Efficiency of Higher Technical Educational Institutions in India

Praveen Kulshreshtha
Associate Professor of Economics
Department of Humanities and Social Sciences
Indian Institute of Technology (IIT) Kanpur

Tapan Kumar Nayak
Associate Professor of Economics
Institute of Management Studies (IMS)

ABSTRACT
There has been a significant growth in the number of higher technical educational institutions (HTEIs) in India during the past six decades. To evaluate the competency of the Indian HTEIs, in terms of number of students produced, quality of students’ training etc., it is necessary to assess the economic efficiency of these institutions. This paper examines the technical efficiency (TE) of eight important HTEIs in India, namely, seven Indian Institutes of Technology (IITs) and Indian Institute of Science (IISc), by applying Stochastic Frontier Analysis (SFA) and Data Envelopment Analysis (DEA) to balanced panel data during 2001-05. The study uses the input-oriented and output-oriented stochastic distance function models, as well as constant returns to scale DEA approach to measure the TE of the above institutions. The paper demonstrates that TE varies across the above HTEIs and highlights the need for strengthening the know-how (concerning higher technical education) of the Indian HTEIs, so that they can exploit the full potential of the existing educational inputs.

Keywords: Data Envelopment Analysis (DEA); Higher Technical Educational Institutions (HTEIs); Stochastic Frontier Analysis (SFA); Technical Efficiency (TE).

EFFICIENCY OF HIGHER TECHNICAL EDUCATIONAL INSTITUTIONS IN INDIA

I. Introduction
Higher education is of vital importance to the education system in any society (Paul (2003)). It is becoming increasingly more significant as knowledge-based industries now occupy the center-stage of development in all contemporary societies. It is also becoming apparent that the growth of the global economy has increased opportunities for countries with good levels of education (Carnoy (1999), Ilon (1994), Stewart (1996) and Tilak (2001)).

Higher education, whether general or technical, must be related to a country’s goals and endeavours. Since technical education determines the socio-economic development of a country, there is a strong need for higher technical educational institutions (HTEIs) to provide high quality technical education to produce technically skilled man-power, especially in developing countries such as India (Palit (1998)). Moreover, globalization of the Indian economy has made consumers and producers more quality conscious. In such a scenario, higher technical education in India needs to be reoriented to face the growing challenge of creating high quality students in teaching and research (Singh (2002)).
The growth of higher technical education in India has contributed immensely to the country’s economic and industrial development during the past six decades. The most significant contribution has been towards the development of a strong technical base, with 7 world-class Indian Institutes of Technology (IITs); Indian Institute of Science, Bangalore (IISc); and 20 National Institutes of Technology (NITs, formerly known as Regional Engineering Colleges or RECs). In 2005, the news group ’Times’ ranked IITs as the third best technical institutions in the world. *IISc is a premier institution of research and advanced instruction in frontier areas of science and technology, which enjoys a high international reputation.*

If we consider the development of the technical base in India, we must understand the efficiency with which the system of higher technical education in India has worked so far. Significant facilities have been created in Indian HTEIs over the past five decades but one still finds that there is a large amount of wastage in the efforts put in the above institutions. For example, the figures of intake and out-turn at the degree and diploma level in Indian HTEIs reveal that the student drop-out rate at the degree and diploma level are about 20 to 22 percent and 52 percent respectively, while the overall drop-out rate is about 30 percent (Source: Official website, Department of Secondary and Higher Education, Government of India).

An intensive study of the problem of student drop-out is very important on account of academic and planning considerations. Any endeavor to improve existing courses, introduce new courses, and diversify programs would be boosted a great deal by the useful information generated from such a study. From the planning perspective, any attempt to reduce wastage of efforts would contribute to the efficiency of the system of higher technical education and thus provide for greater out-turn of additional technical manpower without further inputs (ibid). Thus, the efficiency with which Indian HTEIs manage their resources is an important issue (Abbott and Doucouliagos (2004)).

In general, HTEIs may have a number of objectives concerning the quantity and quality of teaching and research (Stevens (2001)). Indian HTEIs such as the IITs have provided world-class technical education at the under-graduate level, but the above institutions need to improve the quality of post-graduate education, where they have a poor track record, and which means doing more research (Forbes (2003)). As Forbes has pointed out, firms in Indian industries do not carry out much research and development (R&D) activity, unlike the firms in the developed countries, and mostly transfer technology from the developed countries instead. Therefore, unlike the developed countries, strong industry-institution partnerships between firms and HTEIs, which strengthen the management and R&D of firms, are absent in India.

However, there is a need for firms in the Indian industry to increase their R&D activity to become more competitive in the global markets, and develop a strong partnership with Indian HTEIs to promote industrial R&D. Since research in HTEIs is usually carried out at the post-graduate level, a strong industry-institution partnership would also enhance the quality of post-graduate education in India. Also, the engineering professionals graduating from the above institutions should be up-to-date in their know-how, have the willingness and capability

---

1 The 7 IITs are: IIT Kanpur, IIT Kharagpur, IIT Bombay, IIT Delhi, IIT Madras, IIT Guwahati and IIT Roorkee. Moreover, 8 new IITs are being developed, many of which have begun teaching and research activities since 2008-09. The new IITs include IIT Bhubaneswar, IIT Gandhinagar, IIT Hyderabad, IIT Indore, IIT Mandi, IIT Patna and IIT Rajasthan.

to learn new things, and a deep sense of work ethics and must be conversant with the skills needed to perform well in a job (George (2000)).

This study attempts to measure the technical efficiency of prominent Indian HTEIs, namely 7 IITs and IISc. To do so, the study identifies the factors that influence the technical efficiency of IITs and IISc. Given the multiple-input, multiple-output nature of higher educational technology, we estimate the technical efficiency pertaining to IITs and IISc by employing the SFA technique, in particular, the Input Oriented and Output Oriented Distance Function Approaches towards Stochastic Frontier Analysis (SFA), as well as the Data Envelopment Analysis (DEA) methodology for ‘balanced’ panel data from 2001-02 to 2004-05.

The paper is organized as follows. In Section 2, we review the literature pertaining to higher education and higher technical education, as well as technical efficiency, in the context of developed and developing economies. Section 3 delineates the conceptual framework and methodology employed to measure the technical efficiency of Indian HTEIs. Section 4 describes the estimation and interpretation of results concerning the technical efficiency of IITs and IISc. Lastly, Section 5 highlights the policy implications of our study and concludes the paper.

II. Review of the Literature

In this Section, we review various important studies pertaining to: (i) higher education and higher technical education in India, and (ii) measurement of efficiency in higher education (mentioned below). There have been several attempts in the past to study the trends and patterns of higher education in India. However, most of such studies have focused on the financial aspects of higher education in India. The efficiency of higher educational institutions (HEIs) has been analyzed extensively in developed countries such as Australia, Germany, Japan, South Africa, U.K. and U.S.A.

2.1 Studies Concerning Higher Education and Higher Technical Education in India

Altbach (1993) has analysed the changing scenario of higher education in India, especially in the context of economic reforms in the country that began in 1991. He has mentioned that it may be possible to implement relatively small but meaningful reforms in higher education in India, even if it is difficult to achieve systematic change.

Tilak (1993) has critically reviewed privatization and other proposals developed by Government of India and have argued that the higher education in India is not yet ready for privatization. However, Tilak has realized that public budget cannot adequately fund higher education in India, especially when the mass education sector needs special attention. Therefore, he has emphasized on several alternatives such as increasing student fees, revitalizing student loans, imposing graduate tax and privatization to finance higher education in India. Tilak has concluded that the actual mix of public subsidies and non-governmental finances for higher education reflects the economic policies of the government, and the tradition and social policies of the society.

3 There are several studies pertaining to the measurement of efficiency in Indian industries such as agriculture (Battese and Coelli (1995)); banking (Ketkar et. al. (2004)); electricity (Shanmugam and Kulshreshtha (2002, 2005), Shanmugam and Mary (2001)); manufacturing (Agarwal and Goldar (1992), Bhavani (1991), Goldar (1985), Kumar and Mishra (2002), Little et. al. (1987), Page (1984), Patibandla (1998) and Ramaswamy (1990)); and pharmaceuticals (Khan et. al. (2005)).
Tilak (1999) has also evaluated the financial aspects of higher technical education in India. He has stated that higher education is one of the most important components of human capital. It is a specialized form of human capital and the economic returns to higher technical education in India are estimated to be very high. The study indicates that the success of economic reforms in the country depends upon the quality and quantity of technical manpower. Thus, the need for more financial resources, to develop higher technical education in India, is obvious.

Geetha Rani (2004) has studied the impact of economic reforms on financing of higher education in India. She has attempted to examine the financing of higher education in India during 1980s and 1990s, by looking at various sources of funding for higher education, and has highlighted the changes in terms of the hike in student fees, introduction of student loans operated by commercial banks, rapidly increasing role of private sector and self-financing courses within public HEIs.

The study has emphasized that under the influence of globalization and competitions, important economic rationale for public funding of higher education in India are being neglected. It is essential that funding sources are diversified, but cost-sharing with students has social and political limits, and excessive commercialization of higher education should be discouraged.

