Applying DEA Technique to Library Evaluation in Academic Research Libraries

WONSIK SHIM

ABSTRACT

INCREASINGLY, LIBRARIES ARE ASKED TO JUSTIFY their use of resources in terms of producing meaningful services and impacts to the users and the parent organizations. This study applied an analytical technique called Data Envelopment Analysis (DEA) to calculate the relative technical efficiency of ninety-five academic research libraries that are members of the Association of Research Libraries. Instead of providing the average performance among libraries, DEA, with the proper model of library inputs and outputs, can reveal the best practices in the peer groups, as well as the technical efficiency score for each library. The technique was applied to the libraries using the 1996 and 1997 ARL annual statistics. The study also reviews the applications of DEA technique in the library environment.

INTRODUCTION

Researchers recognize two broad aspects of evaluating library performance: "effectiveness" and "efficiency." Effectiveness here means the extent to which library services meet the expectations or goals set by the organization. In the library field, there has been a growing desire to measure effectiveness in terms of impact of library services on their users.

The second aspect of library performance measurement, "efficiency," measures the library's ability to transform its inputs (resources) into production of outputs (services), or to produce a given level of outputs with the minimum amount of inputs. The efficiency aspect of library performance has received less attention in the library literature, but it is an immediate concern for decision-makers at the parent institution.

The success of the library, like that of other organizations, depends on

Wonsik Shim, School of Information Studies, Florida State University, 101 Shores Building, Tallahassee FL 32306-2100
LIBRARY TRENDS, Vol. 51, No. 3, Winter 2003, pp. 312–332

its ability to behave both effectively and efficiently. We can put these two dimensions of library performance in a 2 by 2 matrix as shown in Figure 1.

Performance improvement requires constant and careful monitoring and assessment of library activities and operating environments. This, in turn, requires the development of proper measurement tools or devices. This study assesses the technical efficiency of academic research libraries that are members of the Association of Research Libraries using a complex tool called DEA. While the development of effectiveness is equally important, this study is focused solely on measuring library efficiency.

Figure 1. Library Performance Matrix Using the Levels of Effectiveness and Efficiency as Two Dimensions.

	High Effectiveness		
Low Efficiency	Effective but Excessively Costly	Best All-around Performers	High Efficiency
	Problematic, Underperforming	Efficiently Managed for Insignificant Results	
	Low Effectiveness		

Data Envelopment Analysis

Overview

Data Envelopment Analysis (DEA) measures the relative efficiencies of organizations with multiple inputs and multiple outputs (Charnes et al., 1978). The individual organizations, teams, or units analyzed are called the decision-making units, or DMUs. The basic point of DEA is to identify the so-called efficient frontier in some comparison set of DMUs. All units on this frontier are said to be operating at 100 percent efficiency. DEA provides an efficiency score for each of the inefficient units, as well as a benchmark set of efficient units that lead to that conclusion. The results of the DEA analysis can be used in performance measurement of libraries, especially for benchmarking purposes.

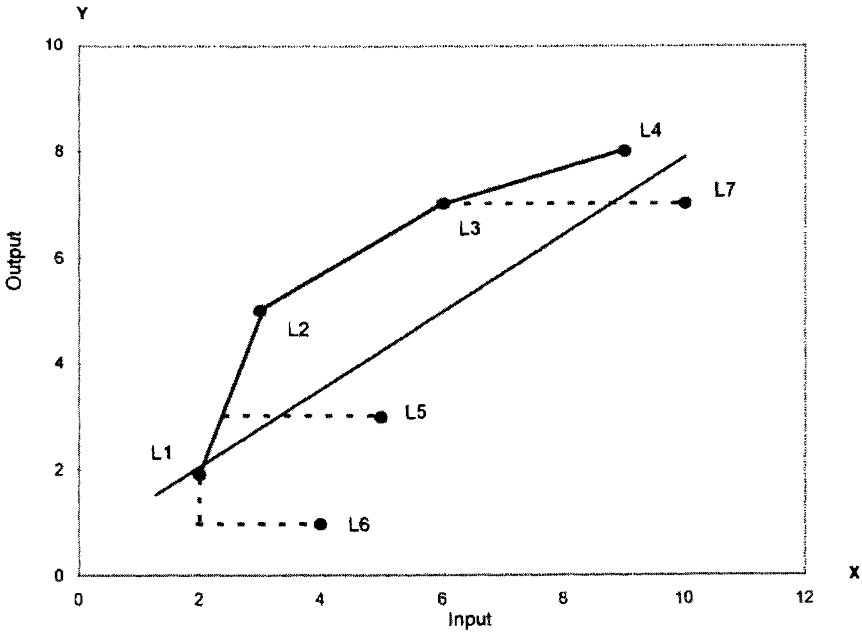
Since the DEA technique was first developed by Charnes, Cooper, and Rhodes in 1978, it has been widely applied to industries as diverse as health care, finance, education, and transportation, as well as many other industries and organizations. The technique is well documented in both the operations research (Banker, Charnes, & Cooper, 1984; Dyson & Thanassoulis, 1988; Golany & Roll, 1989; Cooper, Thompson, & Thrall, 1996) and economics literature (Sengupta, 1987; Banker & Maindiratta, 1988; Seiford & Thrall, 1990; Leibenstein & Maital, 1992). The DEA bibliography compiled by Seiford (1994) includes more than 400 articles, books, and dissertations between 1978 and 1992. A recent bibliography (Emrouznejad, 2001) reports more than 1,000 applications of the DEA technique.

DEA allows the weights of individual inputs and outputs of each DMU to vary until it gives the best possible combination for the focus library. In DEA calculations, through mathematical optimization, each DMU is assigned the weights that maximize its efficiency score. In doing so, DEA gives all the other DMUs "the benefit of the doubt" by allowing them to apply the same weights to see if any of them looks better than the library being evaluated, which is called the "focus" DMU. If the focus DMU looks at least as good as any other DMU, it receives an efficiency score of 1. However, if some other DMU looks better than the focus DMU, even when the weights are calculated in a way that is most favorable to the focus, it will receive an efficiency score less than 1. In DEA, a separate calculation is done for each DMU.

