{\phi}_i}{\phi_i^*} y_i$ be a modified output measure. Compute $\hat{\phi}_i^*$ by solving problem (1), where **Y** is replaced by $Y^* = [y_i^* ... y_n^*]$.

Step 4 Compute the bias-corrected output inefficiency estimators as $\hat{\hat{\phi}}_i = 2 \cdot \hat{\phi}_i - \overline{\hat{\phi}}_i^*$, where $\overline{\hat{\phi}}_i^*$ is the bootstrap average of $\hat{\phi}_i^*$.

Once these first stage bias-corrected measures are produced, the second algorithm continues by replacing $\hat{\phi}_i$ with $\hat{\hat{\phi}}_i$ in the first algorithm, from step 2 onwards.

***Bayesian Stochastic Frontier Analysis***

With reference to the production set $\Gamma$ (see appendix), estimation and inference were undertaken by formulating a prior probability density function (*pdf*) $f(\theta)$, where $\theta$ was unobserved parameters and combining the prior with the likelihood function $f(y \mid \theta)$, where *y* was the set of observable data, and the posterior *pdf* $f(\theta \mid y)$ was formed by using Bayes' theorem. The approach used the Markov Chain Monte Carlo (MCMC) method of Gibbs sampling. The Gibbs sampler allows the marginal posterior distribution of parameter of interest to be approximated by generating a sample drawn from the marginal posterior distribution. $u_i's$ were the part of the set of random quantities from which the joint posterior was derived.

The Gibbs sampler involves the following steps:

Step 0, choose a starting value, $\theta^{(0)}$ for s = 1,..., *S*;
Step 1, take the random draw, $\theta_{(1)}^{(s)}$ from $p\left(\theta_{(1)} \mid y, \theta_{(2)}^{(s-1)}, \theta_{(3)}^{(s-1)}, \ldots, \theta_{(B)}^{(s-1)}\right)$;
Step 2, take the random draw, $\theta_{(2)}^{(s)}$ from $p\left(\theta_{(2)} \mid y, \theta_{(1)}^{(s-1)}, \theta_{(3)}^{(s-1)}, \ldots, \theta_{(B)}^{(s-1)}\right)$;
Step 3, take the random draw, $\theta_{(3)}^{(s)}$ from $p\left(\theta_{(3)} \mid y, \theta_{(1)}^{(s)}, \theta_{(2)}^{(s)}, \theta_{(4)}^{(s-1)}, \ldots, \theta_{(B)}^{(s-1)}\right)$;

.
.
.

Step *B*: Take the random draw, $\theta_{(3)}^{(s)}$ from $p\left(\theta_{(B)} \mid y, \theta_{(1)}^{(s)}, \theta_{(2)}^{(s)}, \ldots, \theta_{(B-1)}^{(s)}\right)$.

These will yield a set of *S* draws, $\theta^{(s)}$ for *s* = 1,2,..., *S*. After dropping the burn-in replications, $S_0$ to eliminate $\theta^{(0)}$, average the set of $S_1$ to create estimates of posterior features of interest. As a weak law of large numbers, if $f(\cdot)$ is a function of interest and

$$\hat{f}s_1 = \frac{1}{S_1} \sum_{s=S_0+1}^{S} f\left(\theta^{(s)}\right) \qquad (5)$$

if $S_1$ goes to infinity then $\hat{g}s_1$ converges to $E(f(\theta) \mid y)$.

## Data and Variables

The substance of the curriculum consists of a body of knowledge, skills or learning processes, values or virtues, morality and ethical behaviour. This substance is assembled into 8 groups: 1) *Thai language*, 2) *mathematics*, 3) *sciences*, 4) *social studies* (*religion and culture*), 5) *foreign language*, 6) *health and physical education*, 7) *arts*, and 8) *occupations and technology*. Nevertheless, only groups 1-5 are included in the national test. The annual time-frame of general secondary education grades 7-9 is 5-6 hours daily. Approximate time shall be allocated to 8 subject groups on more or less an equal basic. Nevertheless, the importance of the Thai language, mathematics, and sciences must be recognized and more time shall be allocated. The learning time framework stipulated in the Basic Core Curriculum 2008 is shown in table 1 below:

**Table 1: Lower general secondary education: learning time framework**

| Learning area | Approximate number of hours per year | | |
|---|---|---|---|
| | Grade 7 | Grade 8 | Grade 9 |
| *Thai language* | 120 | 120 | 120 |
| *Mathematics* | 120 | 120 | 120 |
| *Sciences* | 120 | 120 | 120 |
| *Foreign language* | 80 | 80 | 80 |
| *Social studies* | 80 | 80 | 80 |
| Health and physical education | 40 | 40 | 40 |
| Arts | 20 | 20 | 20 |
| Occupations and technology | 20 | 20 | 20 |
| Total yearly hours (basic) | 600 | 600 | 600 |
| Learning development activities | 400 | 400 | 400 |
| Additional courses | 200 | 200 | 200 |
| Total learning time per year | 1,200 | 1,200 | 1,200 |

**Source**: Ministry of Education, 2001. One credit is equivalent to 40 hours per semester.

Since each subject has a different learning time framework, the test score is calculated based on the number of hours per year. The approximate weight of the Thai language, math, science, foreign language, and social studies was 0.230, 0.230, 0.230, 0.154, and 0.154, respectively.

Data were collected by employing the research instruments called the public expenditure tracking survey (PETS) and the quantitative service delivery survey (QSDS) from October, 2007 to December, 2008. There are 54 expand-opportunity schools in Amnatcharoen; however, only 27 schools were included in the study do to the completeness of the research instruments (representing 50% of the population).

The outputs include the weight test scores of five core subjects from Grade 9 students (Mattayom 3): Thai language, math, science, social studies, and foreign language (English) from the academic year 2006.

The inputs include the variables that can be controlled by the school administrator, which are: *actual* rule-based expenditures (student's capitation grant), which was the main financial resource of school operation; and discretionary funding per student (fundamentally-needed fund), which was the fund allocated to the school and had the main objective to support poor students that attend the school. And average yearly teacher salary per student was another input for the school operation.

The institutional variables were selected within the school-based management framework. The accountability of school principals is upward, to the ministry that holds them responsible for providing services to the clients who, in turn, have put the policymakers/politicians in power and thus have the *voice* to hold the policymakers/politicians accountable for their performance.

In most cases of the SBM, the *management mechanism* of the school changed under the reform process. The clients themselves become part of the management, along with the frontline providers. As a result, the "short route of accountability" becomes short as representatives of the clients, either parents or community members, obtain the authority to make certain decisions and have a voice in decisions that directly affect the students that attend the school. The SBM framework is presented in figure 1, where the school administrators, whether the head teacher alone or a committee of parents and teachers, act as the accountable entity.