### 2.2 Studies Concerning Technical Efficiency in Higher Education

Farrell (1957) suggested that the production function of a firm (which relates inputs to output) can be estimated either by a parametric function, or a non-parametric piece-wise linear technology. The parametric approach led to the stochastic frontier analysis (SFA), which can also be readily applied to the estimation of cost function of a firm (which relates output to cost). The non-parametric approach gave rise to the Data Envelopment Analysis (DEA) methodology, which is a linear programming technique used in the estimation of a production function. Some important studies pertaining to estimation of the technical efficiency in higher education in different countries are discussed below (in chronological order):

Avkiran (2001) has used Data Envelopment Analysis (DEA) approach to examine the technical efficiency of Australian universities. Three performance models have been developed, namely, overall performance, performance on delivery of educational services and performance on fee-payment enrollments. The study states that DEA is particularly appropriate when the measurer of performance is interested in investigating the efficiency of an institution in converting multiple inputs into multiple outputs. The findings show that Australian universities are technically efficient (i.e. exhibit low values of technical inefficiency). Furthermore, the performance models adequately discriminate between technically efficient and inefficient universities.

Stevens (2001) has examined the cost efficiency of HEIs in English and Welsh universities using the method of stochastic frontier analysis (SFA). This paper is unique because it investigates the impact of staff and student characteristics on cost efficiency. The study uses stochastic frontier analysis to estimate the cost function, while allowing for multiple outputs to be produced by HEIs, namely (i) number of under-graduates in arts and science programmes, (ii) number of post-graduate students, and (iii) total research funding attracted. The results suggest that a significant proportion of the error term in the cost function is explained by the inefficiency effect.

Similarly, Abbott and Doucouliagos (2003) have measured the technical efficiency of Australian universities in 1995 using DEA. They have stated that efficiency analysis of
Australian universities is valuable not only for Australian officials and policy makers, who are concerned with expanding higher education while containing costs, but is also of interest to officials in other countries. They found that regardless of the output-input mix, Australian universities have recorded high levels of technical efficiency during the study period.

Moreover, Abbott and Doucouliagos (2004) have estimated the technical efficiency of Australian and New Zealand public universities using stochastic frontier analysis (SFA) and examined the effect of competition for overseas students on the technical efficiency of the above institutions. In the Australian case, they have used two separate samples, first, involving data for 36 Australian government-owned universities that operated during 1995-2002, and second, involving data for 34 business or commerce faculties associated with Australian government-owned universities during 1997-2000. In the case of New Zealand, they have used data for 7 government-owned universities during 1997-2003.

According to economic theory, competition among firms in a market generally induces firms to improve their methods of production, and achieve higher levels of efficiency and productivity. Abbott and Doucouliagos have found that during the study period, competition for overseas students had led to higher technical efficiency in Australian universities. However, competition for overseas students has had no impact on the technical efficiency in New Zealand universities.

Ray and Jeon (2008) have employed the Pareto-Koopmans global efficiency measure to examine the efficiency levels of the MBA programs in Business Week’s list of top-ranking MBA programs in U.S.A. They have computed both input-oriented and output-oriented radial and non-radial efficiency measures for comparison. They found that among three tier groups, the schools from a higher tier group on average are more efficient than those from lower tiers, although variations in efficiency levels do occur within the same tier, which exist over different measures of efficiency. Overall, their findings can aid the administrators of the MBA programs to evaluate their performance in meeting the demands of the graduates, on the one hand and of the employers, on the other.

Moreover, several other studies have explored the efficiency in professional and higher education. For example, Johnes and Johnes (1993) have used DEA to measure research efficiency of a number of Economics departments from British universities based on publications and personnel data collected by the Royal Economic Society. Breu and Raab (1994) have analyzed the data from the Top-25 National universities and Liberal Arts colleges in U.S.A. to measure their efficiency levels using DEA. They have found that several of the best-rated universities such as Cal Tech (rated 5th) and Chicago (rated 10th) operate at less than 90% efficiency.

Other applications of DEA to measure efficiency in higher education include Burton and Phimister (1995), who have applied DEA to evaluate the efficiency of a set of “core journals” identified by Diamond (1989). In yet another study, Tracy and Waldfogel (1997) have pointed out that a valid ranking of MBA programs in U.S.A. should use objective criteria that are comparable across programs and should also be based on “outputs” rather than “inputs”.

Similarly, Haksever and Muragishi (1998) have used output-oriented CCR for the top 20 MBA programs as well as the second 20 MBA programs in U.S.A. to analyze early 1990s data from Business Week, and have found no average efficiency differences between the above two groups of MBA programs, unlike the Ray and Jeon (2008) study. McMillan and Datta (1998) have used DEA to measure efficiency of Canadian universities.

URL: http://dx.doi.org/10.14738/abr.31.794
In another interesting study, Colbert et al. (2000) have developed an alternative ranking of MBA programs in U.S.A. based on DEA, by using the survey response scores reported in Business Week, so as to compare the MBA programs in U.S.A. with three foreign MBA programs. Calhoun (2003) has applied DEA to compare efficiencies of public and privately funded HEIs in the U.S.A, while Jourmady and Ris (2003) have compared the efficiencies of HEIs across a number of countries in Europe. Also, Johnes (2006) has used student level data for several U.K. universities, and decomposed the DEA efficiency scores to isolate the performance of the university from that of a student.

From the above landmark studies, we can identify important variables, which can be used towards the estimation of technical efficiency in higher education. Most of these variables are taken from the studies pertaining to technical efficiency of higher education in developed countries (such as Australia, Germany, Japan, U.K., and U.S.A.) based on Stochastic Frontier Analysis (SFA) and Data Envelopment Analysis (DEA) methodology. These variables can be categorized as input and output variables, as follows:

**Inputs:**
Staff (teaching and non-teaching); Students (UG, PG and Post-doc.); Equipment and library stock; Supply of teaching materials (OHP projector, Transparencies, LCD projector, Stationary etc.); Technical (Lab) equipments (PC, printers, scanners, web cameras, software,…); Total expenditure on education (all educational activities); Number of Academic Staff; Number of Non-Academic Staff; Budget allocation (grant) from the central and state governments for a particular year.

**Outputs:**
Under-graduate student load (number of courses offered/lay-out) in Arts subjects; Under-graduate student load (number of courses offered/lay-out) in Science subjects; Post-graduate student load (number of courses offered/lay-out); Value of research grant and contracts received; Under-graduate enrollments; Post-graduate enrollments; Qualifications completed (degrees, diplomas and certificates); Research output (books, articles in approved (refereed) journals, conference proceedings, patents/licenses and research income); Total placements reported in a year; Number of companies coming for placement in a year.

In the preceding paragraphs, we have listed some important input and output variables, which have been selected from the landmark studies pertaining to technical efficiency in higher education. However, these variables are not fixed and can be changed to suit any given methodological specification. From these studies, it is also found that technical efficiency in higher education has been estimated mostly in developed countries. In particular, no efficiency analysis has been carried out in the case of higher technical education in India.

Some important studies, such as Tilak (1993, 1999) and Geetha Rani (2004) have emphasized the funding and equity aspects of higher education in India. If equity aspect is so emphasized, especially in the light of the educational reform process in India, then the efficiency of HEIs needs to be examined to learn more about the competitiveness of these institutions. Given the above research gap, this study attempts to estimate the technical efficiency of Indian HTEIs by applying the stochastic frontier analysis, in particular, Input-oriented and Output-oriented SFA, and DEA methodology to data concerning higher technical education in India.

**I. Methodology, Data and Variables**
We now delineate an efficiency framework to measure the technical efficiency of higher technical educational institutions in India. Note that efficiency measurement concerns not only
economic units such as firms, but also any form of “productive” units, which transform any kind of inputs to any kind of outputs under the current stage of knowledge. For this reason, the productive units are commonly referred to as Decision Making Units or DMUs in the efficiency literature. In this study, higher technical educational institutions constitute the DMUs under consideration.

3.1 Conceptual Framework Pertaining to Technical Efficiency of Firms in an Industry
To assess the performance of a particular industry, we must measure the economic efficiency of firms in that sector. According to Farrell (1957), technical efficiency can be defined as “the ability and willingness of firms to produce the maximum possible quantity of output with a specified endowment of inputs, given the technology and environmental conditions that surround them” (Kalirajan and Shand (2000)). Precisely, a firm is said to be technically efficient when it realizes its technical potential for the given set of inputs and technology.

Suppose \( x = (x_1, \ldots, x_n) \) denotes the vector of \( n \) inputs, which a firm uses to produce a single output \( y \). The production function \( y = f(x) \) denotes the efficient transformation of inputs to output, i.e. the maximum amount of \( y \) that is obtainable from the given combination of inputs \( x \). Let \( (y', x') \) be the observed production plan of the firm. The observed production plan is said to be technically efficient, if \( y' = f(x') \) and technically inefficient if \( y' < f(x') \). It is important to note that \( y' > f(x') \) is not possible, given the current technology. The measure of technical efficiency \( TE \) can be defined by the ratio of \( y' \) to \( f(x') \), i.e. \( TE = y'/f(x') \), which must lie between 0 and 1.