Graphical Illustration

Suppose, for the sake of illustration, we have seven libraries or DMUs that each have only one input and output. We assign these libraries to the coordinate values associated with the points L1 through L7 in Figure 2 where the input is represented on the horizontal axis (X) and the output is represented along the vertical axis (Y).

Figure 2. Envelopment Surface. Adapted from Charnes et al. (1994), p. 33.



For example, library 1 (L1) uses two units of input and produces two units of output. Library 2 (L2) uses 3 units of input to produce 5 units of output. The best a library can do is the top left section of the graph where input is low but output is high. Using the given data, the DEA identifies a set of units in the comparison set (our seven libraries) whose efficiency score equals 1. In the figure, these are the libraries 1 through 4 (L1–L4) because there is nothing to their left. These libraries are called the efficient frontier and define the limits of what a library can achieve in the given situation. In DEA, determination of whether a unit is part of the efficient frontier is based on the units included in the analysis. The heavy line connecting the efficient libraries is called the "envelopment surface" because it envelops all the cases, thus giving the name "Data Envelopment Analysis." Notice also the regression line (the thin line shown in Figure 2) that represents the average relationship between the input (the independent variable) and the output (the dependent variable).

DMUs L5 through L7 are not on the envelopment surface and thus are evaluated as inefficient in the DEA analysis. There are two ways to explain their weakness. One is to say that, for example, library 5 (L5) could be imagined to produce as much output as it does, but with less input. This could be accomplished by moving horizontally until it hits the line between L1 and L2. It should stop there because, with these data, there is no evidence that any unit can do better than that.

One of the assumptions here is that if L1 and L2 can be attained in the real world, then any point between L1 and L2 is also possible. This is called "convexity," which is almost always assumed in economic theory (Farrell, 1957). Mathematically, any point between L1 and L2 represents the weighted average of the two.

Libraries 1 and 2 (L1 and L2) are called the benchmark set for L5 and are interpreted as peers for L5 in DEA. The term "peers" has a special meaning. It is the set of efficient frontiers with which an inefficient unit is compared. We can also say that the units are compared against a virtual DMU on the envelopment surface which produces the same output as the unit being evaluated (which we call the "focus DMU") but with less input. If DEA finds such a DMU, either a real unit or a weighted average of several units, then the focus DMU is regarded as inefficient. If there is no evidence for a given focus DMU that a better virtual DMU exists, the unit is evaluated "technically" efficient because there is no waste of input.

Another way of looking at efficiency is to say that library 5 could produce more output, consuming the same amount of input. This could be accomplished by moving up vertically until it hits the envelopment surface between L2 and L3. Again, for the same reason, it should stop there. This time libraries 2 and 3 become peer libraries for library 5.

We see that there are two possible definitions of efficiency depending on the purpose of the evaluation. One might be interested in possible reduction of inputs (in DEA this is called the input orientation) or augmentation of outputs (the output orientation) in achieving technical efficiency. No matter how efficiency is defined here, library 5 is not efficient. Depending on the purpose of the evaluation, the analysis provides different sets of peer groups to learn from. In the input-oriented evaluation, the efficiency score is the (proportional) reduction of input required to move a unit onto the envelopment surface. In the output-oriented evaluation, DEA software reports the (proportional) augmentation of output that achieves the same purpose.

However, there are times when reduction of inputs or augmentation of outputs is not sufficient. In our example, even when library 6 reduces its input from 4 units to 2, there is still a gap between it and its peer library 1 in the amount of one unit of output. In DEA, this is called the "slack," which means excess input or missing output still exists even after the proportional change in the input or the outputs.

One could argue that instead of taking either input or output orientation, a DMU could be compared to its peer in the nearest point on the envelopment surface. Frei and Harker (1996) investigated this type of optimal projection of inefficient units onto the envelopment hyperplane. The definition of "nearest" requires establishing a relative importance of inputs and outputs. This approach will not be explored further here.

DEA Formulation

The previous section presented several key concepts in DEA. As an evaluation technique, DEA is fairly easy to understand on the abstract level. However, some of its main subtleties are only appreciated if one examines its computational aspects. At present, various software packages are available to facilitate the complex computation required in DEA applications. While these tools alleviate the need for setting up complicated DEA programming runs, some familiarity with the basic DEA model (Charnes et al., 1978) will be useful for further discussion of DEA application in the libraries.

The CCR Ratio Model[1]

Essentially the Charnes-Cooper-Rhodes ratio model (Charnes et al., 1978) can be thought as an extension of the simple efficiency ratio (ouput/input) to situations with multiple inputs and outputs. The efficiency score for a DMU was previously defined as the ratio of the weighted sum of outputs (virtual output) to the weighted sum of inputs. Suppose DMU (j) consumes a vector $X_j = \{x_{ij}\}$ of inputs (i = 1, ..., m) and produces a vector $Y_j = \{y_{rj}\}$ of outputs (r = 1, ..., s), the score for the particular DMU labeled by j_o can be expressed as follows:

In the formula, μ_r represents a set of weights for the outputs and v_i a set of weights for the inputs.

$$\max h_o = \frac{\sum_r \mu_r y_{rj_o}}{\sum_i \nu_i x_{ij_o}}$$

As was noted, there are two constraints on the model:

(1) $h_o \leq 1$ for $j = 1, ..., n$ (n = number of DMUs)
(2) $\mu_r, \nu_i \geq 0$.

The model is expressed in a fractional form which has an infinite number of solutions. For any optimal solution (μ^*, ν^*), any multiple of it still satisfies the constraints. Charnes and Cooper (1962) developed a transformation technique that converts linear fractional optimization into a linear programming (LP) problem.

