Evaluation of digital libraries : A DEA approach

Wonsik Shim, Paul B. Kantor

Alexandria Project Laboratory
School of Communication, Information and Library Studies
Rutgers University
4 Huntington St. New Brunswick, New Jersey 08903
{wshim, kantor}@scils.rutgers.edu

Abstract

As libraries evolve from paper based to digitized collection, traditional measurement activities must change. To demonstrate the growth in library value during this transition period, libraries must be able to describe how library inputs are transformed into the services libraries render. We apply a complex tool, data envelopment analysis (DEA), to evaluate the relative efficiency of major academic research libraries that are members of the Association of Research Libraries (ARL). An efficient library is defined as the one which produces same output with less input or, for a given input, produces more output. We report the results of a two-year base line study using traditional measures taken from 1995-1996 ARL statistics. We observe the patterns of efficiency scores of both individual libraries and libraries in peer groups (private vs. public). In particular we study the consistency over the years of specific DEA measures. This consistency provides justification for extending DEA as libraries undergo revolutionary digital transformation. The results are also corroborated using standard statistical measures. DEA application in the new digital library environment is discussed.

*This research is supported in part by the Council on Library and Information Resources (CLIR) under Grant #6607, and by the Alexandria Project Laboratory at SCILS, Rutgers University.

1. INTRODUCTION

Higher education has come under increasing pressure to demonstrate its value and performance (Council for Aid to Education, 1997; Michalko, 1993). As a result many universities and colleges have begun to look at a wide array of performance and accountability measures. These institutions are not only interested in their own (absolute) accomplishments but also want to know how well they perform when compared to peer institutions. As a sub-unit or department in universities, academic libraries "feel the pressures" (Kyrillidou, 1998, p.4) from the parent institutions to produce evidence on 1) how well the resources are utilized in terms of producing meaningful outputs, and 2) how well the library compares or competes with other libraries within some peer group.

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 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.

	High Effectiveness		
Low Efficiency	Effective But excessively costly	Best, all-around Performers	High Efficiency
	Problematic, Underperforming	Efficiently managed for insignificant results	
	Low Effectiveness		

Figure 1. Library performance matrix using the levels of effectiveness and efficiency as two dimensions.

2. THE DATA ENVELOPMENT ANALYSIS

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 first of all 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% efficiency. DEA provides an efficiency score for each of the inefficient units, as well as benchmark set of efficient units that lead to that conclusion. The results of the DEA analysis can be used in performance measurement of academic research libraries, especially for benchmarking purpose.

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 and many other industries and organizations. The technique is well documented in both the operations research (Banker, Charnes, & Cooper, 1984; Cooper, Thompson & Thrall, 1996; Dyson & Thanassoulis, 1988; Golany & Roll, 1989) and economics literature (Banker & Maindiratta, 1988; Leibenstein & Maital, 1992; Seiford & Thrall, 1990; Sengupta, 1987). The DEA bibliography compiled by Seiford (1994) includes more than 400 articles, books and dissertations between 1978 and 1992.

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. But 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, then it will receive an efficiency score less than 1. In DEA, a separate calculation is done for each 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% 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% 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 viable alternative where 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 Charnes, Cooper, and Li (1989), Dyson and Thanassoulis (1988), and Thompson, Langemeier, Lee, Lee, and Thrall (1990).

DEA, like other evaluation methods, is not without limitations. First, since the definition of the efficiency score is based on extreme individual observations, the results can be very sensitive to measurement error. Second, the efficiency scores are quite sensitive to variable specification and the size of the comparison set (Dyson, Thanassoulis, & Boussofiane, 1990). Third, in DEA a unit can be made appear efficient simply by adjusting its weights. This situation is not likely in reality. Finally, DEA tends to be very difficult to understand, especially its employment of weighting scheme and the mathematical formulations.

Shim & Kantor (1998b) and the first few chapters in Charnes, Cooper, Lewin and Seiford (1994) provide overview of technical details of DEA.