**Table 2: Descriptive statistics of variables in the model**

| Abbreviation | Variables (at school level) | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| | <u>*Inputs* (X)/academic year</u> | | | | |
| *PERCAP* (*PG*) | Avg. capitation grants received/student | 1,288.68 | 244.67 | 777.22 | 1,918.52 |
| *FUNDNEED* (*FF*) | Avg. fundamental-needs received/student | 269.04 | 69.48 | 102.73 | 412.22 |
| *SALARY* (*SY*) | Avg. teacher salary/student | 18,035.58 | 5,092.43 | 4,178.36 | 27,014.18 |
| | <u>*Outputs* (Y)</u> | | | | |
| *THAI* | Avg. Thai languages test scores | 41.88 | 4.74 | 37.14 | 57.33 |
| *MATH* | Avg. mathematics test scores | 28.06 | 3.86 | 22.83 | 40.00 |
| *SCIENCE* | Avg. science test scores | 35.88 | 5.99 | 24.46 | 51.00 |
| *ENGLISH* | Avg. English languages test scores | 28.69 | 5.68 | 21.15 | 47.67 |
| *SOCIAL* | Avg. Social studies test scores | 37.99 | 4.86 | 28.12 | 48.00 |
| *WCOMPTEST* (*WTS*) | Weight test scores | 34.34 | 3.91 | 29.09 | 48.85 |
| | <u>*Institutional arrangements* (Z)</u> | | | | |
| *POLITICIAN* | Politicians' involvement (*Voice*) | 0.29 | 0.46 | 0.00 | 1.00 |
| *CLASS SIZE* | No. of students/classroom (*Management*) | 15.93 | 4.56 | 7.71 | 25.43 |
| *BOM* | BOM meetings (*Client power*) | 4.30 | 1.68 | 2.00 | 10.00 |
| *INSPECTION* | Number of Inspections (*Compact*) | 7.67 | 5.37 | 2.00 | 25.00 |

The school operation was assumed to behave like the translog production function, which was derived as follows:

$$\begin{aligned}\ln(WTS)_i &= \ln\beta_0 + \beta_{PG}\ln(PG)_i + \beta_{FF}\ln(FF)_i + \beta_{SA}\ln(SY)_i \\ &+\beta_{PG,PG}[\ln(PG)]_i^2 + \beta_{FF,FF}[\ln(FF)]_i^2 + \beta_{SY,SY}\left[\ln(SY)\right]_i^2 \\ &+\beta_{PG,FF}\ln(PG)_i\ln(FF)_i + \beta_{PG,SY}\ln(PG)_i\ln(SY)_i + \beta_{FF,SY}\ln(FF)_i\ln(SY)_i\end{aligned} \quad (6)$$

where output is the weight composite students' test score (W*TS*) and the three inputs are capitation grants (*PG*), fundamental-needed funds (*FF*), and average teacher salary (*SY*). $\beta_0$ is the intercept, $\beta_{PG}$, $\beta_{FF}$, $\beta_{SY}$, $\beta_{PG,PG}$, $\beta_{FF,FF}$, $\beta_{SY,SY}$, $\beta_{PG,FF}$, $\beta_{PG,SY}$, and $\beta_{FF,SY}$ are the parameters to be estimated.

The institutional variable is class size and represents *management*, since it captures the classroom environment in case this affected the student's achievement. The politician's involvement was added to test the strength of "long route of accountability;" if they were accountable to the clients this should significantly explain the schools' efficiency. The SBM framework concerns the creation of community involvement; in this case, the number of meetings of the board of management was proxied for *client power.* In order to capture the *compact*, the relationship of accountability between government and frontline providers was proxied by the number of inspections from higher authorities.

$$Z_i = \delta_0 + \delta_1 POLITICIAN + \delta_2 CLASSIZE + \delta_3 BOM + \delta_4 INSPECTION \quad (7)$$

The Office for National Education Standards and Quality Assessment (Public Organization), or ONESQA, provides information regarding school performance evaluation. The 2006-2010 evaluation round data were compared to the results from the economic model. For basic education, there are 14 measures that assess schools, briefly described as follows (ONESQA, 2006):

***Students' perspective***

1. Students institute the right morals, and values;
2. Students institute hygiene, and physical and mental manners;
3. Students institute aesthetics, arts, music, and sports manners;
4. Students institute the capabilities of analysis, synthesis, consideration, creativity, thoughtfulness, and vision;
5. Students institute the necessary knowledge and skills that comply with the curriculum;
6. Students institute self-learning skills, and are ready to acquire new knowledge and continuously improve themselves;
7. Students institute working skills, are ready to work, are team players, and have a good attitude toward profession;

***Teachers' perspective***

8. Teachers institute knowledge and ability according to their responsibilities, and there are adequate teachers in the school;

9. Teacher institute the ability to deliver services efficiently and with emphasis on a student-centered philosophy;

***School administrator***

10. School administrators institute the leadership skills, and management capability;

***Administration's perspective***

11. School institutes the organizations/structures and systemically administers for meeting educational objectives;

12. School institutes curriculum activities and ervices, with emphasis on student-centered philosophy;

13. School institutes appropriate curriculum for students and community, and teaching resources that facilitate learning objectives;

14. School institute relationships and corporations among communities for educational development.

**Table 3: ONESQA Assessment Criteria**

| Score | | Quality level |
|---|---|---|
| *General measurement* (*Index*) | | |
| Average measure's score lower or equal | 1.74 | Must Improved |
| Average measure's score | 1.75-2.74 | Fair |
| Average measure's score | 2.75-3.49 | Good |
| Average measure's score | 3.50-4.00 | Very good |
| *Only measurement* (*Index*) *no.5* | | |
| Average measure's score lower or equal | 1.74 | Must Improved |
| Average measure's score | 1.75-2.59 | Fair |
| Average measure's score | 2.60-3.49 | Good |
| Average measure's score | 3.50-4.00 | Very good |

Table 3 illustrates details of the assessment criteria. However, there are other criteria; that is, the school had to pass at least 11 out of 14 measures, and they will then be accepted as a "*pass.*" However, if the schools are not "*passing*" in either criteria, no matter what score they receive, the school will be considered as "*failing*" according to ONESQA standards. For example, school A "passes" 10 out of 14 measures. However if they got an average measure score of 2.98, which is above 2.75, the school will "fail" ONESQA's standards.

## Empirical Results

In Table 4, the results of the standard DEA showed that 8 schools (school no. 1, 4, 5, 9, 12, 13, 22, 23) were the most efficient ones. However, school no. 11 seemed to be relatively the least efficient. The DEA analysis indicates that on average, the schools' outputs could be increased by about 14% they were able to utilized resources efficiently.