In linear programming, there is an objective function that serves as the goal to achieve, most often expressed in terms of either maximizing benefits or minimizing costs.

$$\max_{\mu, \nu} h_o = \sum_r \mu_r y_{rj_o}$$

subject to

$$\sum_i \nu_r x_{ij_o} = 1$$

$$\sum_r \mu_r y_{rj} - \sum_i \nu_r x_{ij} \leq 0 \; ; j = 1, \ldots, n$$

$$\mu_r, \nu_i \geq 0$$

Here, the objective function (the first formula) seeks the maximum score of the weighted output. The constraints that accompany the objective function are intended to limit the possible range of the decision variables (μ_r, ν_i), so that the solution is not out of bounds.

DEA calculation requires the solution of n (the number of DMUs) such linear programming problems in the form of a set of m input and s output weights. For each solution, there are n + m + s + 1 constraints to be satisfied. For an analysis of a small number of DMUs, spreadsheet programs such as Microsoft Excel can be used to do the calculations.

For each such linear programming problem (which is called the primal), there is a complementary solution that is calculated from the so-called dual of the problem (Hillier & Lieberman, 1990, pp. 151–191). So the above primal can be converted to:

$$\min \theta$$

subject to

$$\sum_j \lambda_r y_{rj} \geq y_{rj_o}$$

$$\theta j \cdot x_{ij_o} - \sum_j \lambda_j x_{ij} \geq 0$$

$$\lambda_j \geq 0,\ \theta \text{ unconstrained}$$

While both linear programming formulations have equivalent solutions, there are several reasons why solving the dual problem is useful. First, there are only m + s (the number of variables) constraints in the dual problem compared to n + m + s + 1 (the number of variables plus number of DMUs plus one) in the primal problem. So when the analysis involves a large number of DMUs (n), solving the dual is computationally efficient. Second, the variables in the dual have nice interpretations. When a DMU(j_o) is efficient, both θ and λj_o are equal to 1 leaving all the other variables equal to zero. Therefore, θ is the efficiency score for the DMU and tells us that the DMU j_o is efficient. If a DMU is inefficient, then the value for θ will be a positive value less than 1 and the unit will have positive λ values for a set of the other DMUs. In fact, those other DMUs with positive λ are the peers that form the benchmark set for the focus DMU.

DEA contributes to the measurement of efficiency in the following ways. First, in the multiple input-output situations, DEA produces a single technical efficiency score for each unit relative to all other units in the comparison population. If a DMU is operating at 100 percent efficiency, then there is no evidence, at least in the given data, to demonstrate that any other DMU can do better. Second, for each DMU evaluated as less than 100 percent efficient, DEA provides a set of DMUs, which we call the benchmark set, that define the corresponding best practices in the sample. The units included in the benchmark set are efficient, by the DEA definition, and can be used as potential peers from which lessons can be learned. In addition, DEA provides specific recommendations as to how much reduction of inputs or augmentation of outputs, in the form of efficiency gain, would be required to make a unit efficient. It should be noted that the inefficiencies calculated by DEA must be regarded as "potential." Improvement in the efficiency may not be possible due to factors such as significant difference in the service quality or different external operating environments in the compared organizations. To sum up, unlike previous approaches to measuring efficiency, which tend to focus on average performance, DEA provides a viable alternative in which efficiency is defined by units that seem to perform best.

In general, for a given focus, DEA is likely to assign bigger weights to the least-used inputs and to the outputs that are produced most (Sexton, 1986). Units assigning zero weights to some of the inputs and outputs are not uncommon in DEA analysis. This situation is not quite desirable in academic libraries where the production of outputs (services) is not exactly market driven and substitution among outputs or among inputs is not feasible. Several weight restriction schemes have been proposed by Dyson and Thanassoulis (1988), Charnes, Cooper, and Li (1989), and Thompson, Langemeier, Lee, Lee, and Thrall (1990).

The first few chapters in Charnes, Cooper, Lewin, and Seiford (1994) provide an overview of the technical details of DEA.

Comparison of DEA Applications in Libraries

There have been a number of studies that applied DEA technique to the library environment. Table 1 shows a brief comparison of these studies. The table shows that nearly all types of library services have been scrutinized using the technique. It may be difficult to apply the technique to special libraries due to the lack of consistent and comparable data sets. The table also shows that DEA application is not limited to a particular geographic location—different people from different continents have applied DEA to the library environment.

Easun's work appears to be the first one to apply DEA techniques to a library. However, it does not appear that her study influenced subsequent DEA work in libraries; only Shim (2000) cited Easun's dissertation work. The size of the sample varies. For instance, Chen (1997) included all twenty-three university and college libraries in Taipei, Taiwan. Shim (2000) included all U.S. academic libraries that are members of the Association of Research Libraries (ARL). In Worthington's study, 168 public libraries in New South Wales local government were studied. Only in Vitaliano (1998)

Table 1. Comparison of DEA Studies in Libraries.

DEA Application	Library Type	Country	Size of Sample	Data Period	Primary Author's Academic Affiliation
Chen (1997)	Academic	Taipei, Taiwan	23	1995	Economics
Easun (1992)	School	California, USA	74	1985/1986	Library Science
Hammond (Forthcoming)	Public	UK	159	1995/1996	Economics
Sharma, Leung and Zane (1999)	Public	Hawaii, USA	47	1997	Economics
Shim (2000)	Academic	USA	95	1996, 1997	Library Science
Vitaliano (1998)	Public	New York State, USA	184	1992	Economics
Worthington (1999)	Public	Australia	168	1993	Economics

was some form of sampling conducted; only those public libraries that have a single service outlet were evaluated—libraries with branches are omitted due to the difficulty of comparison. Except for Easun (1992) and Shim (2000), the authors of all the other DEA works in libraries have academic affiliation in economics departments. In other words, the library was chosen as a case to apply DEA technique rather that the other way around. Also, most of these works were published outside the library and information science literature, making them difficult to access for library managers who are their intended audiences.

Table 2 allows us to compare the studies in terms of variables included in the DEA models. Except for Worthington (1999), all of the studies have multiple inputs and multiple outputs. Also, four out of seven studies included nondiscretionary input variables. All of these studies included the size of user population as part of the nondiscretionary variables.