3. METHODOLOGY

3.1 Selection of data

It is fortunate that the ARL statistics lend themselves to DEA analysis. The data are widely used by member libraries and by other researchers in the field. There are several specific reasons why we use these statistics for the current research. First, compared to other library statistics, the ARL data are considered more reliable. Second, compared to other sets of library statistics, the ARL statistics are more complete and provide a wide range of information on both inputs and outputs. Third, the data are readily accessible from the ARL at its Internet web site. Thus, the results can be easily replicated and widely used. It is also easy to obtain and transform the data into a format suitable for spreadsheet programs (e.g., Lotus 1-2-3 or Microsoft Excel) and statistical programs (e.g. SPSS). For the purpose of this study, we will limit the analysis to only U.S. institutions (N=95).

We will apply the DEA technique for two consecutive years (95-96 and 96-97) to investigate the stability (and, if libraries are presumed stable, the reliability) of evaluation using DEA. If DEA is revealing a stable characteristic of the libraries, and is measuring it reliably, then the efficiency scores of each should be similar from one year to the next. The libraries in the ARL study population are broken into two groups; public (n=65) and private (n=30) for fair comparison.

3.2 Inputs and Outputs

The current study is based on the use of the following variables. Variables marked as non-discretionary are the ones that are beyond the control of the library managers and are not subject to proportional reduction when a library is being "moved" to the efficient frontier.

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 (LIBINST).
- Number of reference transactions excluding directional questions (REFTRANS).
- Total number of circulation including renewals (TOTCIRC)
-

Input Variables (10):

<u>Collection Characteristics</u> (Discretionary except for VOLS)

- 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 (Non-discretionary)

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

3.3 The constraints on weights

Because 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, we would not expect that the resulting weights will always make much sense. To make the DEA analysis more reasonable, there should be some boundaries (technically called constraints) to limit the relative weight or importance of various inputs, and of various outputs.

In the DEA literature, Charnes et al. (1989), Dyson and Thanassoulis (1988), and Thompson et al. (1990) applied various scheme 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 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 will effectively limit the movement of the weights to 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.

To derive the constraints on the ratios between weights, we use the data themselves and published studies about the relative cost of providing different services We use two types of constraints: the four-fold range and two-fold range. In the four-fold range, the ratio is allowed to vary between one quarter of the observed value and four times the average ratio found in published studies and the data themselves. In the two-fold range, the ratio is allowed to vary by a factor of 4 (1/2 to 2). The two-fold range gives a tighter boundary in the weight selection and is expected to identify a smaller number of efficient units than the four-fold range. Details are given in previous reports (Shim & Kantor, 1998a, 1998b) .

4. RESULTS

4. 1 Efficiency scores without imposing constraints

The libraries are compared to their peer libraries operating under similar conditions or constraints: libraries at the publicly funded universities (n=65) are compared against their publicly funded peers, and libraries at the privately funded universities (n=30) compete among themselves. Without imposing any constraints on the weights, we obtained the results summarized in Table 1.

Table 1 shows that most of the libraries are deemed to be efficient. For example, in 1995, among libraries at the publicly funded universities, more than 72% (=47/65) of the libraries in the group are evaluated by DEA as efficient. In the private group, in 1996, only one library out of 30 was evaluated inefficient. The average efficiency scores in all categories are extremely high, all in the upper .90’s. The fact that many libraries are evaluated efficient with high average efficiency scores does not indicate that libraries included in the analysis are efficient. It is simply the result of adopting a particular model of DEA applied to a particular set of DMUs. Therefore, we can not say that academic research libraries are more efficient than other organizations such as banks and hospitals that are evaluated in other studies.

From an analyst’s point of view, the results seem quite disturbing since the analysis does not have enough discriminating power. There is very little to learn if most libraries are efficient.