The three-stage DEA approach provides some different perspectives. The efficient schools were reduced to 5, and school number 27, which was ranked at number 18 when computed by standard DEA, became the most efficient school when computed by the latter method. However, the schools' output could be increased by about 18% without an increase in inputs used.

**Table 4: Results of Efficiency Scores and ONESQA Assessment**

| School no. | DEA | | | ONESQA assessment | | | Three-stage DEA | | | Fully-corrected bootstrap DEA | | SFA | | BSFA | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Scores | Rank | Peers (School no.) | Scores | Pass | Rank | Score | Rank | Peers (School no.) | Score | Rank | Score | Rank | Score | Rank |
| 1 | 1.000 | 1 | 1 | 3.409 | 13/14 | 5 | 1.0000 | 1 | 1 | 1.3795 | 27 | 1.00217 | 12 | 1.14312 | 25 |
| 2 | 1.138 | 14 | 22 | 2.883 | 14/14 | 19 | 1.6051 | 27 | 1,27 | 1.0905 | 13 | 1.01181 | 25 | 1.02585 | 3 |
| 3 | 1.130 | 13 | 9,4,22 | 2.839 | *9/14* | 22 | 1.1377 | 12 | 1,27 | 1.3704 | 26 | 1.00688 | 22 | 1.05787 | 8 |
| 4 | 1.000 | 1 | 4 | 2.981 | *10/14* | 13 | 1.2092 | 16 | 1,9,27 | 1.0011 | 1 | 1.00999 | 23 | 1.06462 | 10 |
| 5 | 1.000 | 1 | 5 | 2.859 | *10/14* | 21 | 1.1561 | 13 | 1,27 | 1.0736 | 10 | 1.00108 | 3 | 1.12020 | 23 |
| 6 | 1.239 | 24 | 4,5,22 | 3.109 | 13/14 | 10 | 1.0787 | 10 | 1,27 | 1.0512 | 6 | 1.01134 | 24 | 1.06326 | 9 |
| 7 | 1.034 | 10 | 9,5,22 | 2.899 | *9/14* | 16 | 1.3831 | 24 | 1,9,27 | 1.0252 | 4 | 1.00385 | 20 | 1.13148 | 24 |
| 8 | 1.229 | 23 | 9,22 | 2.999 | 11/14 | 11 | 1.0684 | 9 | 9,1,27 | 1.0790 | 11 | 1.00095 | 2 | 1.14495 | 26 |
| 9 | 1.000 | 1 | 9 | 2.785 | *9/14* | 23 | 1.0000 | 1 | 9 | 1.0101 | 2 | 1.00162 | 10 | 1.04734 | 5 |
| 10 | 1.174 | 19 | 1,22 | 3.197 | 14/14 | 7 | 1.1299 | 11 | 1,27 | 1.1444 | 16 | 1.00131 | 7 | 1.06986 | 12 |
| 11 | 1.550 | 27 | 22 | 2.463 | *8/14* | 27 | 1.0277 | 7 | 1,27 | 1.2401 | 20 | 1.00077 | 1 | 1.02480 | 2 |
| 12 | 1.000 | 1 | 12 | 3.469 | 14/14 | 4 | 1.0000 | 1 | 12 | 1.0537 | 9 | 1.00115 | 4 | 1.05563 | 7 |
| 13 | 1.000 | 1 | 13 | 3.121 | 13/14 | 9 | 1.0000 | 1 | 13 | 1.1066 | 14 | 1.00368 | 19 | 1.08342 | 19 |
| 14 | 1.453 | 26 | 22 | 2.881 | 12/14 | 18 | 1.4225 | 26 | 1,27 | 1.3520 | 24 | 1.00262 | 14 | 1.08909 | 21 |
| 15 | 1.205 | 21 | 5,9,22,1 | 3.667 | 14/14 | 3 | 1.0194 | 6 | 9,1,27 | 1.2220 | 19 | 1.00156 | 9 | 1.07968 | 17 |
| 16 | 1.120 | 12 | 22,1 | 2.994 | *10/14* | 12 | 1.3193 | 22 | 13,27,1 | 1.0895 | 12 | 1.00192 | 11 | 1.03093 | 4 |
| 17 | 1.029 | 9 | 9,1,4,22 | 2.870 | 11/14 | 20 | 1.1655 | 15 | 27,9,1 | 1.0520 | 8 | 1.00252 | 13 | 1.07112 | 14 |
| 18 | 1.160 | 16 | 9,22,5 | 3.920 | 14/14 | 2 | 1.2642 | 19 | 27,1 | 1.3414 | 22 | 1.01199 | 26 | 1.09051 | 22 |
| 19 | 1.149 | 15 | 9,4,22 | 3.194 | 14/14 | 8 | 1.2887 | 21 | 9,1,27 | 1.3602 | 25 | 1.00352 | 18 | 1.08213 | 18 |
| 20 | 1.196 | 20 | 9,22 | 2.913 | 11/14 | 15 | 1.2642 | 19 | 1,27 | 1.2900 | 21 | 1.00478 | 21 | 1.08483 | 20 |
| 21 | 1.088 | 11 | 9,5,22 | 2.969 | 12/14 | 14 | 1.0537 | 8 | 1,27 | 1.0276 | 5 | 1.00141 | 8 | 1.18455 | 27 |
| 22 | 1.000 | 1 | 22 | 4.040 | 12/14 | 1 | 1.2151 | 17 | 1,27 | 1.0519 | 7 | 1.00126 | 6 | 1.06906 | 11 |
| 23 | 1.000 | 1 | 23 | 2.726 | *8/14* | 25 | 1.1601 | 14 | 9,12,13,27 | 1.3501 | 23 | 1.00268 | 15 | 1.04965 | 6 |
| 24 | 1.167 | 17 | 22,5 | 2.706 | *10/14* | 26 | 1.4124 | 25 | 1,27 | 1.1361 | 15 | 1.00329 | 16 | 1.07377 | 15 |
| 25 | 1.221 | 22 | 9,22 | 3.384 | 13/14 | 6 | 1.3477 | 23 | 1,27 | 1.0139 | 3 | 1.00332 | 17 | 1.02396 | 1 |
| 26 | 1.276 | 25 | 1,22 | 2.741 | *9/14* | 24 | 1.2500 | 18 | 13,27,1 | 1.2072 | 18 | 1.01352 | 27 | 1.07009 | 13 |
| 27 | 1.171 | 18 | 5,22,4 | 2.886 | *8/14* | 17 | 1.0000 | 1 | 27 | 1.1838 | 17 | 1.00119 | 5 | 1.07434 | 16 |
| Average | 1.138 | | | 3.071 | | | 1.184 | | | 1.1590 | | 1.00416 | | 1.07800 | |

The fully-corrected bootstrap DEA, SFA, and BSFA methods were also used to compute technical efficiency. The results seemed to be highly different for the bootstrap method. The most efficient school was school number 4, which was ranked the same as with the standard DEA. The most efficient schools when computed by the SFA and BSFA were school number 11 and 25.

