Combining DEA and factor analysis to improve evaluation of academic departments given uncertainty about the output constructs.

SClaudina Vargas and Dennis Bricker[1]

Department of Industrial Engineering, University of Iowa, Iowa City, IA 52242, USA
April 2000

Abstract

This paper combines the CCR output-oriented model of data envelopment analysis (DEA) and Factor Analysis (FA) to evaluate the performance of academic units of a university's graduate programs relative to their counterparts nationally. We propose DEA/FA as a means of increasing the utility of DEA for policy decisions when there is uncertainty about the output constructs relevant to the programs. We discuss the concept that an academic program often maximizes the levels of some constructed outputs (CO), which may not themselves be directly observable. By means of FA, these COs can be deduced from the observable outputs, and can be expressed as a linear combination of observed and random components. Using the COs lessen the caveat of extreme specialization without the requirement for value judgements.

Keywords: Data envelopment analysis (DEA); Factor analysis (FA); University programs; Policy decisions.

Introduction

Data Envelopment Analysis (DEA) is a non-parametric approach to evaluating technical efficiency which roots can be traced back to the work of Farrel (1957) on productive efficiency. The breakthrough, however, came in the research work undertaken by Charnes, Cooper and Rhodes (CCR) in 1978. Since its inception, DEA has received widespread acceptability particularly in its application to public sector operations, such as education and health care (Thenassoulis and Dunstan 1994, Tomkins and Green 1988, Sinauny-Stern et al. 1994, Rhode and Southwick 1993, Cooper et al. 1996 and 2000)., where a hierarchy of policy objectives, vaguely defined functional form of the inputs-outputs relationships, absence of market prices for the inputs-outputs and multiplicity of environmental factors are integral to the system structures

In this paper we specifically employ the CCR output-oriented model with a single constant input (CCR-OO-CI), although the equivalent BCC model without inputs (Lovell et al. 1999) could also be used. In the educational sector, for example, the use of the CCR-OO-CI model is appropriate when comparing the performance of a university's internal departments to their external peers with respect to a number of performance

[1] E-mail: scvargas@icaen.uiowa.edu & dbricker@icaen.uiowa.edu

measures. Since it is normally difficult or even impossible to obtain from external academic institutions reliable measures of input with a fine degree of granularity (by department), an acceptable approach in this case could be to assume a single input which is identical across the DMUs.

The CCR-OO-CI model produces meaningful radial efficiency measures (Lovell et al. 1999), but it presents some important drawbacks that can lead to wrong policy decisions. We show that this model generally tends to crowd the frontier with specialists who may hide poor performance by ignoring input-output mixes showing less advantages. This problem is less pervasive when variables are continuous rather than restricted to discrete levels (i.e. L = 1,2,...p). Additionally, the model places strict restrictions on the maximum size of the set of output variables to be included. When the variables are restricted to values in a closed discrete interval, a sizable proportion (over 60%) of DMUs are deemed efficient, even when the number of output variables relative to the number of DMUs (the sample size) is small (i.e. < 1/7). The number of efficient DMUs increases (to 90% or more) as the number of variables relative to the sample size increases (i.e., < 1/4). When variables may assume continuous values, however, the proportion of efficient DMUs decreases significantly (to less than 50%) when the number of variables relative to the sample size is less than 1/7. In both cases, however, a large proportion of extreme specialist DMUs define the frontier.

The obvious alternatives to deal with these drawbacks are: (a) use of continuous variables, (b) a reduction of the number of output variables, (c) constraints on the feasible production possibility or on the weights to discourage specialization and to ensure compliance with higher order policies or controls. Examining these alternatives we immediately notice some difficulties. *First:* alternatives (a) and (b) are not always possible. In reality, output variables are not always continuous and selection of a very reduced number of output variables relative to the sample size can lead to model misspecification by excluding relevant variables while including irrelevant variables. In addition, departmental policies or management requirements can impose some restrictions. *Second*: there are limitations on the imposition of restrictions, such as the necessity of value judgements to account for hierarchy in policy objectives. *Third:* we argue that academic programs often maximize the levels of some constructed outputs

(CO), which may not themselves be directly or only vaguely defined. Uncertainty about the form of these COs constrains the analyst's choice of output variables. Given this uncertainty, the inclusion of a larger set of output variables may be desirable in order to reduce the risk of excluding relevant variables, or when a goal is to evaluate the relative efficiency of the DMUs in term of their performance of some output constructs.