Table 1
Distribution of Efficiency Scores Without Constraints

Efficiency Score	Public (n=65)		Private (n=30)	
	1995	1996	1995	1996
	Number of Libraries			
1	47	49	27	29
.91-.99	5	7	2	0
.81-.90	9	4	0	0
.71-80	3	3	0	0
.61-.70	0	1	0	1
< .60	1	1	1	0
Average	0.96	0.96	0.98	0.99
No. of libraries evaluated inefficient	18	16	3	1

4.2 The effect of imposing constraints

We expect that tightening the range on the constraints and/or adding more constraints will increase the discriminating power of the analysis, and reveal more inefficient libraries. Some of the libraries might have seemed efficient only because there were no constraints and thus unrealistic weights make them look as good as possible. When this freedom is reduced, these libraries should become inefficient. Table 2 summarizes the number of inefficient libraries revealed under each constraint environment. The results from the previous, no-constraint model are appended to facilitate comparison.

As we read the table from left to right, we notice a marked change both in the number of libraries evaluated inefficient (efficiency score < 1) and the average efficiency scores. As the number of inefficient libraries goes up, the average efficiency score goes down. For instance, in 1995, without any constraints, about 28% (=18/65*100) of the libraries in the public group were evaluated inefficient, whereas with the strictest constraint environment (two fold range), 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 1995, the number of inefficient libraries increased from 3 to 11, and the average efficiency score decreased from .98 to .91.

Table 2
Number of Libraries Evaluated Inefficient and Average Efficiency Score under Different Constraints

Year	Group	No Constraint	Constraints	
			Four-fold range (1/4-4)	Two-fold range (1/2-2)
1995	Public	18(0.96)	34(0.90)	43(0.83)
	Private	3(0.98)	7(0.94)	11(0.91)
1996	Public	16(0.96)	33(0.90)	41(0.84)
	Private	1(0.99)	7(0.94)	12(0.89)

Note. Public (n=65), Private (n=30). Numbers in the parentheses are the mean efficiency scores.

We believe that the two-fold range using both input and output ratios seems to give us 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, in the public group, about two thirds of the libraries seem to have some other libraries in the same group to learn from. One the other hand, in the private group, since two-thirds of the libraries are evaluated efficient, only about one third of them will have peers to learn from.

4.3 Consistency of efficiency scores over two year period

A reliable evaluation system gives us stable results over a period of time. Thus, DEA should produce similar scores for the libraries unless there are significant changes in the data or the assigned weights or the configuration of the efficient frontier.

Table 3 summarizes the results for the public group. When there is no constraint, the majority of libraries (44 out of 65) received the same efficiency scores that they received in 1995. There are 5 libraries that were evaluated efficient in 1995, but are evaluated inefficient in the following year. There are fewer libraries who upgraded their efficiency status.

As we move from no constraint to the four-fold range constraint evaluation environment, and to an even stricter two-fold range, the magnitude of changes in terms of the average score change, the maximum change, and the efficiency status, is intensified. Although there are considerably smaller number of libraries that received the same efficiency scores during the two-year period, still the majority of libraries posted less than 5% change regardless of the evaluation environment.

Table 3

Consistency of Efficiency Scores over Two Year Period (95-96) : Public Group

Changes in The efficiency scores	No Ratio	Four-fold range Both Ratios	Two-fold range Both Ratios
Distribution			
No Change	44	25	16
0.01-0.05	8	22	36
0.06-0.10	8	10	4
0.11-0.15	2	4	0
0.16-0.20	2	1	4
0.21-0.25	1	2	2
0.26-0.30	0	0	2
> 0.30	0	1	1
Total	65	65	65
Average Change	0.03	0.04	0.06
Highest Change	0.25	0.36	0.52
Down Status	5	7	8
Up Status	3	6	6

Note. Up Status means that a library was evaluated inefficient in 1995 but evaluated efficient in 1996. Down Status means exactly the opposite.

Table 4

Consistency of Efficiency Scores over Two Year Period (95-96) : Private Group

Changes in The efficiency scores	No Ratio	Four-fold range	Two-fold range
Distribution			
No Change	27	21	15
0.01-0.05	1	4	7
0.06-0.10	2	1	4
0.11-0.15	0	2	0
0.16-0.20	0	0	2
0.21-0.25	0	1	0
0.26-0.30	0	0	0
> 0.30	0	1	2
Total	30	30	30
Average Change	0.01	0.04	0.07
Highest Change	0.09	0.62	0.65
Down Status	1	2	3
Up Status	0	2	4

Note. Up Status means that a library was evaluated inefficient in 1995 but evaluated efficient in 1996. Down Status means exactly the opposite case.