3.2 Empirical Methodology for Estimation of Technical Efficiency in an Industry
The techniques most commonly employed for efficiency and productivity measurement of firms in an industry are: Stochastic Frontier Analysis (SFA), Data Envelopment Analysis (DEA), Total Factor Productivity (TFP) Approach and Malmquist TFP Index Approach. Since the last two approaches are used only for the measurement of productivity and changes in productivity of firms, we focus on the use of SFA technique and DEA methodology to measure the technical efficiency of firms in an industry.

Both of the above techniques can be employed to analyze the technical efficiency of firms in an industry, where firms use multiple inputs and multiple outputs. While DEA is a non-parametric technique which takes multiple inputs and multiple outputs into consideration, a distance function approach towards SFA can be employed to measure the technical efficiency of firms which use a multiple input, multiple output technology.

In 1957, Farrell founded the conceptual framework pertaining to technical efficiency and proposed two alternative methodologies for measuring the technical efficiency of firms in an industry. He pointed out that the production function of a firm is not known in reality and proposed that the production function can be estimated either by (a) a parametric function, such as Cobb-Douglas, trans-log and Constant Elasticity of Substitution (CES) forms or (b) a non-parametric piece-wise linear technology (Coelli et al. (1998)). The first approach was developed further by Aigner and Chu (1968), which led to the stochastic frontier production function approach or SFA technique\(^4\). The second approach was followed by Charnes, Cooper and Rhodes (1978), which gave rise to the DEA methodology. Unlike the SFA method, DEA does not require the specification of a particular functional form for the production function, and

---

\(^4\) Other notable contributions towards this methodology are by Afriat (1972), Richmond (1974), Schmidt (1976) and Timmer (1971).
hence, it provides efficiency estimates that are not conditional on the specific functional form assumed for the production function.

We discuss the above two approaches towards measurement of technical efficiency in the subsections below. More specifically, we discuss the basic SFA technique, as well as Battese and Coelli’s SFA model in the next subsection. The distance function approach towards SFA to measure TE of firms that employ multiple inputs and multiple outputs is discussed in subsection 3.2.2. The DEA approach towards measurement of technical efficiency is delineated in subsection 3.2.3.

3.2.1 Stochastic Frontier Production Analysis (SFA)
The frontier production function can be defined as the maximum possible output that a firm can produce with a given level of inputs and technology. The notion of a deterministic frontier production function was first proposed by Farrell in 1957 and later followed by Aigner and Chu in 1968. The deterministic frontier production function can be expressed as follows:

\[ Q_{it} = f(x_{it}; \beta) \exp(-u_{it}), \quad 0 \leq u_{it} \leq \infty, \quad i = 1,...,n, \quad t = 1,...,T \]  

(3.1)

Where,

\( Q_{it} \): Actual output of \( i^{th} \) firm in period \( t \)

\( x_{it} \): Vector of inputs

\( \beta \): Vector of parameters

\( f(.) \): Frontier production function or potential output of a firm

\( u_{it} \): One-sided (non-negative) residual term

When the operation of a firm is efficient, then the actual output produced by a firm is equal to its potential output and vice versa. Therefore, the technical efficiency (TE) of the firm can be measured by taking the ratio of actual output \((Q_{it})\) and potential output \(f(.)\) of a firm in period \( t \), which equals \( \exp(-u_{it}) \). The residual term \( u_{it} \) becomes zero when the firm produces potential output (i.e. full TE) and is greater than zero in case of less than full TE. Therefore, the TE of a firm and the residual error term \( u_{it} \) are inversely related to each other, and the term \( u_{it} \) is also considered as the efficiency effect or TE effect of firm \( i \) in period \( t \) (Shanmugam and Kulshreshtha (2005)).

Equation (3.1) above implies that the production function is linear in the logs of the variables and can be estimated as:

\[ \ln Q_{it} = \ln f(x_{it}; \beta) - u_{it} \]  

(3.1')

Where,

\[ u_{it} = \ln f(x_{it}; \beta) - \ln Q_{it} = \ln[f(x_{it}; \beta) / Q_{it}] = \ln(1/TE) \]

Technical efficiency can be explained with the help of Figure 1 below, which is drawn from Shanmugam and Kulshreshtha (2003):
Figure 1: Frontier Production Function and Technical Efficiency
In Figure 1 above, the curve OADF represents the deterministic frontier, which traces the maximum level of output that a firm can achieve in a given period at different levels of input, given the existing technology. For instance, if the firm uses $I_1$ units of the input, the maximum output that can be produced during a period with the existing technology is $Q_1$ units, which is represented by point A on the frontier.

Similarly, if the firm uses $I_2$ units of the input, the potential or maximum feasible output is $Q_2$ units, given the existing technology, as illustrated by point D on the frontier. If a firm operates at any point on the frontier, such as point A or D, it is considered to be technically efficient, since it is able to employ the input efficiently to produce the maximum possible output, using the existing technology.

However, a firm may operate at a point below the frontier, such as point B or E, so that its output is less than the potential or maximum feasible output, for the given level of input and existing technology. In this case, the firm is considered to be technically inefficient. This may happen either because firm engineers have incomplete knowledge regarding the best techniques of applying the input with the existing technology or because of various organizational constraints that prevent the plant from reaching the frontier (Kulshreshtha and Nayak (2005), quoting Kalirajan and Shand (1994)).

For instance, at point B in Figure 1 above, which lies inside or below the frontier, the firm uses $I_1$ units of input but is able to produce only $Q_3$ units of output, which is less than $Q_1$ units or the frontier output of the firm at $I_1$ units of input. In this case, the output loss due to technical inefficiency is $(Q_1 - Q_3)$ units. Similarly, at point E, which also lies below the frontier, the firm uses $I_2$ units of input but produces only $Q_4$ units of output, which is less than the frontier output $Q_2$. Therefore, the output loss because of technical efficiency is $(Q_2 - Q_4)$ units.

A random noise $v_t$ can also be introduced in the production function of a firm (equation (3.1) above) to capture the effect of measurement errors/omitted variables, where $v_t$s are independently and identically distributed (i.i.d.) as $N(0, \sigma^2)$. The resulting production function is called the stochastic frontier production function and can be expressed as follows:

$$Q_t = f(.) \exp(v_t) \exp(-u_t)$$  \hspace{1cm} (3.2)

Thus, the technical efficiency (TE) of a firm can be measured by taking the ratio of actual output $(Q_t)$ and potential output $f(.) \exp(v_t)$ of a firm in period $t$, which equals $\exp(-u_t)$. Equation (3.2) above implies that the production function is linear in the logs of the variables and can be estimated as:

$$\ln Q_t = \ln f(x_t; \beta) + v_t - u_t$$  \hspace{1cm} (3.2')

Where,

$$u_t = \ln f(x_t; \beta) + v_t - \ln Q_t = \ln[f(x_t; \beta) \exp(v_t)/Q_t] = \ln(1/TE)$$

---

5 Although it is feasible for the firm to produce $Q_2$ units of output by employing $I_2$ units of input (point D in Figure 1), it is not possible for the firm to achieve $Q_2$ level of output by employing only $I_1$ units of input (point C in Figure 1), given the existing technology. In other words, point C is infeasible under existing technological conditions. However, with a technological advancement, the firm’s production frontier may shift upwards and hence, the firm may be able to achieve point C.
3.2.2 Distance Function Approach towards SFA

A distance function can be defined by specifying the production technology of a firm with the output set \( P(x) \), which represents the set of all output vectors \( y \in \mathbb{R}^M \) that can be produced using the input vector \( x \in \mathbb{R}^K \). That is,

\[
P(x) = \{ y \in \mathbb{R}^M : x \text{ can produce } y \}. \tag{3.3}
\]

It is assumed that the technology satisfies the axioms listed in Fare (1980). The distance function may be specified as either output oriented or input oriented distance function (see, for instance, Coelli and Perelman (1999), Coelli et al. (2003))\(^6\). These are discussed below:

**Output Oriented Distance Function**

The output distance function, introduced by Shepherd (1970), is defined on the output set, \( P(x) \), as,

\[
D_0(x, y) = \min \{ \theta : (y/\theta) \in P(x) \} \tag{3.4}
\]

As discussed in Lovell et al. (1994), \( D_0(x, y) \) is non-decreasing, positively linearly homogeneous and convex in \( y \), and decreasing in \( x \). The distance function \( D_0(x, y) \), will take a value which is less than or equal to one if the output vector, \( y \), is an element of the feasible production set, \( P(x) \). That is \( D_0(x, y) \leq 1 \) if \( y \in P(x) \).

Furthermore, the distance function will take the value of unity if \( y \) is located on the outer boundary of the production possibility set. That is \( D_0(x, y) = 1 \), if \( y \in Isoq P(x) = \{ y : y \in P(x), \omega y \not\in P(x), \omega > 1 \} \), using notation similar to that used by Lovell et al. (1994). Given sample data on \( N \) firms, there are number alternative approaches to calculate the frontier, such as:

1. Construction of a non-parametric piece-wise linear frontier using linear programming (DEA) [e.g., Fare, et al. (1989), Fare et al. (1994b)];
2. Construction of parametric deterministic frontier using linear programming [e.g., Forsund and Hjalmarsson (1987), Fare, et al. (1993)];
3. Estimation of a parametric deterministic frontier using corrected ordinary least square [e.g., Lovell, et al. (1994), Grosskopf, et al. (1997)]; and
4. Estimation of a parametric stochastic frontier using maximum likelihood estimation [e.g., Hetemaki (1996)].