For output variables, total circulation and reference transactions were most often used. Although there is a significant difference in terms of the number of variables included, the selection of output variables is fairly consistent—it is a matter of deciding how many, not which variable(s). For input variables, we see a wide variety of variables that include different aspects of library collection (e.g., book collection, net volumes added, serials, audiovisual materials) and library staff. Chen (1997) used library physical characteristics (e.g., physical space and seating). Library expenditure appears only in Worthington (1999)—it was the only input variable used. One interesting item is library service hours. It was used as an output in Chen (1997) but as an input in Sharma, Leung, and Zane (1999) and Vitaliano (1998).

Easun's approach is unique in that she used a three-stage model where output variables in the earlier stages were used as input variables in later stages. For instance, the variables under the provision of information and resource-based instruction were used as output variables in the first stage of her analysis. But in the second stage, those variables were treated as input variables to produce output variables related to library use. The final outputs in her study were student performance in standardized tests. The study may be overly aggressive in the sense that the final outputs are school-related outcomes that are outside the context of DMUs (media centers) under consideration.

In summary, DEA technique has been applied to various types of libraries over the past ten years without being noticed and assessed by researchers and practitioners in the library science field.

Selection of Data

This study used the annual statistics (1996 and 1997) from the Association of Research Libraries (ARL) for the population of ninety-five academic research libraries in the U.S. For the purpose of valid peer comparison,

Table 2. Variables Chosen in Library DEA Studies.

	Outputs	Discretionary Inputs	Nondiscretionary Inputs*
Chen (1997)	Library Visits; Book Circulation; Reference Transactions and Online Search; Patron Satisfaction; Annual Service Hours; Interlending Service	Library Staff; Book Collection; Book Acquisition Expenditure; Library Physical Space; Seating Capacity	None
Easun (1992)	Final Outputs: Student Achievement in Standardized Tests (Math, Reading, and Writing) Intermediate Outputs: Provision of Information (3 Variables); Resource-based Instruction (4); Library Use (3)	Initial Inputs: Human Resources (4 Variables); Material Resources (3 Variables)	None
Hammond (Forthcoming)	Total Circulation; Reference Transactions; Items Requested Processed	Opening Hours; Monographs, Audiovisual Materials; Serials; Newly Added Items	Population Density; Area Size; Resident Population; Outlet Type
Sharma, Leung, & Zane (1999)	Total Circulation; Library Visits; Reference Transactions	Book Collection; Library Staff; Days Open; Total Library Expenditure	None
Shim (2000)	Total Circulation; Reference Transactions; Interlibrary Lending; Interlibrary Borrowing; Library Instruction	Volumes Held; Net Volumes Added; Monographs Purchased; Total Serials; Professional Staff; Support Staff; Student Staff	Total Students; Total Graduate Students; Total Faculty
Vitaliano (1998)	Total Circulation; Reference Transactions	Total Holdings; Weekly Hours; New Books Purchased; Serial Subscriptions	Population Served; Librarian Starting Salary; Director's Salary
Worthington (1999)	Total Circulation	Total Library Expenditure	Population; Area; Non-English Speaking Background; Aged Population; Student Population; Nonresidential Borrowers; Socioeconomic Index

* Nondiscretionary inputs are the inputs that are beyond the control of library administrators. These inputs are included in the DEA formula but are not subject to proportional reduction during the efficiency score calculation.

the libraries are grouped by the main funding source of the parent institutions (publicly funded versus privately funded). A total of five output variables were selected, encompassing all the service measures reported in the statistics: interlibrary loans, interlibrary borrowings, reference transactions, total circulation, and library instruction. On the input side, the study includes two types of variables, discretionary and nondiscretionary. "Discretionary" variables include two main resources libraries use to provide services: materials (4 variables), and staff (3 variables). "Nondiscretionary" variables, which are beyond the control of the library administrator, include measures of the number of library users in several categories. They are treated as input variables because they help to determine how much service the library can provide. While the inclusion of the user populations as input variables seems to suggest that the market being served is used as an input, the rationale for their inclusion is that the level of use is a function of the size of the user population being served and that the DEA model accommodates these variables as a special kind of input variable and does not alter (or manipulate) the figures of user populations in its computations of best possible scenarios for each DMU. This study focused on inefficiencies in inputs; the DEA recommendations are represented as in the calculated input reduction for libraries deemed inefficient:

Output Variables (5):

- Total number of interlibrary lending transactions filled (ILLTOT).
- Total number of interlibrary borrowing transactions filled (ILBTOT).
- Number of people who participated in group presentations or instructions (PRESPTCP).
- Number of reference transactions excluding directional questions (REFTRANS).
- Total number of circulation including renewals (TOTCIRC).

Input Variables (10):

Collection Characteristics (Discretionary)

- Total volumes held (VOLS).
- Net volumes added during the period (VOLSADN).
- Monographs purchased, in volumes (MONO).
- Total number of current serial copies (CURRSER).

Staff Characteristics (Discretionary)

- Number of full-time, professional staff (PRFSTF).
- Number of full-time, support staff (NPRFSTF).
- Number of full-time equivalents of hourly student employees (STUDAST).

University Characteristics (Nondiscretionary)

- Total full-time student enrollment (TOTSTU).
- Total full-time graduate student enrollment (GRADSTU).
- Total full-time instructional faculty (FAC).

Scaling of Data

The data values are in a wide range; volumes held are in the millions whereas the numbers of professional staff and staff assistants are in the hundreds or in the tens. The wide range of values—in one input and output, or in a particular variable across the units—can produce a so-called ill-conditioned matrix that causes computational difficulties (Ali, 1994). Therefore, the study applied scale changes for each variable, so that the scaled data fall below 100. Table 3 shows the ranges of each variable before and after scaling. The same scaling was applied to both 1996 and 1997 data.

Constraints on Weights

Because DEA allows the weights of both the inputs and the outputs of each DMU to vary until it gives the best possible combination for the focus library, the resulting weights will not always make much sense. To make the DEA analysis more reasonable, there should be some boundary (technically called a constraint) to limit the relative weight or importance of various inputs and of various outputs.