The results of the private group, shown in Table 4, give us a similar pattern of changes in the efficiency scores. Again there was one library which had a major change in the efficiency score and the efficiency status. This particular library was evaluated efficient in 1995 in all three evaluation environments. But in 1996, it has one missing variable: the total circulation transaction. This caused the efficiency score to tumble from 1.00 to .38 in the four-fold range and to .35 in the two-fold range. Interestingly, the library was evaluated efficient in 1996 when there was no constraint imposed on the weights. It was able to look efficient with one missing output by giving relatively big weights to its other output variables. When this freedom is removed, the effect of missing value becomes quite evident.

4.4 Sensitivity analysis

The number of variables as well as the size of the comparison group affect the efficiency scores (Dyson, Thanassoulis & Boussofiane, 1990). Since we have fixed number of libraries in the comparison groups, we can vary the number of variables and observe the effect on the efficiency scores. Based on the observation of results so far, we would expect that the number of libraries evaluated efficient as well as the efficiency scores will go down as we take out variables from a base model where all 15 variables are utilized. Table 5 shows, for the 1995 public group, the results of employing alternative models where one output variable at a time is taken out of the base model.

Table 5
Effects of dropping a variable on the efficiency scores : Output Variables (1995, Public)

Changes in the efficiency scores	Inter Lib. Lending	Inter. Lib. Borrowing	Lib. Instruction	Reference	Circulation
Distribution					
No Change	23	23	24	18	18
0.01-0.05	40	40	41	23	22
0.06-0.10	2	2	0	10	6
0.11-0.15	0	0	0	5	6
0.16-0.20	0	0	0	4	3
0.21-0.25	0	0	0	2	4
0.26-0.30	0	0	0	2	2
> 0.30	0	0	0	1	4
Total	65	65	65	65	65
Mean Efficiency Score	0.83	0.83	0.83	0.86	0.74
Average Change	0.01	0.01	0.00	0.06	0.09
Highest Change	0.11	0.07	0.01	0.36	0.63
Down Status	0	0	0	5	5
Up Status	0	2	0	6	0

Note. The mean efficiency score for the base model is .83. The changes are in absolute terms.

It shows, for example, that when the library instruction (measured by the number of people who participated in it) variable is removed from the base model, hardly anything changes. More than one third of the libraries (24 out of 65) are assigned the same efficiency scores as they were with the variable present in the base model. All the rest of the libraries (41 libraries) posted less than .01 score change, which seems quite inconsequential. There was not a single library that switched its efficiency status, either from inefficient to efficient or vice versa.

Removing either interlibrary lending or interlibrary borrowing also has a very small impact on the efficiency scores. On the other hand, dropping either reference transactions or circulation seems to cause substantial changes.

Table 6 summarizes the impact of the input variables for the public group in 1995. Among three variables that represent library users, only the number of graduate students has some impact, though small, on the efficiency scores. Among the staff variables, professional staff and full-time support staff variables have more impact than does the student assistants variable. The net volumes added variable seems to have more impact than other collection-related input variables. When that variable is removed from the base model, the mean efficiency score decreases from .83 to .79. It is the most important of the input variables, but it falls short when compared to the changes produced by two output variables, reference and circulation.

Table 6
Effects of dropping a variable on the efficiency scores : Input Variables. 1995, Public Group

Changes in the efficiency scores	Total Stud	Total Grad	Fac	Prof. Staff	Supp. Staff	Stud Staff	Vols Held	Net Vols Added	Mono-Graphs	Current Serials
Distribution										
No Change	60	51	59	25	25	25	52	27	23	23
0.01-0.05	4	8	5	37	31	40	9	22	32	35
0.06-0.10	0	2	1	2	7	0	4	6	2	5
0.11-0.15	1	1	0	1	2	0	0	3	7	1
0.16-0.20	0	2	0	0	0	0	0	3	1	1
0.21-0.25	0	1	0	0	0	0	0	2	0	0
0.26-0.30	0	0	0	0	0	0	0	1	0	0
> 0.30	0	0	0	0	0	0	0	1	0	0
Total	65	65	65	65	65	65	65	65	65	65
Mean Efficiency Score	0.83	0.82	0.83	0.82	0.84	0.83	0.83	0.79	0.81	0.84
Average Change	0.00	0.02	0.00	0.02	0.02	0.01	0.01	0.04	0.03	0.02
Highest Change	0.12	0.22	0.07	0.13	0.16	0.05	0.10	0.30	0.19	0.20
Down Status	1	5	0	0	0	0	4	7	2	2
Up Status	0	0	0	1	2	1	0	0	1	3