However, it is necessary to decide an appropriate functional form for parametric empirical analysis. The functional form for the distance function would ideally be

1. Flexible;
2. Easy to calculate; and
3. Permits the imposition of homogeneity (Coelli and Perelman, 1996).

\(^6\)Coelli and Perelman (1999) have outlined the distance function solution to the multi-output problem by using data on European Railways and estimating output oriented, input oriented and constant returns to scale distance functions. Coelli et al. (2003) have measured efficiency relative to a stochastic input distance function by using survey data on private and cooperative Indian dairy processing plants, and have found that private dairy processing plants are not more cost-efficient than the cooperative dairy processing plants.
Compared to the Cobb-Douglas distance function, the trans-log distance function has been used by majority of the experts [e.g., Lovell et al. 1994; Grosskopf et al. 1997], since it satisfy the above three conditions. The trans-log distance function for the case of M outputs and K inputs is specified as follows:

\[
\text{LnDO}_i = \alpha_0 + \sum_{m=1}^{M} \alpha_m \ln y_{m_i} + \frac{1}{2} \sum_{m=1}^{M} \sum_{n=1}^{M} \alpha_{mn} \ln y_{m_i} \ln y_{n_i} + \sum_{n=1}^{K} \beta_{n} \ln x_{n_i} + \frac{1}{2} \sum_{n=1}^{K} \sum_{l=1}^{K} \beta_{n} \beta_{l} \ln x_{n_i} \ln x_{l_i} + \sum_{m=1}^{M} \delta_{mn} \ln x_{n_i} \ln y_{m_i}
\]

Where \( i \) denotes the i-th firm in the sample, \( i = 1, 2, N \). Note that to obtain the frontier surface (i.e. the transformation function), one would set \( \text{D}_0 = 1 \), which implies the left hand side of equation (3.5) is equal to zero.

**The restriction required for homogeneity of degree +1 in outputs are**

\[
\sum_{m=1}^{M} \alpha_m = 1 \tag{3.5a}
\]

And,

\[
\sum_{m=1}^{M} \alpha_{mn} = 0, \ m = 1, 2, M, \text{ and } \sum_{m=1}^{M} \delta_{mn} = 0, \ k = 1, 2, ..., K \tag{3.5b}
\]

And those required for symmetry are:

\[
\alpha_{nm} = \alpha_{mn}, \ m, n = 1, 2, ..., M, \text{ and } \beta_{kl} = \beta_{lk}, \ k, l = 1, 2, ..., K \tag{3.6}
\]

It is also noted in passing that the restrictions required for separability between inputs and outputs are

\[
\delta_{km} = 0, \ k = 1, 2, ..., K, \ m = 1, 2, ..., M \tag{3.7}
\]

These last restrictions will be used when we test for separability in the following section.

A suitable method of imposing the homogeneity constraint upon equation (3.5) is to follow Lovell, et al. 1994 and observe that homogeneity implies that:

\[
D_0(x, \omega y) = \omega D_0(x, y), \text{ for any } \omega > 0 \tag{3.8}
\]

Hence, if we arbitrarily choose one of the outputs, such as the M-th output, and set \( \omega = 1/y_M \), we obtain

\[
D_0(x, y/y_M) = D_0(x, y)/y_M \tag{3.9}
\]

Hence, for trans-log form we found:

\[
\text{Ln} \ (D_{0i}/y_M) = \\
\alpha_0 + \sum_{m=1}^{M-1} \alpha_m \ln y_{m_i} + \frac{1}{2} \sum_{m=1}^{M-1} \sum_{n=1}^{M-1} \alpha_{mn} \ln y_{m_i} \ln y_{n_i} + \sum_{n=1}^{K} \beta_{n} \ln x_{n_i} + \frac{1}{2} \sum_{n=1}^{K} \sum_{l=1}^{K} \beta_{n} \beta_{l} \ln x_{n_i} \ln x_{l_i} + \sum_{m=1}^{M-1} \delta_{mn} \ln x_{n_i} \ln y_{m_i}
\]

\( i = 1, 2, N, \)

Where \( y_{m_i} = (y_{m_i}/y_M) \) and \( D_{0i} \) denotes the trans-log output oriented distance function. It is observed that when \( y_{m_i} = y_{M_i} \) the ratio \( y_{m_i} \) is equal to one and hence the log of the ratio is zero. Thus, all terms involving the M-th output become zero. Hence, the summation involving \( y_{m_i} \) in the above expression is over \( (M-1) \) and not over \( M \).

Note that a single output production function is equivalent to an output distance function, when production involves only one output. Hence, if we set \( M = 1 \) and \( D_{0i} = 1 \), so as to trace the production surface, we find that:
\[-\ln (y_{M0}) = \alpha_0 + \sum_{i=1}^{K} \beta_i \ln x_{i0} + \frac{1}{2} \sum_{i=1}^{K} \sum_{k=1}^{K} \beta_{ik} \ln x_{ik} \ln x_{ik}, i = 1,2,\ldots,N. \tag{3.11}\]

This is the (negative of the) very familiar trans-log production function.

**Input Oriented Distance Function**

A trans-log input distance function is obtained by imposing homogeneity of degree 1 in inputs (instead of outputs) upon the transformation function. Thus, instead of obtaining equation (3.11), we will obtain

\[\ln \left( \frac{D_{ii}}{x_{ki}} \right) = \]

\[\alpha_0 + \sum_{i=1}^{M} \alpha_i \ln y_{i0} + \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \alpha_{ij} \ln y_{i0} \ln y_{j0} + \frac{1}{2} \sum_{i=1}^{M} \sum_{k=1}^{K} \beta_{ik} \ln x_{i0} \ln x_{k0} + \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \delta_{ij} \ln x_{i0} \ln y_{j0} \]

\[i = 1, 2, N, \tag{3.12}\]

Where \(x^{*}_{ki} = (x_{ki}/x_{K0})\) and \(D_{ii}\) denotes the *trans-log input oriented distance function*. Note that when \(x_{ki} = x_{K0}\) the ratio \(x^{*}_{ki}\) is equal to one and hence the log of the ratio is zero. Thus, all terms involving the \(K\)-th input become zero. Hence, the summation involving \(x^{*}_{ki}\) in the above expression is over (K-1) and not over \(K\).

Again, analogous to the relationship between production functions and output distance functions in the single output case, we observe that an input distance function will be equivalent to an input requirement function when a single input is used in the production process. Therefore, if we set \(K=1\) and also set \(D_{ii}=1\), so as to trace the production surface, equation (3.12) becomes

\[-\ln (x_{K0}) = \alpha_0 + \sum_{i=1}^{M} \alpha_i \ln y_{i0} + \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \alpha_{ij} \ln y_{i0} \ln y_{j0}, i = 1, 2, N. \tag{3.13}\]

This is the (negative of the) trans-log input requirement function.

With the selection of a suitable functional form for the output distance function, we must now adopt an appropriate method for obtaining estimates of the unknown parameters of the function. Hence, we must obtain estimates of the parameters of the function such that the function is good fit to the data. This may be underscored using simple algebra by rewriting equation (3.10) as:

\[\ln(D_{0i}/y_{M0}) = TL(x_{bi}/y_{M0} \alpha \beta), i=1,2,\ldots,N, \tag{3.14}\]

Or

\[\ln (D_{0i}) - \ln (y_{M0}) = TL (x_{bi}/y_{M0} \alpha \beta), i=1,2,\ldots,N, \tag{3.15}\]

And hence

\[-\ln (y_{M0}) = TL (x_{bi}/y_{M0} \alpha \beta) - \ln(D_{0i}), i=1,2,\ldots,N. \tag{3.16}\]

What is required here is the selection of parameter values for the trans-log function, which ensures the function fits the observed data “as closely as possible” while maintaining the requirement that \(0 < D_{0i} \leq 1\), which implies that \(-\infty < \ln(D_{0i}) \leq 0\). Since the output distances are

**URL:** http://dx.doi.org/10.14738/abr.31.794
bounded by zero and one, their logarithms must then be zero and negative. Thus, this is equivalent to minimizing the sum of the deviations of the observations below the frontier, where the deviations are defined as the logarithms of the inverse of the distances (Coelli and Perelman (1996)).

Here, it is important to mention that under constant returns to scale (CRS), the input distance function is equivalent to the inverse of the output distance function (i.e. \( D_0 = 1/D_i \)) (Fare et al. (1994a)). In case of the trans-log output distance function, CRS is imposed by imposing homogeneity of degree (-1) in inputs. The resulting function will obviously be exactly equal to the negative of the input distance function in which homogeneity of degree (-1) has been imposed in outputs.

**ML Estimation of Stochastic Distance Function**

The above mentioned two methods of fitting a parametric distance function explicitly assume that all deviations between observed production points and the production surface are due to the technical inefficiency. However, these deterministic frontier methods do not account for the possible influence of data noise upon the shape and positioning of the frontier and hence that the method may be sensitive to the influence of outliers.