In the DEA literature, Charnes et al. (1989), Dyson & Thanassoulis (1988), and Thompson et al. (1990) applied various schemes for restricting the relative size of the possible weights. We follow the "Assurance Region" approach developed by Thompson et al. In this approach, instead of

Table 3. Scaling of Data.

		Original Data (1996)		Applied	After Scaling (1996)	
Category	Variable Name	High	Low	Scale	High	Low
Input	VOLS	13,143,330	1,606,642	200,000	65.72	8.03
	VOLSADN	248,156	22,381	3,000	82.72	7.46
	MONO	138,406	—	2,000	69.20	0.00
	CURRSER	96,353	10,284	1,000	96.35	10.28
	PRFSTF	402	36	5	80.40	7.20
	NPRFSTF	589	53	8	73.63	6.63
	STUDAST	222	6	3	74.00	2.00
	TOTSTU	52,637	3,988	600	87.73	6.65
	GRADSTU	11,592	1,198	150	77.28	7.99
	FAC	3,186	390	40	79.65	9.75
Output	ILLTOT	248,741	1,988	3,000	82.91	0.66
	ILBTOT	74,598	1,702	1,000	74.60	1.70
	PRESPTCP	42,222	—	1,000	42.22	0.00
	REFTRANS	1,161,212	—	15,000	77.41	0.00
	TOTCIRC	2,690,871	—	30,000	89.70	0.00

imposing a single set of weights, which is unrealistic, a range of weights in the form of the ratios between the weights is applied to the weight selection process. This approach will effectively limit the movement of the weights in a more realistic range and potentially improve the validity of the DEA analysis. The introduction of the constraints on the weights is expected to decrease the number of efficient DMUs.

Halme, Joro, Korhonen, Salo, and Wallenius (1999) argued against the use of constraints on the weights, proposing instead to use the explicit preferences of the decision-makers. This would make sense in a situation where the DMUs included in the comparison set are all under the control of the same centralized decision-makers. However, this is not applicable to this study population, as the data do not include the information regarding the preferences of library directors or decision-makers at the universities on the proposed inputs and outputs.

While DEA permits each library to "rearrange the world" so that it looks as efficient as possible, there are nonetheless some limitations on the distortions that are permitted. For example, if a staff person costs $40,000/ year (the person's yearly salary) and a book costs $50 (purchasing), it would be unreasonable to let the DEA program set their weights or multipliers equal in determining the combined virtual input. A sensible approach might be to examine available data, and allow large, but not outrageous, variation around the median value reported in the literature. For example, the numbers given would lead to a nominal ratio of 40,000/50 = 800. In applying this ratio, we will adopt two approaches. One is to permit a range from 200 (one quarter of the observed value) to 3,200 (four times the observed value). We call this the *four-fold range.* This seems extremely generous. Under a two-fold range this ratio would be allowed to vary from a low of 400 (half of the observed value) to a high of 1,600 (two times the observed value). The justification for varying degrees of range is based on the reports in the benchmarking literature that the observed performance difference among different organizations could be as large as a factor of several hundredfold (Boxwell, 1994; Zairi, 1996).

The literature reports a wide range of cost figures for the same service category. The studies listed in Table 4 were consulted for guidelines in deriving service costs. Please note that this study uses the cost of each service as the basis for its relative weight in comparison to other services. Similarly, the cost of inputs and their ratios were obtained directly from the ARL statistics. These are summarized in Table 5.

Analysis of Results

The data was analyzed using the commercial program called IDEAS.[2] Additional statistical analyses were conducted to delineate the characteristics of libraries evaluated to be efficient.

Table 4. Cost of Services with Consulted Sources.

Source	Description	Cost Reported	Year	Adjusted for 1997*
(1) Reference				
Cable (1980)	Average Cost of Search (Excluding Hidden Costs)	$5.18	1980	$16.36
Spencer (1980)	Reference Queries	$2.52	1980	$7.96
	Extended Reference Queries	$4.57	1980	$14.44
	Consultation, Training, Tours	$9.09	1980	$28.71
Kantor (1986)	Query	$14.00	1982/3	$37.34
Cochrane & Warmann (1989)	Full Cost Reference	$9.22	1989	$15.84
Robinson & Robinson (1994)	Average Total Cost per Reference Question Handled	$6.84	1994	$8.38
				$18.43 (Average)
(2) Interlibrary Loans				
Roche (1993)	Borrowing	$18.62	1992	$26.12
[data from 1992] ARL/RLG average	Lending	$10.93	1992	$15.33
(3) Circulation				
Kantor (1986)	Per Circulation Cost (Includes Collection Cost)	3.72	1982/3	6.13
(4) Group Presentation				
From ARL Statistic (per participant)	Average Hourly Rate of Professional Staff (1996)	$34.96	1996	$37.41
	Assuming 2 Hours and 14 Attending per Session	$4.99	1996	$5.34

Note: * Applied 7 percent annual increase except for circulation (3.5 percent).

Table 5. Cost Information for Inputs.

Year	Category	Units*	Total Cost*	Unit Cost
1996	Professional Staff	8,242	$332,752,579	$40,373
	Nonprofessional Staff	14,705	$313,687,653	$21,332
	Student Assistants	7,469	$74,137,023	$9,926
	Monographs Purchased	2,889,585	$173,567,824	$60
	Serials (Current)	2,762,558	$319,589,674	$116
1997	Professional Staff	8,349	$350,265,615	$41,953
	Nonprofessional Staff	14,702	$326,773,412	$22,226
	Student Assistants	7,667	$76,831,246	$10,021
	Monographs Purchased	2,815,990	$176,298,928	$63
	Serials (Current)	2,783,810	$346,120,125	$124

Note: * Total of 95 libraries.

Efficiency Scores

Table 6 summarizes the number of inefficient libraries revealed in different evaluation environments.