Note. The mean efficiency score for the base model is .83.

Close examination of the data used in the analysis shows that if removal of a variable has a significant effect on the efficiency of a particular library, then that variable has a high weight, for that library. But the converse is not true, as there are many cases in which a variable has a high weight but, after its removal, the library in question is assigned new weights which maintain essentially the same efficiency score or status.

From the point of view of an economist or analyst, reducing the number of variables in the model is a Good Thing. But from the perspective of a library director, the fall from efficient to inefficient status would be unacceptable if it could be traced to our pruning of the model. For this reason we propose that in practical applications, the full set of variables be retained in the analysis.

4.5 Noise analysis

In addition to sensitivity analysis, we added random noise in the data and observed the resulting changes in the efficiency scores and the efficiency status. We conducted four *Monte Carlo* simulations of noise for each year. In each simulation, every value of every variable was subject to a random distortion, causing it to vary uniformly between 90% and 110% of its true value. Table 7 shows the results for the public group. Overall, the mean efficiency score increased by .03 to .05 points. The number of libraries that changed their efficiency status was from 4 to 12. The results demonstrate that while the effect of data error might be inconsequential on the mean efficiency score, they can have significant impact on the configuration of the efficient frontier. In the private group, there was slightly larger fluctuation in the mean efficiency scores and also in the proportion of the libraries that switched their efficiency status. Given this range of variation we hope that library data are accurate to better than 10%.

4.6 Regression Analysis

We ran regression analyses to determine whether the DEA efficiency scores can be predicted from the variables included in the DEA analysis. Note that many of the variables are highly correlated with each other. So there is a potential problem of multicollinearity. Table 8 shows the results for the 1995 data.

In all four cases, the linear models were found to be statistically significant predictors. But the results may have been influenced by the fact that we have large number (15) of highly correlated predictors. Nonetheless, in most cases, 75% of the variation in the efficiency scores can be accounted for by all 15 variables included in the DEA analysis. Further regression analyses reveal that 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 measure deteriorates quite rapidly.

Other independent (predictor) variables, such as the total operating budget, per faculty or per student expenditure, were also used to predict the efficiency scores. However, no such models were statistically significant.

Table 7
Effects of simulated random noise in the data : Public Group

Changes in the efficiency scores	1995				1996			
	Run 1	Run 2	Run 3	Run 4	Run 1	Run 2	Run 3	Run 4
Distribution								
No Change	19	21	20	21	19	20	22	21
0.01-0.05	33	24	31	32	32	32	31	31
0.06-0.10	11	16	8	7	6	8	5	8
0.11-0.15	0	2	6	3	2	1	1	1
0.16-0.20	1	1	0	1	2	1	1	2
0.21-0.25	1	1	0	1	1	1	2	0
0.26-0.30	0	0	0	0	2	0	1	2
> 0.30	0	0	0	0	1	2	2	0
Total	65	65	65	65	65	65	65	65
Mean Efficiency Score	0.84	0.85	0.84	0.86	0.83	0.85	0.87	0.84
Average Change	0.03	0.04	0.03	0.03	0.05	0.04	0.05	0.04
Highest Change	0.25	0.24	0.14	0.24	0.32	0.45	0.57	0.29
Down Status	3	1	2	1	5	2	2	3
Up Status	5	3	5	4	5	7	10	6

Note. The mean efficiency score for the base model is .83 in 1995 and .84 in 1996.