Aigner et al. (1977) have proposed the stochastic frontier approach, which can be used to account for the influence of data noise upon an estimated frontier. This involves the specification of a frontier function with an error term with two components: a symmetric error to account for noise and an asymmetric error to account for inefficiency. To start with we apply a symmetric error term, \( \sigma_u^2 \), to the predicted value of the output distance for the \( i \)-th firm, \( D_{Oi} \). The stochastic output distance function thus obtained as

\[
-\ln(y_{Oi}) = TL(x_i, y_i/y_{Mi}, \alpha, \beta) + v_i + u_i \quad i=1,2,...,N. \tag{3.17}
\]

Given the distributional assumptions for \( v_i \) and \( u_i \), the parameters of this stochastic trans-log distance function can be estimated using the maximum likelihood method and assume that \( v_i \) are iid \( N(0, \sigma_v^2) \) and distributed independently of the \( u_i \) which are assumed to be iid \( N(0, \sigma_u^2) \) (Aigner, et al. (1977)).

However, the predicted value of the output distance for the \( i \)-th firm, \( D_0 = \exp(-u_i) \), is not directly observable because \( u_i \) only appears as part of the composed error term, \( e_i = (v_i + u_i) \).

The predictions may be obtained using a modification of the conditional expectations formulae described in Jondrow et al. (1982), and Battese and Coelli (1988). The output distance function value for the \( i \)-th firm may be obtained using conditional expectation as

\[
D_{Oi} = E[\exp(-u_i)|e_i] = \frac{1 - \Phi(\sigma_1 = \gamma \sigma_2, \sigma_2 = \gamma \sigma_2^2)}{1 - \Phi(\gamma \sigma_2 / \sigma_1)} \exp(\gamma \sigma_2^2 / 2) \tag{3.18}
\]

Where \( \sigma_1 = \sqrt{\gamma (1-\gamma) \sigma_v^2} \), \( \sigma_2 = \sigma_u^2 + \sigma_v^2 \), \( \gamma = \sigma_u^2 / 2 \), and \( \Phi(.) \) represents the distribution function of a standard normal random variable.

The maximum likelihood estimates of the unknown parameters and the distance function predictions can be obtained using the computer program called FRONTIER, version 4.1 (Coelli, 1996). Moreover, the stochastic frontier method can also be applied in a way similar to a trans-log input distance function, where the non-positive error term will be subtracted from the equation. Hence, the stochastic input distance function can be expressed as:
\[-\ln \left( x_{ki} \right) = TL \left( x_{i}/x_{ki}, y_{i}, \alpha, \beta \right) + v_{i} - u_{i} \quad (3.19)\]

And the input distances would be predicted as:

\[D_{ii} = E \left[ \exp \left( u_{i} \right) \right] e_{i} \quad (3.20)\]

Where, \( e_{i} = (VI - u_{i}) \)

### 3.2.3 Data Envelopment Analysis (DEA) Approach

Data Envelopment Analysis (DEA) is widely used to compare the relative efficiency of decision making units (DMUs) such as universities, hospitals, libraries, banks etc. Usually, such DMUs are characterized by a vector of multiple inputs and multiple outputs. Hence, absolute comparisons of such DMUs are difficult. In order to aggregate information about such inputs and output quantities, DEA makes use of fractional and corresponding linear programmes to measure the relative performance of DMUs (Klein (2004)).

Analogous to stochastic frontier analysis, DEA estimates the maximum potential output for a given set of inputs and has primarily been used in the estimation of technical efficiency. It is also used to measure capacity utilization of firms in an industry. A range of DEA models have been developed which measure technical efficiency and capacity utilization in different ways. These models largely fall under the categories of input-oriented or output-oriented models.

With input-oriented DEA, the linear programming model is configured so as to determine how the input use of a firm could contract, if the input is used efficiently in order to achieve the same output level. Input-oriented DEA is illustrated in Figure 2 below, which is borrowed from Dionysios et al. (2005). Using this figure, we can examine the technical efficiency of four DMUs (A, B, C and D), which use two inputs \((X_1, X_2)\) to produce a single output \(y\).

In Figure 2 above, the piecewise linear frontier constructed by the segments that pass through the points A and C approximates the efficient isoquant. The points on and in the area to the right of the efficient frontier (isoquant) define the input requirement set. DMUs A and C which determine the frontier are considered to be technically efficient, while DMUs B and D are considered relatively inefficient, as they lie in the interior of the efficient frontier. The projected point B’ is a linear combination of points A and C and defines the efficient operation of DMU B.

Geometrically, a measure of the technical efficiency of DMU B in Figure 2 above refers to the proportional reduction of both \(X_1\) and \(X_2\) along the radius defined by points O and B up to the point B’, which lies on the frontier. The deviation of the input quantities utilized by DMU B from the frontier (distance BB’) may be viewed as a measure of the degree of technical inefficiency of DMU B. Accordingly, the degree of technical efficiency for DMU B can be calculated by the ratio of distances \((OB'/OB)\). Measuring technical efficiency in this fashion is known as input-oriented technical efficiency, because it addresses the question: “By how much can the input quantities of a DMU be proportionally reduced without changing the output production?”

---

\(^{7}\) The input requirement set is defined as the locus of all input combinations, which produce at least a particular output level. In production economics, the input requirement set can be considered as a formal way to represent the production technology.

**URL:** http://dx.doi.org/10.14738/abr.31.794
In contrast, with output-oriented DEA, the linear programming model is configured to determine a firm’s potential output given its inputs, if the firm is operated efficiently, similar to the firms along the “best practice” frontier\(^8\). This is analogous to the stochastic frontier production function approach, which estimates the potential output for a given set of inputs and measures technical efficiency as the ratio of actual and potential output.

Hence, DEA is a multi-factor productivity analysis model, which measures the relative efficiency of a homogenous set of decision making units. The efficiency score in the presence of multiple inputs and outputs is defined as:

\[
\text{Efficiency} = \frac{\text{Weighted sum of outputs}}{\text{Weighted sum of inputs}}
\]

Assuming that there are \(n\) DMUs, each with \(m\) inputs and \(s\) outputs, the relative efficiency score of the \(P^{TH}\) DMU is obtained by solving the model given by Charnes et al. (1978) as follows:

\[
\max \sum_{j=1}^{s} \nu_j y_{jP} \quad \text{subject to} \quad \sum_{j=1}^{m} \lambda_j x_{jP} = \gamma \sum_{j=1}^{s} \nu_j y_{jP},
\]

\(\nu_j, \lambda_j, \gamma \geq 0\)

\(^8\) Output-oriented technical efficiency can be defined as the amount by which all the output quantities can be proportionally increased without changing the input quantities.
The fractional programme shown in equation (3.22) above can be converted to a linear programme as below:

\[
\begin{align*}
\text{Max} & \quad \sum_{k=1}^{s} v_k y_{ki} \\
\text{s.t.} & \quad \sum_{j=1}^{m} u_j x_{ji} = 1 \\
& \quad \sum_{k=1}^{s} v_k y_{ki} - \sum_{j=1}^{m} u_j x_{ji} \leq 0 \quad \forall i \\
& \quad V_k \geq 0, \quad u_j \geq 0 \quad \forall k, j
\end{align*}
\]

(3.23)

The above programme can be run \( n \) times to identify the relative efficiency scores of all DMUs. Each DMU selects input and output weights that maximize its efficiency score. In fact, a DMU is considered to be efficient if it obtains a score of 1 and less than 1 score implies that the DMU is inefficient.

Traditional DEA model does not allow for ranking DMUs, especially the efficient ones. In addition, it is possible that some of the inefficient DMUs are in fact better overall performers than certain efficient ones. This is because of the unrestricted weight flexibility problem in DEA. Cross-efficiencies in DEA is one method that could be utilized to identify good overall performers and effectively rank DMUs (Sexton et al., 1986).
3.3 Data, Sample and Variables
The dataset for the study includes observations pertaining to 7 IITs and IISc. More specifically, our sample consists of IIT Kanpur, IIT Kharagpur, IIT Bombay, IIT Delhi, IIT Madras, IIT Guwahati, IIT Roorkee and IISc\(^9\).

The data is taken from the Annual Reports of IITs and IISc for four years (i.e. academic year 2001-02 to 2004-05). Most of the annual reports were obtained from Department of Secondary and Higher Education, Technical Section I (for IITs and IISc) in Ministry of Human Resource Development (MHRD), Government of India, New Delhi, while some annual reports have been obtained from the institutions’ administrative offices directly. The dataset used in the study consists of 4-year ‘balanced’ panel data pertaining to IITs and IISc, for academic years 2001-02 to 2004-05\(^{10}\).

From the studies on technical efficiency in higher education discussed in Section 2 above, we found examples of several important variables that can act as determinants of technical efficiency in higher technical education. To calculate the technical efficiency of higher technical educational institutions in India, we use two output variables and three input variables.

The output variables are: (i) Total Research Publications (TRP), which measures the research output of an institution in a given academic year (henceforth, year); (ii) Total Enrollment (ENR), which measures the teaching output of an institution in a given year. The input variables pertain to Academic Staff (AS), Non-Academic Staff (NAS) and Library Stock (LS) of an institution in a given year.