The methods compared to the scores computed by the ONESQA assessment were highly different. Only school number 16 and 22 provided the same rank as the ONESQA. The three-stage DEA method presented a different picture since only school number 6 matched the ONESQA assessment results. Regarding the bootstrap method, school number 8 and 27 had the same ranking compared to the ONESQA method; and regarding the parametric method, the SFA provided the same rank for school number 3, 10 and 12. On the other hand, no school provided a ranking that matched the ONESQA results.

**Table 5: SFA School Inefficiency Effect (n=27)**

| | Coefficient | t-ratio | | Coefficient | t-ratio |
|---|---|---|---|---|---|
| Intercept | -15.91 | 15.92 | | | |
| Ln(*PG*) | 1.01 | 1.03 | *POLITICIAN* ($\delta_1$) | 0.008 | 0.01 |
| Ln(*FF*) | 12.34 | 12.44*** | *CLASSSIZE* ($\delta_2$) | -0.008 | 0.04 |
| Ln(*SY*) | -3.86 | 3.95*** | *PARTICIPATION* ($\delta_3$) | 0.01 | 0.04 |
| Ln(*PG*)$^2$ | -0.30 | 0.50 | *INSPECTION*,($\delta_4$) | 0.0053 | 0.07 |
| Ln(*FF*)$^2$ | -0.26 | 0.35 | $\sigma^2$ | 0.007 | 1.16 |
| Ln(*SY*)$^2$ | 0.19 | 0.53 | $\gamma$ | 0.05 | 0.05 |
| Ln(*PG*) Ln(*FF*) | -0.42 | 0.49 | | | |
| Ln(*PG*) Ln(*SY*) | 0.56 | 0.70 | | | |
| Ln(*FF*)Ln(*SY*) | -0.66 | 0.86 | | | |
| Log likelihood ratio | | | | 29.68 | |

**Note:** ***Significant at 1% level

The SFA method was the ranking method that most matched the ranking results of the ONESQA assessment method. Then SFA was employed using the data from 27 schools. Table 5 depicts the results of the SFA schools' inefficiency effects.

There were 2 parameters that were significant for the schools' efficiency: fundamentally-needed funds and teacher salary. Fundamentally-needed funds included the per-capitation grants allocated to the school, aiming to support students in need (the criterion was household income less than or equal to 40,000 baht per year). This fund enabled these groups to obtain the necessary and sufficient resources for learning. It is no surprise that this parameter significantly explains school efficiency. However, none of the parameters negatively explained school efficiency. In public service, salary is proportionate to years of services. This means that if the school has a lot of highly-experienced teachers, the student achievement in this group of schools will be lower. However, no other inefficiency parameters significantly explained the schools' efficiency. The low efficiency of these schools seemed to lack a sufficient degree of accountability strength.

**Table 6: ONESQA School Recommendations**

| Ranking | Recommendation |
|---|---|
| 27 (Fail school) | Improve curriculum and increase community participation |
| | Promote nourishing school lunches |
| | Professional teacher development |
| 26 | Improve instruction in math and English subject groups |
| | Increase board of school management (BOM) participation |
| | Increase community participation |
| 25 | Professional teacher development |
| | Increase teaching aids and educational technology in classroom |
| | Increase community participation |
| | Increase board of school management (BOM) participation |
| 24 | Raise community and board of school management (BOM) participation |
| | Promote nourishing school lunches |
| | Improve curriculum and teacher development |
| 23 (Fail school) | Increase teaching aids and educational technology in classroom |
| | Increase utilization of school resources |
| | Increase community participation |
| | Professional teacher development |
| 22 (Fail school) | Professional teacher development |
| | Increase school infrastructure |
| | Increase board of school management (BOM) participation |
| | Improve curriculum and increase community participation |

Table 6 shows the recommendations for low-ranking schools according to the ONESQA assessment. It is obvious that the schools lacked professional teachers and that they needed to be urgently developed. The SBM framework seemed to prove itself as the critical institutional management tool that had to be improved. All of these are needed in order to improve community and BOM participation. There were no concerns with the inspection and policymakers'/politicians' involvement from the ONESQA assessor's point of view. According to the SFA and qualitative data above, it is noted that short route accountability played the key role in increasing school efficiency. However, the accountability relationships were not strong enough.

## Conclusions

In this paper, the technical efficiency of small-size schools in Amnatcharoen was analyzed by assessing the outputs (student performance) against the inputs directly used in the education system (capitation grant, fundamentally-needed funds), and institutional variables stemmed from the SBM framework (politicians' involvement, class size, board's school meetings, inspections).

None of the methods provided the same school ranking. For this group of samples, the SFA was selected as the model to analyze the function of educational production. However, there were some matches between the SFA and ONESQA assessment results, and this can support the employment of the SBM framework.

It is noted that there was room in this study to combine the quantitative and qualitative methods. The quantitative method could be used to compute the education production process (output), and the qualitative method can be used to assess the identity and outcomes of the school. Identity may include the vision, mission, and core values of the school, and the outcomes may include the high productivity worker and citizenship of the students. The goal of the educational system was the ability to build citizens'capability so that they can have a competitive advantage in this era of globalization. This combination can be used to assess the entire education system: input-output analysis and outcome quality measurement.

## References


Ahmad, Munir and Bravo-Ureta, Boris. (1996). Technical Efficiency Measures for Dairy Farms Using Panel Data: A Comparison of Alternative Model Specifications. *Journal of Productivity Analysis*. 7(4), 399-415.

Balcombe, Kevin, Fraser, Iain and Kim, Jae. (2006). Estimating Technical Efficiency of Australian Dairy Farms Using Alternative Frontier Methodologies. *Applied Economics.* 38(19), 2221-2236.

Banker, Rajiv, Charnes, Abraham and Cooper, William. (1984). Some Models For Estimating and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*. 30(9), 1078-1092.

Barrera-Osorio, Felipe, Fasih, Tazeen, Patrinos, Harry and Santibanez, Lucrecia. (2009). *Decentralized Decision-Making in Schools: The Theory and Evidence on School-Based Management*. Washington, D.C.: World Bank.

Barro, Robert and Lee, Jong—Wha. (2001).Schooling Quality in a Cross-section of Countries. *Economica.* 68, 465-488. Battese, George and Coelli, Timothy. (1992). Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India. *Journal of Productivity Analysis.* 3(1-2), 153-169.

Battese, George and Coelli, Timothy. (1995). A Model for Technical Inefficiency Effects in A Stochastic Frontier Production Function for Panel Data. *Empirical Economics*. 20(2), 325-332.