Table 8
Regressions Predicting the Efficiency Scores from Data Variables : 1995

Group	Evaluation Model from which DV is drawn	R^2	*F*	Statistically Significant Predictors
Public	No Constraint	.53	3.614***	MONO***, NPRFSTF*, FAC*, ILBTOT*, LIBINST*, REFTRANS**, TOTCIRC***
	Most Strict Constraints (Two-fold range)	.78	11.274***	VOLSADN***, PRFSTF*, REFTRANS***,TOTCIRC***
Private	No Constraint	.76	2.889*	MONO*, NPRFSTF*, STUDAST*
	Most Strict Constraints (Two-fold range)	.80	3.807**	STUDAST*, REFTRANS*, TOTCIRC**

Note. * p<.05, ** p<.01, *** p<.001

5. CONCLUSIONS AND DISCUSSIONS

The results shown in this paper demonstrate that the DEA technique can be applied to the evaluation of academic research libraries. When external information, in the form of the constraints on the ratios between weights, is added to the DEA, it seems to increase the discriminating power of the technique quite significantly. DEA also gives us quite consistent results over a two-year period (1995-1996) that we tested. We also observed the effect of individual variable on the efficiency scores to get some idea about an ideal set of variables, which has yet to be determined. The results also show that the technique is fairly robust despite the presence of 10% random noise in the data. And finally a series of regression runs indicate that the efficiency scores are explained only by including all 15 variables in a linear model, and then only approximately.

One crucial point about the DEA is that the efficiency scores can vary wildly depending on the assignment of weights to the variables. Even when sensible constraints are imposed, some of the results do not seem sensible in the eyes of the investigators. Therefore, to be used as a practical tool for evaluation, we must pay careful attention to how the weights are assigned and how they can be interpreted.

One problem encountered during the course of this study is that the present data do not capture (1) ease of access (2) preservation function (3) academic libraries' serials burden. In particular, in light of the fact that libraries are spending increasingly a significant portion of their materials budget to keep up with expensive journals, it is unfortunate that we do not have a single measure that tells something about the level of serials use. However, with the proliferation of electronic journals and digital collections, we will be better equipped to capture this kind of information. DEA seems to have a flexibility and expandability that other traditional measures lack. While it will take some years for the academic library community to come up with measures that capture user activities in the networked, digital environment, the work reported here provides a base line approach that can be continued in the new era.

Finally, the study is based on secondary, statistical data. Thus, the results of this study should be regarded as documentary evidence about the problem of technical efficiencies in the libraries examined. We recognize that fuller assessment of the efficiencies at the academic libraries would require understanding the local context of individual libraries.

REFERENCES

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.

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

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

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

Council for Aid to Education (1997). *Breaking the Social Contract: The Fiscal Crisis in Higher Education* [Online]. Available: http://www.rand.org/publications/CAE/CAE100/index.html

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: http://www.warwick.ac.uk/~bsrlu/.

Golany, B., & Roll, Y. (1989). Application procedure for DEA. *Omega, 17(3),* 237-250.

Kyrillidou, M. (1998). An overview of performance measure in higher education and libraries. *ARL, 197(April),* 3-7.

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

Michalko, J. (1993). Higher education, the production function, and the library. *Journal of Library administration, 19(3/4)*, 11-22.

Seiford, L. M. (1994). A DEA bibliography (1978-1992). In A. Charnes, W. W. Cooper, A. Y. Lewin, and L. M. Seiford (Eds.), *Data Envelopment Analysis: theory, methodology, and application* (pp. 437-469). Norwell, MA: 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 R. H. (Ed.) *Measuring efficiency: An assessment of Data Envelopment Analysis.* (pp. 7-29). San Francisco: Jossey Bass.

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.

Shim, W., & Kantor, P. B. (1998a). *A novel approach to the evaluation of libraries. Final report to the Council on Library and Information Resources (CLIR).* New Brunswick, NJ: Alexandria Project Laboratory, SCILS, Rutgers University.

Shim, W., & Kantor, P. B. (1998b). A novel economic approach to the evaluation of academic libraries. *Proceedings of the Annual Meeting of the American Society for Information Science, 35*, 400-410.