As an alternative, we propose combining DEA and factor analysis (DEA/FA) as a viable and effective approach for overcoming the limitations of the CCR-OO-CI model and to improve discrimination between the efficient DMUs. FA is a multivariate analytical technique which is widely used in social science and other research. *First*, we use FA both (a) to extract a parsimonious set of variables from the observable output variables which explains a substantial proportion of the variance, and (b) to generate a set of new variables defined in a continuous domain. The extracted output levels are expressed as linear combinations of the original levels plus some random components. *Second*, we claim that this economical set of variables is a good representation of the relevant latent output constructs (COs) being maximized, and that evaluating academic departments with respect to these COs is more congruent with departmental policies. It also renders a more appropriate measure of efficiency. *Third*, we argue that using the CO scores in lieu of the original data effectively constrains the optimal choice of weights (policy) available to the programs or DMUs while still allowing "unrestricted choice" of policy in the DEA model. This form of constraining the weights serves to obtain efficiency measures that reflect higher-order policies or controls, accounting for externality effects and equity considerations. In this sense, DEA/FA can be valuable for controlling extreme specialization, a known drawback of DEA while still maintaining its advantages.

Although DEA differs from long-standing statistical techniques in some important ways, they need not be mutually exclusive. For some time, an extensive body of work has emerged considering alternative ways in which DEA and some statistical techniques can be combined to improve the results or produce complementary solutions to the problem (Rhodes and Southwick, 1993; Arnold et al. 1996, Bordhan 1998, Ueda and Hoshiai 1997, Zhu 1998). The combined use of DEA and multivariate analysis techniques is particularly notable. For instance, Zilla and Friedman (1998) combine DEA and Discriminant Analysis to *develop* a new method, "Discriminant Data Envelopment Analysis of Ratios (DR/DEA)", that allows the determination of the "best common weight" in order to rank all units on the same scale. Ehreth(1994) uses factor analysis as a means for selecting the most appropriate measures for further analysis while Zhu (1998) compares DEA with principal component analysis (PCA) as an alternative or even complementary ranking technique. Ueda and Hoshiai (1997), on the other hand, propose using principal component analysis as a means of weighting inputs and/or outputs and summarizing them parsimoniously rather than selecting them. Our approach of combining DEA and FA is similar to that of Ueda and Hoshiai (1997), but differs from theirs in come important concepts and the purposes for combining both techniques.

The remainder of this paper is organized as follows: first we explain the characteristics of the OO-CCR-CI model. Second we explain our rationale for combining DEA and FA. Third, we present the FA model. Fourth, we discuss the data, following which we discuss the findings. Finally, we present the conclusion.

The output-oriented CCR model with a single constant input: use and limitations

The Output-oriented Charnes-Cooper-Rhodes model with constant input (OO-CCR-CI) produces meaningful results when the objective is to measure the relative efficiency of the decision-making units (DMUs) in the production of a set Y of variable outputs given a vector $\overline{X}$ of constant inputs:

$$\begin{aligned} &\max_{\phi,\lambda} \ \phi && (M1) \\ &s.t.\ \ Y^T \lambda \geq Y_0^T, \\ &\qquad \overline{X}^T \lambda \leq 1, \\ &\qquad \lambda \geq 0, \end{aligned}$$

Note that ϕ, λ are the decision variables to be optimized. Let $P = \{(\overline{X}, Y) \in \Re^{+} / g(\overline{X}) = Y\}$ be the production possibilities set, where $Y = (y_1, y_2, \ldots y_n)$ is an $n\,x1$ vector of outputs that can be technically produced with $\overline{X} = (\overline{x}_1, \overline{x}_2, \ldots \overline{x}_s)$, a $s\,x1$ vector of constant inputs, with the technology g, which does not have to be specified *a priori* in DEA. The form of g determines the type of returns to scale of the system. Since rescaling the vector of constant input to level 1 will not affect the optimal efficiency score obtained by DEA, we can set $\overline{X}^T$ to 1s. Under the assumption of diminishing marginal productivity of the inputs, P is convex. And for the sector or units for which g is specified, P is specific. For any $Y \in P$ the short-run conditions $\left|Y\right| \leq n$ and $Y \leq \max g(\overline{X})$ hold. These conditions limit the size of the set of output variables that could be used for evaluating the DMUs relative efficiency using *M1,* and precludes the assumption of infinity (∞) elasticity of output (production) with respect to the input. Infinite output elasticity implies unbounded production, which is unrealistic even in an educational environment. DEA partitions P into two subsets, namely, the inefficient production set $P(i) = \{(Y^o, \overline{X}) / Y^o = g(\overline{X}) \prec g(\overline{X}) = Y^*\}$ enveloped by the frontier, and the efficient production set $P(e) = \{(Y^*, \overline{X}) / Y^* = g(\overline{X})\}$ that defines the frontier. Where Y^* is an observed output for which $f^* = 1$ for some $l^* \in \Lambda^*$, the optimal policy set, and Y^o is an observed output for which $f^* \prec 1$ for some $l^* \in \Lambda^*$.