Bauer, Paul, Berger, Allen, Ferrier, Gary and Humphrey, David. (1998). Consistency Conditions for Regulatory Analysis of Financial Institutions: A Comparison of Frontier Efficiency Methods. *Journal of Economics and Business*. 50(2), 85-114.

Borger, Bruno and Kerstens, Kristiaan. (1996). Radial and Nonradial Measures of Technical Efficiency: An Empirical Illustration for Belgian Local Governments Using an FDH Reference Technology. *Journal of Productivity Analysis*. 7(1), 41-62.

Cummins, David and Zi, Hongmin. (1998). Comparison of Frontier Efficiency Methods: An Application to the U.S. Life Insurance Industry. *Journal of Productivity Analysis*. 10(2), 131-152.

Farrell, M. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A.* 120, 253-290.

Fried, Harold, Lovell, Knox, Schmidt, Shelton, and Yaisawarng, Suthathip. (2002). Accounting for Environmental Effects and Statistical Noise in Data Environment Analysis. *Journal of Productivity Analysis.* 17, 157-174.

Giannakas, Konstantinos, Tran, Kien C. and Tzouvelekas,Vangelis. (2003). On the Choice of Functional Form in Stochastic Frontier Modelling. *Empirical Economics.* 28(1), 75-100.

Gocht, Alexander, and Balcombe, Kelvin. (2006). Ranking Efficiency Units in DEA Using Bootstrapping an Applied Analysis for Slovenian Farm Data. *Agricultural Economics.* 35(2), 223-229.

Hanushek, Eric, and Kimko, Dennis. (2000). Schooling, Labor Force Quality, and Economic Growth. *American Economic Review.* 90, 1184-1208.

Hanushek, Eric., Luque, Javier. (2003). Efficiency and Equity in Schools Around the World. *Economics of Education Review.* 22, 481-502.

Kneip, Alois, Park, Byeong, and Simar, Leopold. (1998). A Note on the Convergence of Nonparametric DEA Estimators for Production Efficiency Scores. *Econometric Theory.* 14(6), 783-793.

Kim, Yangseon, and Schmidt, Peter. (2000). A Review and Empirical Comparison of Bayesian and Classical Approaches to Inference on Efficiency Levels in Stochastic Frontier Models with Panel Data. *Journal of Productivity Analysis.* 14(2), 91-118.

Koop, Gary, Osiewalski, Jacek, and Steel, Mark. (1997). Bayesian Efficiency Analysis Through Individual Effects: Hospital Cost Frontiers. *Journal of Econometrics.* 76(1-2), 77-105.

Krueger, Alan, and Lindahl, Mikael. (2001). Education and Growth: Why and for Whom? *Journal of Economic Literature.* 39, 1101-1136.

Ministry of Education. (2001). Office of the Basic Education Commission. *Basic Education Core Curriculum B.E. 2544 (A.D. 2001).*

Sharma, Khen, Leung, Pingsun, and Zaleski, Halina. (1997). Productive Efficiency of the Swine Industry in Hawaii: Stochastic Frontier vs. Data Envelopment Analysis. *Journal of Productivity Analysis.* 8(4), 447-459.

Sickles, Robin. (2005). Panel Estimators and the Identification of Firm-Specific Efficiency Levels in Parametric, Semiparametric and Nonparametric settings. *Journal of Econometrics.* 126(2), 305-334.

Simar, Leopold, and Wilson, Paul. (2000). A General Methodology for Bootstrapping in Non-Parametric Frontier Models. *Journal of Applied Statistics.* 27(6), 779-802.

Simar, Leopold, and Wilson, Paul. (2004). *Estimation and Inference in Two-Stage, Semi-Parametric Models of Production Process*. April, mimeo.

The Office for National Education Standards and Quality Assessment. (2006). *Standard, Index, and Criteria for Quality Assessment, Basic Education Round 2 (2006-2010).*

Van den Broeck, Julien, and Koop, Gary. (1994). Stochastic Frontier Models: A Bayesian Perspective. *Journal of Econometrics.* 61(2), 273-303.

World Bank, (2003). *World Development Report 2004: Making Services Work for Poor People.* Oxford University Press for The World Bank, Washington D.C.

Yangseon, Kim, and Schmidt, Peter. (2000). A Review and Empirical Comparison of Bayesian and Classical Approaches to Inference on Efficiency Levels in Stochastic Frontier Models with Panel Data. *Journal of Productivity Analysis.* 14(2), 91-118.

## Appendix

***Data envelopment analysis framework***

The DEA, which originated from Farrell's (1957) seminal work and which was popularized by Charnes, Cooper, and Rhodes (1978), assumes the existence of a production frontier. Inputs ( $x \in R_{+}^{n}$ ) and outputs ( $y \in R_{+}^{n}$ ) are represented by the production set $\Gamma$ of attainable input-output combinations:

$$\Gamma = \{(x,y) \in R_{+}^{n} : x \text{ can produce } y\} \tag{A.1}$$

The technology was assumed to satisfy the set of axioms discussed in, for example, Shephard (1953, 1970). In short, the axioms are: (*i*) inactivity is allowed; (*ii*) 'free lunch' is not allowed; (*iii*) strong disposability of inputs and outputs; and the (*iv*) production frontier is convex. The representation of the technology is the input set defined as:

$$L(y) = \{x : (x, y) \in \Gamma\}. \tag{A.2}$$

The production frontier is completely characterized by Farrell's (1957) scalar-values output-based technical efficiency measure, defined as:

$$\phi(x, y)^{-1} = \theta(x, y) = \max\{\theta : \theta y \in L(y)\}. \tag{A.3}$$

The measurement is the reciprocal of the output distance functions defined by Shephard (1953, 1970). $\theta(x, y) \le 1$ if and only if $y \in L(y)$. The value of the efficiency measure is given by $\|y\| / \|y^{I}\|$, where $y^{I} \in IsoqL(y) = \{y : y \in L(y), \iota y \notin L(y), \iota < 1\}$ is the production frontier output. Since the direction of the output vector is held fixed, $\phi$ is said to be a radial measure; it gives the maximum feasible, proportionate increase of outputs for a firm operating at $(x, y) \in \Gamma$. Clearly $\phi(x, y) \ge 1$ if and only if $y \in \theta y$; if $\phi(x, y) = 1$, then $(x, y) \in IsoqL(y)$ and point $(x, y)$ is said to be output-efficient. It will be useful later to denote the efficient level of output, corresponding to input level *x* and the output vector direction determined by *y*, as:

$$y'(x) = \frac{y}{\phi(x,y)} \tag{A.4}$$

Note that $y'(x)$ is the intersection of $IsoqL(y)$ and ray $(x,\theta y),\theta \in [0,\infty]$. Typically, $\Gamma$, $y^I$ and $IsoqL(y)$ are unknown; hence, for the firm producing at $(x,y)$, $\phi(x,y)^{-1}$ is also unknown. The DEA technique provides a consistent estimator of $\phi(x,y)^{-1}$ from a random sample $\Im = \{(x_i,y_i) \mid i = 1,...,n\}$.