The above output and input variables of a higher educational institution in a given academic year have been defined as follows:

TRP (Total Research Publications) = (Number of Journal Articles + Number of Conference Papers + Number of Chapters in Books + Number of Edited Books and Books)

ENR (Total Enrollment) = (Number of Under-graduate Students Enrolled + Number of Post-graduate Students Enrolled + Number of Research Scholars Enrolled)

AS (Academic Staff) = (Number of Classroom Teachers + Number of Research Associates + Number of Visiting Faculty)

NAS (Non-Academic Staff) = (Number of Non-Academic Staff, which includes Type A, B, C and D Staff)

LS (Library Stock held in the various libraries of the given institution) = (Number of Books + Number of Bound Volumes + Number of Journals + Number of Periodicals + Number of E-journals)

3.4 Estimation of Technical Efficiency in Higher Technical Education in India: Specification of Input-Oriented and Output-Oriented SFA Models
In the present study, we have attempted to measure the technical efficiency of higher technical educational institutions in India using input-oriented and output-oriented SFA approach, along

\(^9\) As pointed out earlier, 8 new IITs which have come up recently (i.e. during 2008-09) have not be included in our analysis due to lack of data.

\(^{10}\) ‘Balanced’ panels were generated by considering only those institutions for which data was available during all years of the study period.
with the DEA model. Following our earlier discussion in Section 3.2 above, we fit the following equations to estimate the TE of Indian HTEIs:

**Output-Oriented Distance Function SFA Approach**

To follow this approach, we employ equation (3.10) in Section 3.2 above, with two outputs (i.e. \( M = 2 \)) and three inputs (i.e. \( K = 3 \)), where \( y_{1t} = \text{ENR}_t, y_{2t} = \text{TRP}_t, x_{1t} = \text{AS}_t, x_{2t} = \text{NAS}_t \) and \( x_{3t} = \text{LS}_t \). Moreover, to impose the homogeneity constraint, we set \( \omega = (1/y_2) = (1/\text{TRP}) \). Furthermore, to obtain the frontier surface (i.e. the transformation function), we set \( D_{0t} = 1 \), so that the left-hand side of equation (3.10) has a negative sign (see below). Hence, the transformation (i.e. frontier) function can be written as:

\[
\ln(\text{TRP}_t) = \alpha_{0t} + \alpha_{1t} \ln(\text{ENR}_t/\text{TRP}_t) + \alpha_{11t} \ln(\text{ENR}_t/\text{TRP}_t) \ln(\text{ENR}_t/\text{TRP}_t)
\]

\[
+ \beta_{1t} \ln(\text{NAS}_t) + \beta_{2t} \ln(\text{NAS}_t) + \beta_{3t} \ln(\text{LS}_t) + \beta_{11t} \ln(\text{NAS}_t) \ln(\text{AS}_t)
\]

\[
+ \beta_{22t} \ln(\text{NAS}_t) \ln(\text{NAS}_t) + \beta_{33t} \ln(\text{LS}_t) \ln(\text{LS}_t) + \beta_{12t} \ln(\text{NAS}_t) \ln(\text{NAS}_t)
\]

\[
+ \beta_{13t} \ln(\text{AS}_t) \ln(\text{LS}_t) + \beta_{23t} \ln(\text{NAS}_t) \ln(\text{LS}_t) + \delta_{11t} \ln(\text{ENR}_t/\text{TRP}_t) \ln(\text{NAS}_t)
\]

\[
+ \delta_{21t} \ln(\text{ENR}_t/\text{TRP}_t) \ln(\text{NAS}_t) + \delta_{31t} \ln(\text{ENR}_t/\text{TRP}_t) \ln(\text{LS}_t) + (v_{it}-u_{it})
\]

\( i = 1, 2 \ldots \ldots N, t = 1, 2, T. \)

**Input-Oriented Distance Function SFA Approach**

To pursue this approach, we use equation (3.12) in Section 3.2 above, with two outputs (i.e. \( M = 2 \)) and three inputs (i.e. \( K = 3 \)), where \( y_{1t} = \text{TRP}_t, y_{2t} = \text{ENR}_t, x_{1t} = \text{NAS}_t, x_{2t} = \text{LS}_t \) and \( x_{3t} = \text{AS}_t \). Note that the notations associated with the output and input variables in the input-oriented distance function approach differ slightly from the notations used in the output-oriented distance function approach above.

Again, to impose the homogeneity constraint on the input-oriented distance function, we set \( \omega = (1/x_3) = (1/\text{AS}) \). Moreover, to obtain the frontier surface (i.e. the transformation function), we set \( D_{0t} = 1 \), so that the left-hand side of equation (3.13) has a negative sign (see below). Hence, the transformation (i.e. frontier) function can be written as:

\[
\ln(\text{AS}_t) = \alpha_{0t} + \alpha_{11t} \ln(\text{TRP}_t) + \alpha_{12t} \ln(\text{ENR}_t) + \alpha_{111t} \ln(\text{TRP}_t) \ln(\text{TRP}_t)
\]

\[
+ \alpha_{22t} \ln(\text{ENR}_t) \ln(\text{ENR}_t) + \alpha_{12t} \ln(\text{TRP}_t) \ln(\text{ENR}_t) + \beta_{11t} \ln(\text{NAS}_t/\text{AS}_t)
\]

\[
+ \beta_{2t} \ln(\text{LS}_t/\text{AS}_t) + \beta_{11t} \ln(\text{NAS}_t/\text{AS}_t) \ln(\text{NAS}_t/\text{AS}_t)
\]

\[
+ \beta_{22t} \ln(\text{LS}_t/\text{AS}_t) \ln(\text{LS}_t/\text{AS}_t) + \beta_{12t} \ln(\text{NAS}_t/\text{AS}_t) \ln(\text{LS}_t/\text{AS}_t)
\]

\[
+ \delta_{11t} \ln(\text{NAS}_t/\text{AS}_t) \ln(\text{TRP}_t) + \delta_{12t} \ln(\text{NAS}_t/\text{AS}_t) \ln(\text{ENR}_t)
\]

\[
+ \delta_{21t} \ln(\text{LS}_t/\text{AS}_t) \ln(\text{TRP}_t) + \delta_{22t} \ln(\text{LS}_t/\text{AS}_t) \ln(\text{ENR}_t) + (v_{it}-u_{it})
\]

\( i = 1, 2 \ldots \ldots N, t = 1, 2 \ldots \ldots T. \)
I. Estimation and Interpretation of Results

In Table 1 below, we present the descriptive statistics (Mean and Standard Deviation (SD)) pertaining to the output and input variables of our study:

Table 1: Descriptive Statistics Pertaining To Input and Output Variables

<table>
<thead>
<tr>
<th>Period</th>
<th>TRP</th>
<th>ENR</th>
<th>AS</th>
<th>NAS</th>
<th>LS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001-02 Mean</td>
<td>1080.5</td>
<td>3352.67</td>
<td>408.67</td>
<td>1122.34</td>
<td>244376</td>
</tr>
<tr>
<td>SD</td>
<td>377.77</td>
<td>1127.57</td>
<td>46.48</td>
<td>512.22</td>
<td>198301.47</td>
</tr>
<tr>
<td>2002-03 Mean</td>
<td>1225.5</td>
<td>3566</td>
<td>410</td>
<td>1044.17</td>
<td>247284.83</td>
</tr>
<tr>
<td>SD</td>
<td>361.63</td>
<td>1225.98</td>
<td>48.53</td>
<td>487.92</td>
<td>197397.69</td>
</tr>
<tr>
<td>2003-04 Mean</td>
<td>1042.34</td>
<td>4332</td>
<td>410.83</td>
<td>975.34</td>
<td>228625.34</td>
</tr>
<tr>
<td>SD</td>
<td>212.60</td>
<td>1632.15</td>
<td>45.26</td>
<td>464.39</td>
<td>199755.67</td>
</tr>
<tr>
<td>2004-05 Mean</td>
<td>1352.17</td>
<td>3885.17</td>
<td>405.67</td>
<td>973.5</td>
<td>258016.67</td>
</tr>
<tr>
<td>SD</td>
<td>359.95</td>
<td>1184.83</td>
<td>39.73</td>
<td>457.36</td>
<td>197279.65</td>
</tr>
<tr>
<td>Overall Mean</td>
<td>1175.13</td>
<td>3783.96</td>
<td>408.79</td>
<td>1028.83</td>
<td>244575.71</td>
</tr>
<tr>
<td>SD</td>
<td>336.48</td>
<td>1276.47</td>
<td>42.12</td>
<td>452.81</td>
<td>185121.5</td>
</tr>
</tbody>
</table>

We now estimate the technical efficiency of IITs and IISc, by using the input-oriented and output-oriented distance function SFA approaches as well as the DEA methodology.

4.1 Technical Efficiency of IITs and IISc using Input-Oriented and Output-Oriented Distance Function SFA Approaches

Table 2 below presents the estimation results pertaining to the TE of IITs and IISc obtained by using the input-oriented and output-oriented distance function SFA approaches, based on panel data pertaining to IITs and IISc. The table includes the results of ML (Maximum Likelihood) estimation based on the trans-log input-oriented and output-oriented distance functions\(^\text{11}\).