Reading the table from left to right, there is a marked change both in the number of libraries evaluated inefficient (efficiency score $q < 1$) and the average efficiency scores. As the number of inefficient libraries goes up, the average efficiency score goes down. For instance, in 1996, without any constraints, about 28 percent (= 18/65*100) of the libraries in the public group were evaluated inefficient, whereas with the strictest constraint environment (two fold range, both input and output ratios), about two thirds (= 43/65) of the libraries are evaluated inefficient. The average efficiency score fell from .96 to .83 accordingly. In the private group, again in 1996, the number of inefficient libraries increased from 3 to 11, and the average efficiency score decreased from .98 to .91.

Another noticeable change is that, as we expected, the narrower range (two-fold) will always find more inefficient libraries than the more generous range (four-fold). For instance, in 1997, imposing the four-fold range revealed thirty-three inefficient libraries in the public university group while the two-fold range revealed forty-one inefficient libraries.

The two-fold range seems to provide the reasonable discriminating capability that is required of an evaluation tool. Still, there are some differences in the two comparison groups. Under this particular constraint environment about two-thirds of the libraries in the public group seem to have some other libraries in the same peer group to learn from. On the other hand since two-thirds of the libraries are evaluated efficient in the private group, only about one-third of them will have peers to learn from. This difference should not be interpreted as an indication that academic libraries at the privately funded universities are better managed than their peers are at the publicly funded institutions.

Table 6. Number of Libraries Evaluated Inefficient and Average Efficiency Score under Different Constraints.

			Constraints	
Year	Group	No Constraint	Four-fold range (1/4–4)	Two-fold range (1/2–2)
1996	Public	18	34	43
		(0.96)	(0.90)	(0.83)
	Private	3	7	11
		(0.98)	(0.94)	(0.91)
1997	Public	16	33	41
		(0.96)	(0.90)	(0.84)
	Private	1	7	12
		(0.99)	(0.94)	(0.89)

Note: Public (n = 65), Private (n = 30). The numbers in the parentheses are the average efficiency scores.

The difference might have been simply due to the relative number of units included in the analysis and the density of the observed data values. If the number of units in the analysis is large, then the competition among the units is more severe than with a smaller number of units. Also, if the observed data values are not concentrated, meaning that there is a greater variation of the size of the libraries, more libraries are likely to become somehow unique, and thus become efficient for no other merit. It is expected that the libraries in the public group are more homogeneous in terms of their observed data values than the libraries in the private group.

Tables 7 and 8 show the rankings of the ARL libraries in terms of their efficiency scores. Random codes are used in place of the names of the institutions to keep their identities confidential. One of the considerations for not revealing the identities is that the DEA technique is only one way of measuring library efficiency. The DEA results need to be accompanied by other measures and data collection methods (e.g., site visits or interviewing library staff) to get a detailed picture of the libraries.

Through a series of sensitivity analyses, this study explored the relative impacts of the variables included in the study on the efficiency scores. Among output variables, removal of reference transactions and circulation variables made the biggest changes on the efficiency scores. All input variables seemed

Table 7. Rank Order by Efficiency Score for the Public Group (1996).

Rank	Library	Efficiency Score	Rank	Library	Efficiency Score	Rank	Library	Efficiency Score
1	L01	1.00	23	L26	0.99	45	L42	0.74
1	L02	1.00	24	L04	0.99	46	L37	0.73
1	L03	1.00	25	L90	0.97	47	L84	0.72
1	L08	1.00	26	L22	0.97	48	L83	0.72
1	L10	1.00	27	L19	0.96	49	L12	0.71
1	L17	1.00	28	L46	0.96	50	L18	0.71
1	L20	1.00	29	L25	0.94	51	L38	0.71
1	L23	1.00	30	L09	0.93	52	L72	0.70
1	L28	1.00	31	L62	0.93	53	L44	0.69
1	L30	1.00	32	L69	0.88	54	L35	0.67
1	L31	1.00	33	L50	0.87	55	L75	0.66
1	L34	1.00	34	L81	0.85	56	L32	0.61
1	L47	1.00	35	L15	0.85	57	L85	0.60
1	L48	1.00	36	L70	0.83	58	L71	0.59
1	L65	1.00	37	L16	0.83	59	L07	0.56
1	L68	1.00	38	L45	0.80	60	L40	0.55
1	L73	1.00	39	L27	0.79	61	L41	0.52
1	L78	1.00	40	L33	0.78	62	L63	0.50
1	L79	1.00	41	L57	0.78	63	L89	0.48
1	L87	1.00	42	L55	0.77	64	L51	0.48
1	L92	1.00	43	L64	0.77	65	L91	0.44
1	L94	1.00	44	L21	0.75			

Table 8. Rank Order by Efficiency Score for the Private Group (1996).

Rank	Library	Efficiency Score	Rank	Library	Efficiency Score	Rank	Library	Efficiency Score
1	L05	1.00	1	L61	1.00	21	L52	0.94
1	L06	1.00	1	L66	1.00	22	L82	0.89
1	L11	1.00	1	L67	1.00	23	L49	0.83
1	L13	1.00	1	L76	1.00	24	L39	0.81
1	L29	1.00	1	L77	1.00	25	L74	0.81
1	L43	1.00	1	L80	1.00	26	L93	0.73
1	L56	1.00	1	L86	1.00	27	L53	0.65
1	L58	1.00	1	L88	1.00	28	L14	0.64
1	L59	1.00	1	L95	1.00	29	L54	0.57
1	L60	1.00	20	L24	0.94	30	L36	0.40

to affect the efficiency scores to more or less the same degree. However, taking out a variable sometimes can have a huge effect on individual libraries either by decreasing the efficiency scores substantially or by changing their efficiency status, from efficient to inefficient. The selection of variables is not purely a technical issue. For practical, wide applications of DEA, it is recommended that the full set of variables be retained in the analysis.