In the educational sector, for example, the use of *M1* is appropriate when comparing the performance of a university's internal academic programs (herein the DMUs) with respect to units of similar discipline externally. Since it is normally difficult or even impossible to obtain reliable input measures from external academic institutions with a fine degree of granularity (by department), an acceptable approach in this case could be to assume a single input, which is identical across the units. Using partial performance outputs for the Y (the ratios of the observed performance of the DMU to the average performance of its external peers) the DEA ranking of the DMUs can be compared using *MI*, which then expresses the degree to which DMU *k* has achieved a more favorable competitive position, compared with external peers.(the relative efficiency of DMU *k*).

It can be seen that DMU *k* with a vector Y of outputs such that for some output levels y_{mk} $(m \leq n)$ in the vector, the observed performance is advantageous, the DMU will appear efficient by choosing $l^* \in \Lambda^*$ (l^*_{sp}) that exploits its competitive advantage in the production of y_{mk} and dominates all DMUs showing lower performance. Although l^*_{sp} appears optimal under DEA, not only can it hide low performance but it may also violate higher-order policies. Furthermore, it can lead to wrong policy decisions, such as wrong target setting. Thus, the utility of the results is an important concern when specifying a DEA model.

In general, the likelihood that DMU k is deemed efficient by DEA is conditional upon the effectiveness and efficiency of its production policy (i.e., its competitiveness). The adequacy of the frontier determined by the DEA model, however, still depends on the model specification, such as the set of inputs and outputs included, a choice regarding returns to scale, assumptions regarding higher-order policies or controls, consideration of externalities and value judgements. That is, the adequacy of the DEA result is model-dependent.

Although the type of output (discrete or continuous) should have no effect on the discriminating power of the DEA model, this is not corroborated by the results we have observed using the CCR-OO-CI model first with discrete and then with continuous outputs. The model's ability to discriminate between efficient and inefficient DMUs was observed to be reduced when output variables are discrete. In addition, the results of the model are highly sensitive to the cardinality of Y . The generally accepted criterion to ensure satisfactory discrimination of the DEA analysis is $(s+n) \leq 1/3k$, where $(s+n)$ represents the combined number of inputs and outputs and k is the number of DMUs in the sample. With the CCR-OO-CI model, however, this upper limit on the combined number of input and output levels appears to be substantially smaller. Furthermore, it appears to be dependent upon the type of output variables (discrete or continuous).

Our finding is that when the outputs variables assume discrete levels (i.e. L (*1,2,…p*)) a sizable proportion (over 60%) of DMUs are deemed efficient, even when the proportion of output variables relative to the sample size is small (i.e. < 1/7). This value increases (to 90% or more) as the number of variables relative to the sample size increases. In the case with continuous variables, the proportion of DMUs ranked efficient decreases significantly (< 50%) when the number of variables relative to the sample size is small (<1/7). In both cases, however, virtually all DMUs effectively ignored some or most of the input-output mixes in order to present themselves in the best possible light. Policies of extreme specialization, although technically optimal, are input-output mix sub-optimal, which raise serious concerns about the adequacy of the efficiency frontier for policy development.

Several approaches can be immediately advanced to deal with these apparent drawbacks of the CCR-OO-CI model: (a) the use of continuous variables, (b) reduction of the number of output variables, and (c) the imposition of restrictions on the feasible production possibility or on the weights. There are, however, limitations with these approaches. *First:* in reality, output variables are not always continuous. In some social environment, like education, measuring the outcomes in some discrete scale may even be common.

Second: selection of a very reduced number of output variables relative to the sample size can be risky. Relevant variables can be excluded while irrelevant variables may be included, which leads to model misspecification (Smith, 1997). Furthermore, the inclusion of a larger number of variables relative to the sample size may be desirable in the analysis of educational units when there is uncertainty about the functional form of constructed outputs that are being maximized. We claim that vagueness of the construct is almost pervasive in educational environment because DMUs do not openly reveal their policies.