In particular, we have considered the impact of the three input variables, namely, Academic Staff (AS), Non-Academic Staff (NAS) and Library Stock (LS) on the two output variables, i.e. research output (TRP) and teaching output (ENR) of an institution, and have used this relationship to estimate the technical efficiency of IITs and IISc, based on 4-year 'balanced' panel data (2001-02 to 2004-05), following the input-oriented and output-oriented stochastic distance function approaches (see equations (3.25) and (3.24) respectively in Section 3 above):

\(^{11}\) Mostly, the studies on efficiency measurement estimate both the input and output distance functions and present both set of results. The results of the two approaches are usually expected to be different (See Kumbhakar et. al. (2007) for a detailed discussion).
Table 2: ML Estimates of Input-Oriented and Output-Oriented Stochastic Distance Functions for IITs and IISc, ‘Balanced’ Panel Data (2001-02 to 2004-05)

<table>
<thead>
<tr>
<th>Dependent Variables: Log of AS (for Input-Oriented Distance Function)</th>
<th>MLE (t-statistics)</th>
<th>Independent Variables: Output-Oriented Distance Function</th>
<th>MLE (t-statistics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>43.1059 (1.0935)</td>
<td>Constant</td>
<td>-160.677 (-114.645)*</td>
</tr>
<tr>
<td>lnTRP</td>
<td>-4.3098 (-0.9795)</td>
<td>ln(ENR/TRP)</td>
<td>-8.1458 (-1.8402)*</td>
</tr>
<tr>
<td>lnENR</td>
<td>-8.2486 (-1.3096)</td>
<td>ln(ENR/TRP)*ln(ENR/TRP)</td>
<td>0.3597 (2.4997)**</td>
</tr>
<tr>
<td>lnTRP*lnTRP</td>
<td>-0.2115 (-1.1506)</td>
<td>lnAS</td>
<td>73.1998 (17.3743)*</td>
</tr>
<tr>
<td>lnENR*lnENR</td>
<td>0.1010 (0.3874)</td>
<td>lnNAS</td>
<td>-4.3031 (-1.4532)</td>
</tr>
<tr>
<td>lnTRP*lnENR</td>
<td>0.8974 (3.0591)**</td>
<td>lnLS</td>
<td>-4.9642 (-4.1459)*</td>
</tr>
<tr>
<td>ln(NAS/AS)</td>
<td>-5.6176 (-1.5054)</td>
<td>lnAS*lnAS</td>
<td>-5.7586 (-9.8977)*</td>
</tr>
<tr>
<td>ln(LS/AS)</td>
<td>0.8147 (1.0639)</td>
<td>lnNAS*lnNAS</td>
<td>0.1006 (1.5738)</td>
</tr>
<tr>
<td>ln(NAS/AS)*ln(NAS/AS)</td>
<td>0.2841 (2.2738)**</td>
<td>lnLS*lnLS</td>
<td>0.0229 (1.1693)</td>
</tr>
<tr>
<td>ln(LS/AS)*ln(LS/AS)</td>
<td>-0.0268 (-2.2673)**</td>
<td>lnAS*lnNAS</td>
<td>-0.7884 (-2.2121)*</td>
</tr>
<tr>
<td>ln(NAS/AS)*ln(LS/AS)</td>
<td>0.0180 (0.2039)</td>
<td>lnAS*lnLS</td>
<td>0.0475 (0.2852)</td>
</tr>
<tr>
<td>ln(NAS/AS)*lnTRP</td>
<td>0.1111 (1.2504)</td>
<td>lnNAS*lnLS</td>
<td>0.5684 (8.3012)*</td>
</tr>
<tr>
<td>ln(NAS/AS)*lnENR</td>
<td>0.5739 (1.4878)</td>
<td>ln(ENR/TRP)*lnAS</td>
<td>0.6476 (0.9489)</td>
</tr>
<tr>
<td>ln(LS/AS)*lnTRP</td>
<td>-0.0055 (-0.1299)</td>
<td>ln(ENR/TRP)*lnNAS</td>
<td>0.3441 (3.2142)**</td>
</tr>
<tr>
<td>ln(LS/AS)*lnENR</td>
<td>-0.0638 (-0.8085)</td>
<td>ln(ENR/TRP)*lnLS</td>
<td>0.0508 (1.3173)</td>
</tr>
<tr>
<td>(\sigma^2 = (\sigma_u^2 + \sigma_v^2))</td>
<td>0.0113 (1.4697)</td>
<td>(\sigma^2 = (\sigma_u^2 + \sigma_v^2))</td>
<td>0.0328 (1.3716)</td>
</tr>
<tr>
<td>(\gamma = (\sigma_v^2 / \sigma^2))</td>
<td>0.9296 (15.2108)*</td>
<td>(\gamma = (\sigma_v^2 / \sigma^2))</td>
<td>0.9331 (23.8945)*</td>
</tr>
</tbody>
</table>

URL: http://dx.doi.org/10.14738/abr.31.794
<table>
<thead>
<tr>
<th></th>
<th>μ</th>
<th></th>
<th></th>
<th>μ</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>η</td>
<td>-0.6680 (-1.8795)*</td>
<td></td>
<td>η</td>
<td>-1.9627 (-1.1463)</td>
<td></td>
</tr>
<tr>
<td>Log likelihood (d.f.)</td>
<td>44.9972 (09)</td>
<td></td>
<td>Log likelihood (d.f.)</td>
<td>33.8542 (09)</td>
<td></td>
</tr>
<tr>
<td>LR test</td>
<td>3.6305</td>
<td></td>
<td>LR test</td>
<td>17.1537</td>
<td></td>
</tr>
<tr>
<td>R² (OLS)</td>
<td>0.85</td>
<td></td>
<td>R² (OLS)</td>
<td>0.90</td>
<td></td>
</tr>
</tbody>
</table>

*statistically significant at 1% level; **statistically significant at 5% level; *statistically significant at 10% level

Note: All output distance function parameters have been multiplied by (−1) in order to be comparable with other results (Coelli and Perelman, 1996).

In Columns 2 and 4 of Table 2 above, we provide the ML estimates of the transformation (or frontier) function, based on the trans-log input-oriented and output-oriented distance functions respectively. As shown in Column 2, all first order terms in the input-oriented distance function approach have coefficients with the expected sign (negative for output terms and positive for input terms), except the ln(NAS/AS) term (whose coefficient has a negative sign). However, only the second order terms in the input-oriented approach have statistically significant coefficients (specifically, lnTRP*lnENR, ln(NAS/AS)*ln(NAS/AS) and ln(LS/AS)*ln(LS/AS)).

In contrast, as depicted in Column 4, the coefficient of the first order output term in the output-oriented distance function approach has the expected sign (negative) and is weakly statistically significant (i.e. at 10% level of significance). Moreover, two of the first order input terms in the output oriented approach, namely, ln(AS) and ln(LS), have coefficients that are highly statistically significant (i.e. at 1% level of significance), but only of them (i.e. ln(AS)) has a coefficient with the expected sign (positive). Also, the coefficients of several second order terms in the output oriented approach are statistically significant (namely, ln(ENR/TRP)*ln(ENR/TRP), lnAS*lnAS, lnNAS*lnLS and ln(ENR/TRP)*lnNAS), while the coefficient of one second order term (i.e. lnAS*lnNAS) is weakly statistically significant.

In both Columns 2 and 4, σ² is weakly statistically significant, while γ is highly statistically significant and indicates that approximately 93% of the variation in the dependent variable can be explained by the independent variables in both input-oriented and output-oriented SFA approaches. Note that to obtain a better fit, we have restricted μ to zero in both approaches, while η is negative and weakly statistically significant only in the input-oriented SFA approach. Hence, the above findings suggest that according to the input-oriented SFA approach, TE of IITs and IISc is decreasing over time, while according to the output-oriented SFA approach, there is no change in TE of the above institutions over time. Lastly, R² is fairly high in both approaches (0.85 and 0.90 respectively), which indicates a good fit.

Moreover, in the output-oriented distance function approach, the generalized likelihood ratio statistic $\chi^2 (=17.1537)$ exceeds the critical $\chi^2$ value with 3 degrees of freedom (= 7.81 at 95% level). Therefore, we can reject the null hypothesis of traditional half-normal distribution model with full and time invariant TE (i.e. $\gamma = \mu = \eta = 0$). Table 3 below presents the TE estimates of IITs and IISc during the study time period, along with the mean efficiency estimates, obtained by using the input-oriented and output-oriented distance function SFA approaches:
Table 3: Estimates of Technical Efficiency of IITs and IISc by Using Input-Oriented and Output-Oriented Distance Function SFA Approaches

<table>
<thead>
<tr>
<th>Year/Institutions</th>
<th>Input-Oriented Distance Function Approach</th>
<th>Output-Oriented Distance Function Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIT Kanpur</td>
<td>0.9975</td>
<td>0.9951</td>
</tr>
<tr>
<td>IIT Kharagpur</td>
<td>0.9761</td>
<td>0.9540</td>
</tr>
<tr>
<td>IIT Bombay</td>
<td>0.9797</td>
<td>0.9609</td>
</tr>
<tr>
<td>IIT Delhi</td>
<td>0.9982</td>
<td>0.9966</td>
</tr>
<tr>
<td>IIT Roorkee</td>
<td>0.9933</td>
<td>0.9870</td>
</tr>
<tr>
<td>IISc Bangalore</td>
<td>0.9927</td>
<td>0.9859</td>
</tr>
<tr>
<td>Mean Efficiency</td>
<td>0.9896</td>
<td>0.9799</td>
</tr>
</tbody>
</table>

According to the input-oriented distance function approach, IIT Delhi exhibits the highest levels of technical efficiency during the study period (between 99% and 100%), followed by IIT Kanpur (between 98% and 100%), IIT Roorkee and IISc Bangalore (between 95% and 99%), and IIT Bombay (between 86% and 98%). The lowest levels of technical efficiency are attained by IIT Kharagpur (between 84% and 98%). Moreover, the mean technical efficiency for IITs and IISc is fairly high (ranging from 93% to 99%).