In addition to sensitivity analysis, this study added random noise in the data and observed the resulting changes in the efficiency scores and the efficiency status. Four simulations of noise were conducted for each year. In each simulation, every observed data element was subject to a random distortion, causing it to vary according to a normal distribution in which the mean is the original value and the standard deviation is 5 percent of its true value. The results are remarkably consistent in terms of changes in the mean scores (.02–.03 for public, .01–.05 for private). The number of libraries that changed their efficiency status was from 4 to 7 in the public group, from 1 to 5 in the private group. Furthermore, the technique is fairly robust despite the presence of random dummy variables.

In conclusion, the DEA technique can be successfully implemented in research libraries in the U.S. This study provides a baseline approach, as well as results that can be further extended to studies using similar techniques to investigate the problem of assessing library efficiency.

Fluctuation of Efficiency Scores Over Time

Library statistics are extremely stable. The biggest median change of all fifteen variables over a two-year (1996–1997) period was 5 percent. All the input variables, on average, changed by less than 3 percent during the same period, most of them by less than 1 percent. Therefore, it would be logical to expect that the efficiency scores will stay more or less the same. If there was too much fluctuation, it would be a threat to the technique's reliability and validity.

Table 9 shows the consistency of efficiency scores and the efficiency status over a two-year period.

Table 9. Consistency of Efficiency Scores over Time.

	Public (n = 65)	Private (n = 30)
Mean Efficiency Score Change	.06	.07
Libraries With Less Than .05 Change	52	22
Efficiency Status Change	14	7

The mean efficiency score changed on average by 6 percent for the public group and by 7 percent for the private group. For the majority of libraries, there was either no change or less than a 5 percent change. However, the composition of the efficient frontier, measured by the number of libraries that change their efficiency status, shows a moderate change. Close examination of the results shows that significant changes accompanied changes in the observed data values of a similar magnitude. These results demonstrate that the DEA technique produces quite reliable results and can be used to track efficiency over an extended period of time.

Characteristics of Efficient Libraries

This study looked for the variables or library characteristics that are closely associated with libraries with high efficiency scores.

For the public group, libraries with large net volumes added and professional staff tend to have lower efficiency scores. On the other hand, libraries producing more reference transactions and circulation are more likely to be assigned higher efficiency scores. For the public group, the total circulation was the only statistically significant predictor of efficiency scores over a two-year period.

When all fifteen variables included in this study were used in the regression analyses to predict efficiency scores, a substantial portion of variation in the scores (in both groups) was accounted for by the model with R^2 values ranging between .72 and .80. However, when only a subset of the variables is used, such as input variables, output variables, staff variables, collection variables, or user variables, the R^2 measures deteriorate quite rapidly.

The amounts of library expenditures per student and per faculty were not significant predictors of efficiency scores. However, in the public group, the size of the library budget was a significant predictor. Libraries with a smaller budget were more likely to be assigned higher efficiency rating. This was not the case in the private group.

Interestingly, none of the per-user activities, measured by the number of various service outputs per student, was a significant predictor of the efficiency scores in either of the comparison groups.

Among the measures of library resource utilization, for the public group, the number of reference transaction handled per professional staff was a significant predictor. For the private group, libraries a with larger proportion of total volumes actively circulated tend to have higher efficiency scores.

Finally, as expected, university libraries that have both law and medical libraries tend to have lower mean efficiency scores (.79 for public, .87 for private) than libraries with neither (.87, 1.00 respectively) due to increased resource requirement. However, the differences were not statistically significant.

Conclusions and Next Steps

DEA seems to have the flexibility and expandability that other traditional measures lack. It provides a technical means to take a closer look at ways in which libraries can improve their performance. The approach is to look at other libraries, not the ones that are simply big or conventionally "good," but the ones that function efficiently and from which better ways of doing things can be learned. However, this does not mean that the results are directly transformed into actionable recommendations in the real world. On the contrary, there are a host of issues that need to be considered. Two practical areas can be addressed to make progress on these issues.

First and foremost, although the DEA technique has some intuitive appeal, it is difficult to understand its formulations and some of the subtleties related to interpretation of key measures produced. This is the same problem that other applications of operations research techniques have suffered. McDonald & Micikas (1994) noted that the complexity of the models, the arbitrary, unverified assumptions, and the lack of adequate definitions involved in such research are the main stumbling blocks that hinder widespread use of the tools of operations research. One of the ways to address this issue is to form a small group of libraries that agree to adopt DEA as a model to assess the library as a whole or a specific service and collaborate with researchers who are familiar with the technique. As previously noted, most of the DEA applications in libraries were initiated by economists without much interaction with the libraries being evaluated. It is conceivable that the technical complexities can be overcome once the library field has the initiative and forms a nucleus of practitioners who are versed in the applications of the technique.

The second practical issue is that while DEA can provide a way to identify best practices for the purpose of benchmarking, the results need to be verified through followup examination—for example, case studies. The results from DEA analyses in most cases are suggestive rather than confirmatory. A followup is necessary to find out how the best practicing libraries (i.e., efficient libraries) achieve what they do and how other libraries can learn useful lessons by observing and adopting the processes that enabled the efficient libraries. For this reason, it is recommended that instead of assessing the library as a whole—which was the case for all DEA applications in the libraries identified—it might be more meaningful to investigate a particular library operation or function (e.g., cataloging, reference service, digital content creation, and so on). This way, the libraries being evaluated can determine input and output variables more precisely and gain more useful results. After all, what goes in determines what comes out, and this is especially true in DEA applications.

Acknowledgments

This research was supported in part by the Council on Library and Information Resources (CLIR) under Grant #6607. The author would like to thank Professor Paul B. Kantor for his encouragement and thoughtful discussions during the study.

Notes

1. Here we use the input orientation model for the purpose of illustration. An analogous formulation is possible for the output orientation model.
2. Version 5.1, available from Software 1 Consulting Inc., P.O.Box 2453, Amherst, MA 01004–2453.

References

Ali, A. I. (1994). Computational aspects of DEA. In Charnes, A., Cooper, W. W., Lewin, A. Y., & Seiford, L. M. (Eds.), *Data Envelopment Analysis: theory, methodology, and application,* (pp. 63–88). Boston: Kluwer Academic Publishers.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science, 30*(9), 1078–1092.