This production frontier in the DEA approach was constructed using linear programming methods, the term "envelopment" stemming from the fact that the production frontier envelops the set of observations. Following Banker et al. (1984), the DGP of variable return to scale (VRS) output-orientated measure of technical efficiency (*TE*) was the solution for the linear programming problem, which can be expressed as:

$$\begin{aligned} &\max_{\phi,\lambda} \phi, \\ \text{st} \quad & -\phi y_i + Y\lambda \geq 0, \\ & x_i - X\lambda \geq 0, \\ & I1'\lambda = 1 \\ & \lambda \geq 0, \end{aligned} \tag{A.5}$$

where $\phi$ is a scalar, $y_i$ and $x_i$ are the column vector of outputs, and inputs for the *i*-th school, respectively. $\lambda$ is a $N \times 1$ vector of constants. Variable *Y* is an $M \times N$ output matrix, while *X* is a $K \times N$ matrix in which $1 \leq \phi < \infty$, and $\phi - 1$ is the proportional increase in outputs that could be achieved by the *i*-th firm, with input quantities held constant; $I1$ is a $I \times 1$ vector of ones.

***Stochastic frontier analysis***

The basic idea behind the SFA is that the error term is composed of two parts; (*i*) the systematic component that captures the effect of the measurement error, and others were statistical noise and random events, and the one-sided component that captures the effects of inefficiency. Several extensions of the SFA have been proposed (Battese and Coelli, 1992). The general formulation of the model is:

$$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + ... + \beta_k x_{ik} + \varepsilon_i \quad \text{(A.6)}$$

where $y_i$ is the output and $x_j$ are input. $\varepsilon_i = v_i - u_i$ ,where $v_i \sim N(0, \sigma_v^2)$ and $u_i \sim N(0, \sigma_u^2)$, $u_i \geq 0$, and the $u_i$ and $v_i$ are assumed to be independent. $\varepsilon_i$ is the difference between the standard white-noise disturbance; $(v_i)$ allows to captures the effect of measurement error, other statistic noise, and random events, and the one-sided component $(u_i)$ captures the effect of inefficiency. The method is used for decomposing the residual, which are defined as the functional form of the distribution of the one-sided inefficiency component, and for deriving the condition of the distribution of $[u_i \mid v_i - u_i]$ such as the half normal distribution.

Following Battese and Coelli (1992), for the half normal distribution assumption, consider the generalized production frontier for education as:

$$y_i = \exp(x_i \beta + v_i - u_i) \quad \text{(A.7)}$$

where $y_i$ denotes the output of the *i*-th school, $x_i$ represents a $(1 \times k)$ vector of inputs and other institutional variables for the *i*-th school, $\beta$ is the $(k \times 1)$ vector of unknown parameters to be estimated, $v_i s$ are assumed to be *iid*~ $N(0, \sigma_v^2)$ random variables, and $u_i s$ are non-negative unobserved random variables associated with the technical efficiency of production, where $u_i$ is defined as $u_i = \{\exp[-n(t-T)]\} u_T$; *i* = 1,2,...I. *t* = 1,2,...T, where $u_i \sim N^+(\mu, \sigma^2)$ and $\eta$ is the parameter to be estimated. As $t \to T$, $u_i \to u_T$. Then the inefficiency in periods prior to *T* depends on the parameter $\eta$. If $\eta$ is positive then $\exp\{-\eta(t-T)\} = \exp\{\eta(T-t)\}$ is always greater than 1, and increases with the distance of the period *t* and the last period *T*. If $\eta$ is positive, then it implies that technical inefficiencies decrease over time and if negative, then it implies that technical inefficiencies increase over time. $TE_i$ is the output-orientated technical efficiency of producer *i*, and can be expressed as:

$$TE_i = \frac{y_i}{f(\mathrm{x}_i; \beta) \cdot \exp(v_i)}. \quad \text{(A.8)}$$

In order to explore the sources of inefficiency, Battese and Coelli (1995) extended the model, where $u_i$ is defined as:

$$u_i = z_i\delta + c_i, \tag{A.9}$$

where $z_i$ is the $(1\times M)$ vector of institutional variables associated with technical inefficiency, $\delta$ is the $(M\times 1)$ vector of unknown parameters, and $c_i$ is the non-negative unobserved random variable obtained by truncation of the $c_i \sim N^+\left(0,\sigma_c^2\right)$ such that $c_i \geq -z_i\delta$. This is the specification of $u_i$ being a non-negative truncation of the $N(z_i\delta,\sigma^2)$. The general interest in this model concerns the testing of the null hypothesis as to whether inefficiency effects were present in the model, which is expressed as $H_0:\gamma=0,$ where $\gamma=\sigma_u^2/(\sigma_v^2+\sigma_u^2)$.

***Adjusted DEA: A three-stage procedure***

Following Fried et al. (2002), the three-stage procedure was used for incorporating environmental effects and statistical noise into producer-performance evaluation based on the DEA. The technique involves a three-stage procedure. In the first stage, the DEA is applied to inputs and outputs only to obtain an initial measurement of the producer-performance using the above linear programming problem. In the second stage, the SFA or Tobit model is used to regress the first stage performance measure against the set of environmental variables, a three-way decomposition of the variation in performance into a part attributable to environmental effects, a part attributable to managerial inefficiency, and a part attributable to statistical noise. The Tobit model is specified as follows:

$$\begin{aligned} s_{ni} &= r_{ni}\beta^n + v_{ni} \qquad \text{if } s_{ni} > 0 \\ &= 0, \end{aligned} \tag{A.10}$$

where $s_{ni}$ is the output surplus of school *i*-th obtained from stage 1, $r_{ni}$ = [1 $z'_{ni}$] is an $(1\times(L+1))$ vector of institutional factors plus one, $\beta$ is an $((L+1)\times 1)$ transpose vector of parameters, $v_{ni}$ is statistical noise, identically and independently distributed with zero mean and constant variance $\sigma^2$.