However, according to the output-oriented distance function approach, IIT Kanpur exhibits the highest levels of technical efficiency during the study period (between 97% and 100%), followed by and IIT Bombay (between 96% and 100%), IISc Bangalore and IIT Delhi (between 95% and 100%), and IIT Kharagpur (between 73% and 100%). The lowest levels of technical efficiency are achieved by IIT Roorkee (between 71% and 100%). Again, the mean technical efficiency for IITs and IISc is fairly high (ranging from 88% to 100%)\(^\text{12}\).

4.2: Technical Efficiency of IITs and IISc Using DEA (CRS) Methodology

Table 4 below provides the TE values of IITs and IISc obtained by using DEA (CRS) methodology, based on ‘balanced’ panel data pertaining to IITs and IISc. In particular, we have considered the impact of the three input variables, namely, Academic Staff (AS), Non-Academic Staff (NAS) and Library Stock (LS) on two output variables, i.e. research output (TRP) and teaching output (ENR), and have used this relationship to estimate the technical efficiency of IITs and IISc, based on 4-year ‘balanced’ panel data (2001-02 to 2004-05), following DEA (CRS) methodology:

\(^{12}\)In general, the rankings based on input and output oriented approaches are different and most efficiency studies present both set of results. As pointed out by Kumbhakar et al. (2007), “when econometric techniques are used to capture the effect of firms’ inefficiency on the data, assuming that it is either output or input oriented, both the efficiency scores and the estimated technologies differ.”
According to the DEA (CRS) approach, IIT Kanpur, IIT Roorkee and IISc Bangalore exhibit the relatively highest levels of technical efficiency during the study period (i.e. at 100%), followed by IIT Delhi (between 97% and 100%), IIT Kharagpur (between 97% and 100%) and IIT Bombay (between 97% and 98%). The mean efficiency of IITs and IISc is fairly high and ranges from 98.7% to 99.6% during the study period.

4.3 Technical Efficiency of IITs And IISc Using Stochastic Distance Function (SFA) Approach versus DEA Approach

Table 5 below shows the TE estimates of IITs and IISc during the study time period, along with the mean efficiency estimates, obtained by using the input-oriented and output-oriented SFA as well as the DEA (CRS) approaches:

All the methodologies (i.e. input-oriented and output-oriented SFA as well as the DEA (CRS) methodology) yield comparable technical efficiency estimates for IITs and IISc over the study period. For instance, IIT Kanpur shows highest technical efficiency values (between 97% and 100%), followed by and IISc Bangalore (between 97% and 100%), IIT Delhi (between 95% and 97%)

---

### Table 4: Technical Efficiency of IITs and IISc Using DEA (CRS) Methodology, ‘Balanced’ Panel Data (2001-02 to 2004-05)

<table>
<thead>
<tr>
<th>Institutions</th>
<th>Relative Technical Efficiency Using DEA (CRS)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2001-02</td>
</tr>
<tr>
<td>IIT Kanpur</td>
<td>1.000</td>
</tr>
<tr>
<td>IIT Kharagpur</td>
<td>0.971</td>
</tr>
<tr>
<td>IIT Bombay</td>
<td>0.972</td>
</tr>
<tr>
<td>IIT Delhi</td>
<td>0.976</td>
</tr>
<tr>
<td>IIT Roorkee</td>
<td>1.000</td>
</tr>
<tr>
<td>IISc Bangalore</td>
<td>1.000</td>
</tr>
<tr>
<td><strong>Mean Efficiency</strong></td>
<td><strong>0.987</strong></td>
</tr>
</tbody>
</table>

### Table 5: TE of IITs & IISc using Input-Oriented & Output-Oriented SFA & DEA (CRS) Approaches

<table>
<thead>
<tr>
<th>Technical Efficiency</th>
<th>Input-Oriented Distance Function Approach</th>
<th>Output-Oriented Distance Function Approach</th>
<th>DEA (CRS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIT Kanpur</td>
<td>0.9975</td>
<td>0.9951</td>
<td>0.9906</td>
</tr>
<tr>
<td>IIT Kharagpur</td>
<td>0.9761</td>
<td>0.9540</td>
<td>0.9123</td>
</tr>
<tr>
<td>IIT Bombay</td>
<td>0.9797</td>
<td>0.9609</td>
<td>0.9253</td>
</tr>
<tr>
<td>IIT Delhi</td>
<td>0.9821</td>
<td>0.9966</td>
<td>0.9934</td>
</tr>
<tr>
<td>IIT Roorkee</td>
<td>0.9933</td>
<td>0.9870</td>
<td>0.9749</td>
</tr>
<tr>
<td>IISc Bangalore</td>
<td>0.9927</td>
<td>0.9859</td>
<td>0.9727</td>
</tr>
<tr>
<td><strong>Mean Efficiency</strong></td>
<td><strong>0.9896</strong></td>
<td><strong>0.9799</strong></td>
<td><strong>0.9615</strong></td>
</tr>
</tbody>
</table>
IIT Roorkee also exhibits high TE values (between 95% and 100%) for most years during the study period, followed by IIT Bombay (between 86% and 100%) and IIT Kharagpur (between 73% and 100%). However, the TE values pertaining to the SFA approaches exhibit more variability than the TE values obtained via the DEA (CRS) methodology. Moreover, all the three methodologies exhibit high TE values, with mean TE ranging from 88% to 100%.

Furthermore, as noted earlier, the technical efficiency of IITs and IISc is decreasing over time according to the input oriented SFA approach, while there is no change in the technical efficiency of the above institutions over time according to the output oriented SFA approach. Hence, the above TE values suggest that both SFA as well as DEA (CRS) methodologies can be used to measure the technical efficiency of IITs and IISc.

II. Policy Implications and Conclusion
The results of our empirical analysis based on the distance function SFA and DEA (CRS) techniques have important policy implications, which are highlighted below:

IITs and IISc exhibit high levels of technical efficiency (with mean TE values ranging from 88% to 100%), regardless of the methodology employed (SFA or DEA). Hence, the above results suggest that there is ample scope for IITs and IISc to improve their technical efficiency via proper application of the educational resources that are already available to them. Therefore, the above institutions should be encouraged, through better funding mechanisms and other incentives, to improve their technical efficiency levels.

The study also underscores both ‘time-varying’ TE (either increasing or decreasing over time) and ‘time-invariant’ TE in case of the above institutions. We hope that the study will provide useful clues to the concerned institutions and policy makers in India towards raising the performance levels of the above institutions. However, the present study is not free from limitations.

First, given the data availability, we have restricted the selection of input and output variables, which may not be adequate for accurate measurement of technical efficiency of Indian HTEIs. Second, in the case of research output (TRP), equal weights have been given to the different research outputs (i.e. research paper published, books, chapters in books etc.). However, some researchers (such as Cohn et al., 1989) have argued that total research output of a HEI should be measured as a ‘weighted’ average of all the different research outputs of the institution.

However, as Salerno (p. 26, paragraph 2) has pointed out, “specifying the weights a priori requires value judgments as to the respective worth of any given output and there is no substantive basis in the literature for making such judgments.” Johnes (1995), for example, showed in their study of economics departments in the United Kingdom that efficiency scores are highly sensitive to the weight assignments given to different publications like journal articles, books, and book reviews.”

Third, some researchers (such as Nelson and Hevert (1992)) have suggested that a “quality weight” must be attached to the total number of students enrolled (or teaching output (ENR)) in an institution, to account for the variation in the quality of teaching or instruction provided by different higher educational institutions. However, as Salerno (p. 25, paragraph 2) has indicated, “… the opportunity costs of obtaining the necessary data (to assign such a weight) are simply too great for it (the assignment of quality weight) to be feasible.” (Brackets added)

Fourth, in the case of input variables such as library stock (LS), many efficiency studies on higher education have used total library expenditure of an institution as an input variable as
well. However, data pertaining to total library expenditure has not been consistently reported in the annual reports of IITs and IISc over the study period. Therefore, we have not been able to include total library expenditure as an input variable in our study.

REFERENCES


Rediff News (2005) 'IITs, World’s 3rd Best Tech Universities', October 10, 19:01 IST.