Banker, R. D., & Maindiratta, A. (1988). Nonparametric analysis of technical and allocative efficiencies in production. *Econometrica, 56*(6), 1315–1332.

Cable, L. G. (1980). Cost analysis of reference service to outside users. *Bulletin of the Medical Library Association, 68*(2), 247–248.

Charnes, A., & Cooper, W. W. (1962). Programming with linear fractional functionals. *Naval Research Logistics Quarterly, 9*(3/4), 181–185.

Charnes, A., Cooper, W. W., Lewin, A. Y., & Seiford, L. M. (1994). *Data envelopment analysis: Theory, methodology and application.* Boston: Kluwer Academic Publishers.

Charnes, A., Cooper, W. W., & Li, S. (1989). Using DEA to evaluate the efficiency of economic performance by Chinese cities. *Socio-Economic Planning Sciences, 23,* 325–344.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operations Research, 2,* 429–444.

Chen, T. Y. (1997). An evaluation of the relative performance of university libraries in Taipei. *Library Review, 46*(3), 190–201.

Cochrane, L.S., & Warmann, C. (1989). Cost analysis of library services at Virginia Polytechnic Institutite and State University. In Fennell, J. C. (Ed.), *Building on the first century: Proceedings of the fifth national conference of the Association of College and Research Libraries,* (pp. 55–62). Chicago, IL: Association of College and Research Libraries.

Cooper, W. W., Thompson R. G., & Thrall, R. M. (1996). Introduction: Extensions and new developments in DEA. *Annals of Operations Research, 66,* 3–45.

Dyson, R. G., & Thanassoulis, E. (1988). Reducing weight flexibility in Data Envelopment Analysis. *Journal of the Operational Research Society, 39*(6), 563–576.

Dyson, R. G., Thanassoulis, E. & Boussofiane, A. (1990). *A DEA (Data Envelopment Analysis) tutorial* [Online]. Available from http://www.warwick.ac.uk/~bsrlu/.

Easun, M. S. (1992). "Identifying efficiencies in resource management.: An application of data envelopment analysis to selected school libraries in California." Ph.D. Diss., University of California, Berkeley.

Emrouznejad, A. (2001). *An extensive bibliography of Data Envelopment Analysis (DEA), Volume I: Working Papers.* Business School, University of Warwick, Coventry CV4 7AL, England. [Online]. Available from http://www.deazone.com/bibliography/index.htm.

Frei, F. X. & Harker, P. T. (1996). *Projections onto efficient frontiers: Theoretical and computational extensions to DEA.* Philadelphia: Financial Institutions Center at the Wharton School, University of Pennsylvania.

Halme, M., Joro, T., Korhonen, P., Salo, S., & Wallenius, J. (1999). A value efficiency approach to incorporating preference information in Data Envelopment Analysis. *Management Science, 45*(1), 103–115.

Hammond, C. J. (2002). Efficiency in the provision of public services: A Data Envelopment Analysis of UK public library systems. *Applied Economics, 34*(5), 649–657.

Hillier, F. S., & Lieberman, G. J. (1990). *Introduction to operations research.* New York: McGraw-Hill.

Kantor, P. B. (1986). Three studies of the economics of academic libraries. In McCabe, G. B., & Kreissman, B. (Eds.), *Advances in Library Administration and Organization, 5,* 221–286. Greenwich, CT: JAI Press.

Leibenstein, H., & Maital, S. (1992). Empirical estimation and partitioning of X-inefficiency: A Data-Envelopment approach. *American Economic Review, 82*(2), 428–433.

McDonald, J. A., & Micikas, L. B. (1994). *Academic libraries: The dimensions of their effectiveness.* Westport, CT: Greenwood Press.

Robinson, B. M., & Robinson, S. (1994). Strategic planning and program budgeting for libraries. *Library Trends, 42,* 420–447.

Roche, M. M. (1993). *ARL/RLG Interlibrary Loan Cost Study.* Washington, DC: Association of Research Libraries.

Seiford, L. M. (1994). A DEA bibliography (1978–1992). In Charnes, A., Cooper, W. W., Lewin, A. Y., and Seiford, L. M. (Eds.), *Data Envelopment Analysis: theory, methodology, and application* (pp. 437–469). Boston: Kluwer Academic Publishers.

Seiford, L. M., & Thrall, R. M. (1990). Recent developments in DEA: The mathematical programming approach to frontier analysis. *Journal of Econometrics, 46,* 7–38.

Sengupta, J. K. (1987). Production frontier estimation to measure efficiency: A critical evaluation in light of Data Envelopment Analysis. *Managerial and Decision Economics, 8*(2), 93–99.

Sexton, T. R. (1986). The methodology of Data Envelopment Analysis. In Silkman, R. H. (Ed.), *Measuring efficiency: An assessment of Data Envelopment Analysis,* (pp. 7–29). San Francisco: Jossey Bass.

Sharma, K. R., Leung, P., & Zane L. (1999). Performance measurement of Hawaii state public libraries: An application of Data Envelopment Analysis (DEA). *Agricultural and Resource Economics Review, 28*(2), 190–198.

Shim, W. (2000). Assessing technical efficiency of research libraries. *Advances in Library Administration and Organization, 17,* 243–339.

Spencer, C. C. (1980). Random time sampling with self-observation for library cost studies. *Bulletin of the Medical Library Association, 68*(1), 53–57.

Thompson. R. G., Langemeier, L. M., Lee, C-T., Lee, E., & Thrall, R. M. (1990). The role of multiplier bounds in efficiency analysis with application to Kansas farming. *Journal of Econometrics, 46,* 93–108.

Vitaliano, D. F. (1997). X-efficiency in the public sector: The case of libraries. *Public Finance Review, 25*(6), 629–643.

Vitaliano, D. F. (1998). Assessing public library efficiency using Data Envelopment Analysis. *Annals of Public and Cooperative Economics, 69*(1), 107–122.

Worthington, A. (1999). Performance indicators and efficiency measurement in public libraries. *Australian Economic Review, 32*(1), 31–42.