The methodology that decomposes the compose error terms from the conditional estimators for managerial inefficiency is given by $\hat{E}[u_{ni} \mid v_{ni} + u_{ni}]$; it derives the estimators for statistical noise residually by means of:

$$\hat{E}[v_{ni} \mid v_{ni} + u_{ni}] = s_{ni} - z_i \hat{\beta}^n - \hat{E}[u_{ni} \mid v_{ni} + u_{ni}], n = 1, ..., I, \tag{A.11}$$

which provide conditional (on $v_{ni} + u_{ni}$) estimators for the $v_{ni}$ in equation (A.11). Since the $\hat{E}[u_{ni} \mid v_{ni} + u_{ni}]$ depends on $(\hat{\beta}^n, \hat{\sigma}^2_{vn}, \hat{\sigma}^2_{un}, \hat{\mu}^n)$, so do the $\hat{E}[v_{ni} \mid v_{ni} + u_{ni}]$. The elements of $\hat{\beta}^n$ provide estimates of the contributions of each observable environmental variable to the surplus of the *n*th output, while the parameters $(\hat{\sigma}^2_{vn}, \hat{\sigma}^2_{un}, \hat{\mu}^n)$ characterize the separate contribution of managerial inefficiency and statistical noise to output of the *n*th output. However, in this case, the variation in managerial inefficiency plays no role in producing the output, and that variation in predicted output is due exclusively to statistical noise. Observed output was adjusted for the influences of the noise, which can be calculated as:

$$\hat{E}[v_{ni}] = s_{ni} - z_i \hat{\beta}^n, n = 1, ..., I. \tag{A.12}$$

Note that $\hat{E}[v_{ni}]$ was not bounded from below at zero; it can take on negative, zero, or positive values. If $\hat{E}[v_{ni}]$ is zero, then the school performs as well as the average firm with the same set of institutional factors. If $\hat{E}[v_{ni}]$ is not equal to zero, then its performance differs from the average firm with the same set of institutional factors. The parameter estimates for output surplus are based upon Tobit's specification. The schools' adjusted outputs were constructed from the results of the Stage 2 Tobit regression by mean of :

$$y^A_{ni} = y_{ni} + [\max_i \left\{ z_i \hat{\beta}^n \right\} - z_i \hat{\beta}^n] + [\max \left\{ \hat{v}_{ni} \right\} - \hat{v}_{ni}], n = 1, ..., N, i = 1, ..., I, \tag{A.13}$$

where $y^A_{ni}$ and $y_{ni}$ were adjusted and observed output quantities, respectively. The first adjustment on the right side of equation puts all schools into a common operating environment, which was the least favourable environment. The second adjustment puts all schools into common state of nature, which was the unluckiest situation encountered. Thus firms with relatively least favourable operating environments and/or relatively the unluckiest have their outputs adjusted upward by a relatively small amount, while producers with relatively favourable operating environments and/or relatively good luck have their outputs adjusted upward by a relatively large amount. These adjustments vary both across producers and across outputs. In the third stage, outputs were adjusted to account for the impact of the environmental effects and statistical noise uncovered in the second stage, and the DEA was used to re-evaluate producer performance. The analysis emphasis is on outputs, rather than on radial efficiency scores, as appropriate measures of producer performance.

***Bootstrap and non-discretionary inputs***

A perturbation on an observation located on the DEA estimate frontier will shift the production frontier. As a result, some firms will find themselves near or far from the frontier, and their efficiency scores will change accordingly. The bias came from the correlation derived from the non-discretionary inputs and outputs, which are the ingredients in estimating the scores. Thus, standard approaches to inference are usually not valid in small samples. In order to overcome this, Simar and Wilson (2004) proposed an alternative estimation and inference procedures based on bootstrap methodology.

It is straightforward to prove that $\hat{\delta}$ is a consistent estimator of $\delta$. Kneip et al. (1998) showed that:

$$\hat{\delta} = \delta + \zeta_p \left( n^{-\frac{2}{p+q+1}} \right). \tag{A.14}$$

The asymptotic distribution of the output distance function is the special case of one input and one output $(p = q = 1)$. In the more general multivariate setting, where $p + q > 2,$ the radial nature of the distance functions and the complexity of the estimated frontier made the derivation complex. So far, the bootstrap appears to offer the only way to approximate the asymptotic distribution of the distance function estimators in multivariate settings.

For arbitrary point $(x, y) \in R_+^{p+q}$ and $y = \left[ y_1 ... y_q \right],$ since radial distance is used, to the polar coordinates of *y* defined by its modulus $\omega = \omega(y) \in R^+$, where $\omega(y) = y'y$ and its angle $\eta = \eta(y) \in \left[ 0, \frac{\pi}{2} \right]^{q-1}$, where, for $j = 1, ..., q-1,$ and *p* is prime number, $\eta_j = \arctan\left( \frac{y_j + 1}{y_1} \right)$ if $y_1 > 0$ or $\eta_j = \frac{\pi}{2}$ if $y_1 = 0.$ The Farrell efficiency measure can be expressed as:

$$\delta(x,y\mid\Gamma)=\frac{\omega\left(\delta\left(x,y\mid\Gamma\right)y\right)}{\omega\left(y\right)}.$$

Since $\Gamma$ is fixed, one can characterize $y$ by $(\eta,\delta)$, where $\eta=\left[\eta_1...\eta_{p-1}\right]$ and $\delta=\delta(x,y\mid\Gamma)$.

The joint density $f\left(x,y,z\right)$ can now be described by a series of conditional densities and in terms of cylindrical coordinates:

$$f(x_i,\eta_i,\delta_i,z_i)=f\left(x_i,\eta_i\mid\delta_i,z_i\right)f\left(\delta_i\mid z_i\right)f(z_i). \tag{A.15}$$

The order of the conditioning on the right-hand side of (A.15) reflects the sequential nature of the DGP. Firm *i* was faced with institutional variable $z_i$ drawn from $f(z)$. Given this $z_i$, an efficiency level $\delta_i$ was drawn from $f\left(\delta_i\mid z_i\right)$, and then $x_i$ and $\eta_i$ were drawn from $f\left(x,\eta\mid\delta,z\right)$, resulting in realization $\left(x_i,y_i,z_i\right)$ from the joint density $f\left(x,y,z\right)$ after transforming the polar coordinates $\left(\eta_i,\delta_i\right)$ to Cartesian coordinates $y_i$.

In the real world, the firms face certain institutional variables $z$, and this constrains their choices of inputs $x$ and outputs $y$ and also confronts them with a set of observations $\Theta_n=\left\{\left(x_i,y_i,z_i\right)\right\}_{i=1}^{n}$. The two-stage studies that have appeared in the literature typically specified $\psi\left(z_i\beta+\varepsilon_i\right)=z_i\beta$ and can be written as:

$$\delta_i=z_i\beta+\varepsilon_i\geq 1, \tag{A.16}$$

where $\delta_i=\delta\left(x_i,y_i\mid\hat{\Gamma}\right)$. This study then: (*i*) used the observed pairs $\left(x_i,y_i\right)$ in $\Theta_n=\left\{\left(x_i,y_i.z_i\right)\right\}_{i=1}^{n}$ to estimate $\delta_i$ for all $i=1,...,n$, yielding a set of estimates $\left\{\hat{\delta}_i\right\}_{i=1}^{n}$; (*ii*) replaced the unobserved $\delta_i$ on the left-hand side of (16), which estimated $\hat{\delta}_i$ obtained from step (*i*); and then (*iii*) estimated:

$$\hat{\delta}_i = z_i\beta + \xi_i \geq 1, \tag{A.17}$$

using the Tobit regression, or in a few cases, ordinary least squares. However, Simar and Wilson (2004) revealed an additional problem. Note that:

$$\hat{\delta}_i = E\left(\hat{\delta}_i\right) + u_i, \tag{A.18}$$

where $E\left(u_i\right) = 0$. In addition, the bias of the estimator $\hat{\delta}_i$ is defined by:

$$\text{BIAS}\left(\hat{\delta}_i\right) \equiv E\left(\hat{\delta}_i\right) - \delta_i. \tag{A.19}$$

From (A.18) and (A.19), rearranging terms yields:

$$\delta_i = \hat{\delta}_i - \text{BIAS}\left(\hat{\delta}_i\right) - u_i. \tag{A.20}$$

Substituting for $\delta_i$ in (A.16) gives:

$$\hat{\delta}_i - \text{BIAS}\left(\hat{\delta}_i\right) - u_i = z_i\beta + \varepsilon_i \geq 1. \tag{A.21}$$

Since $\hat{\delta}_i$ was a consistent estimator, the $u_i$ becomes negligible asymptotically, as does $\text{BIAS}\left(\hat{\delta}_i\right)$.

These facts provide justification for writing (A.21), the equation that is typically estimated in two-stage applications. The bias of $\hat{\delta}_i$ is always strictly negative in finite samples. $u_i$ was unknown and cannot be estimated, but the bias term can be estimated by the bootstrap method. The bootstrap bias estimate equals the true bias plus a residual:

$$\widehat{\text{BIAS}}\left(\hat{\delta}_i\right) = \text{BIAS}\left(\hat{\delta}_i\right) + v_i. \tag{A.22}$$

The variance of the residual $v_i$ diminishes as $n \to \infty$, and hence $v_i$ is typically of smaller magnitude than $\text{BIAS}\left(\hat{\delta}_i\right)$ for reasonable sample sizes *n*. The bootstrap estimator of the bias can in turn be used to construct a bias-corrected estimator of $\delta$:

$$\hat{\hat{\delta}}_i = \hat{\delta}_i - \widehat{\text{BIAS}}\left(\hat{\delta}_i\right). \tag{A.23}$$

Rearranging the term (A.20), (A.21) and (A.22) yields:

$$\hat{\hat{\delta}}_i + v_i - u_i = z_i\beta + \varepsilon_i \geq 1. \tag{A.24}$$

As noted, both terms $v_i$ and $u_i$ become negligible asymptotically; hence maximum likelihood is:

$$\hat{\hat{\delta}}_i \approx z_i\beta + \varepsilon_i \geq 1. \tag{A.25}$$

will yield consistent estimates.

The efficiency scores that solve (A.1) $\hat{\phi}_i$, were then considered as an estimate for $\phi_i$, and this was the first stage in the procedure. The second stage was designed to assess the influence of non-discretionary inputs on efficiency.

***Bayesian stochastic frontier analysis***

Following Koop et al. (1997), as in a classical exponential case, it was assumed that $v$ is normally distributed with mean zero and constant variance $(h_v)$, and *u* was the Gamma distribution with a shape parameter *j* and an unknown scale parameter $\lambda$. When *j*=1 this yields an exponential probability distribution, i.e. $u_i \sim F_G(u_i, 1, \lambda^{-1}) \propto \lambda^{-1}\exp(-u_i\lambda^{-1})$. Van den Broeck and Koop (1994) found the exponential probability distribution to be the most robust model with respect to assumption of the prior median efficiency. Balcombe et al. (2006) suggest that, assuming the prior of $\beta$ as:

$$p(\beta) \propto I(\beta \in \Lambda), \tag{A.26}$$

where $I(\cdot)$ is an indicator function. In this context $\Lambda$ was the region of the parameter space where the constraints implied by economy (i.e., monotonicity and curvature) were satisfied. A few papers have estimated flexible functional forms; in this paper the translog production function was used, and imposed monotonicity and quasi-concavity via the indicator function in equation (A.26).

The prior for $\lambda$ has the following form:

$$p(\lambda^{-1}) = f_G(1, -\ln(r^*)), \tag{A.27}$$

where $r^*$ was the prior median of the efficiency distribution. The result for the informative prior $(r^*)$ of 0.875 is presented; however, Koop et al. (1997) employed 0.850, and Kim and Schmidt (2000) employed 0.80. Finally, the choice of the prior for *h*, was:

$$p(h_v) = h_v^{(n_0-2/2)} \exp(-h_v a_0) \tag{A.28}$$

with $n_0 \geq 0$ and $a_0 > 0$. $n_0 = 0$ and $a_0$ were set equal to zero or very small numbers. In order to conduct Bayesian inference on the model, using Gibbs sampling sequential draws from the following conditional posteriors:

$$p(\lambda^{-1} \mid y, \beta, h_v, u) = f_G(\lambda^{-1} \mid N\bar{u}\lambda^{-1} - \ln(r^*)) \tag{A.29}$$

$$p(h_v \mid y, \beta, \lambda^{-1}, u) = f_G\left(\frac{N}{2} + \frac{n_0}{2}, \left(\frac{v'v}{2} - a_0\right)\right) \tag{A.30}$$

$$p\left(\beta \mid y, h_v, u, \lambda^{-1}\right) \propto f_N\left(b, h_v(\sum x_i x_i')^{-1}\right) \times I\left(\beta \in \Lambda\right) \tag{A.31}$$

$$p(u_i \mid y, \beta, \lambda^{-1}, h_v) \propto f_N(y_i - x_i'\beta - \frac{\lambda^{-1}}{h}, h_v^{-1}) \times I(u_i > 0),$$

$$p(u \mid y, \beta, \lambda^{-1}, h) = \prod_{i=1}^{n} p(u_i \mid y_i, \beta, \lambda^{-1}h). \tag{A.32}$$

The result of interest will focus on $\beta$ marginal density functions and technical efficiency measurement derived by taking MCMC draws from the joint posterior density. Gibbs sampling can be briefly described here, for the *S* replication; however, the first $S_0$ of these were discarded as so-called "burn-in replications," and the remaining $S_1$ was retained for the estimate of $E(f(\theta) \mid y)$, where $S_0 + S_1 = S$. This was the case for blocks (*B*) but can be extended to more blocks.