Third: imposing restrictions, such as placing constraints on the feasible production possibilities or placing constraints on the weights, can present some difficulties. Adoption of restrictions on DEA is prompted by the need to reflect higher-order policies or controls, externality effects and equity consideration. The imposition of constraints, however, requires not only knowledge of these policies and controls and the form of externalities, but also requires careful selection of the constraints in order to avoid rendering the model infeasible or contradicting the purpose of the restrictions. Imposition of restrictions, therefore, implies incorporating prior views or information regarding the assessment of efficiency of the DMUs (Allen et al. 1997, Roll et al. 1991, Wong and Beasley 1990).

As an alternative, we explore combining DEA and factor analysis (FA) to help overcome the limitations of the CCR-OO-CI model.), FA is a multivariate statistical technique. (See an appendix to this paper for a short explanation of FA) We will next further elaborate our rationale for combining DEA and FA (DEA/FA).

Rationale for DEA/FA when evaluating educational units

In a university environment many different output-outputs are generally produced. Often, however, decision-makers are interested in output-constructs (COs) not directly observable. We contend that these COs constitute the critical elements of the departmental policies and it is with respect to their performance in these constructs that the unit's relative efficiency should be compared. For example, an academic program may be interested in comparing itself to its peers with respect to the quality of candidate it produces, or the program selectivity of applicants rather than test scores and GPA.. Even when the program may have knowledge of the specific composition of the COs, including the functional relationships between the COs and the inputs, as well as the relative importance of the input-output mixes, this information is not generally known by the analyst. Furthermore, some departments may have only an ambiguous definition of the COs. In an academic setting this is an unavoidable phenomenon when internal departmental policies are not openly disclosed. Since departments are not necessarily homogenous, these policies are likely to vary across academic departments, which further limits the analyst's ability to construct the correct COs.

The approach generally followed is to use the directly observable outputs to draw conclusions about the effectiveness or overall performance of the academic programs. For example, test scores and GPA and GPA are generally used as proxy measures of 'quality'. Gender, race, and income, are generally used as proxy measures of 'diversity' levels. With unrestricted weights, however, this approach can render an unrealistic frontier because a DMU will seek to exploit any advantages in performance to position itself on the frontier by relying on one or only a few of the input-output mixes. As previously explained, the CCR-OO-CI model specialization seems to be a serious concern. And, given that the departmental policies are heterogeneous, the difficulty of imposing unbiased weight restrictions is much too great.

As we have previously argued, meaningfully selecting a reduced number of observable outputs in order to increase discrimination among efficient DMUs can lead to model misspecification. For example, consider the rule: if GPA plays a relevant role in all departmental policies, then it should be included; however, if only a few departments consider GPA relevant to their decisions, then GPA may not be included. In reality, the analyst typically doesn't know the departmental policies. He/she, therefore, can only hope to draw adequate inferences from experience and prior research findings to reduce the risk associated with the selection of variables. As an alternative, the analyst may employ other methodologies or techniques that could assist him/her in the selection of an appropriate set of variables, or the extraction of the underlying output structures from the observable outputs. One such technique is factor analysis.

DEA/FA implies ranking the DMUs with respect to their overall performance on some implicit constructed output (the extracted factors) rather than on individual outputs. Each constructed output (CO) may be defined by several observable outputs. Consequently, they must be appropriately and substantially interpreted. Although this can pose some difficulties, it also can lead to more meaningful interpretation of the results. For example, suppose that academic capability (AC) of graduate candidates, an CO extracted using FA, is characterized by three observable outputs: (a) EntMaScore, (b) AppMaScore, and (c) AppTotalScore (see Table 8C for explanation). If 80% of the DMUs on average assign higher weight to this CO, then it is appropriate to infer that AC plays a highly important role in the DMU's policy. It is not appropriate, however, to infer such a conclusion if 80% of the departments assign higher weight to 'EntMaScore' and zero or very low weight to the other two outputs in order to maximize their efficiency scores.

In this sense, an important advantage of DEA/FA for evaluating the efficiency of academic departments is that the policy alternatives that can emerge would not only be congruent with the departmental policies, but would also be more effective since they will target input-CO mixes that are objectives of the departmental policies.

Another advantage of DEA/FA is that it offers an effective way to curtail extreme specialization that may be model-driven while still allowing "unconstrained" flexibility in the choice of optimal pricing policy. Additionally, the advantages of using DEA are still maintained. Such advantages include: (a) No *a priori* specification of the technology, (b) The ability to test for returns to scale using DEA, and (c) Determining a reference set for each inefficient DMUs, which can provide an inefficient DMU with important information for target setting and efficiency improvement. In this sense, DEA/FA can be a valuable tool for improving the educational policy development process.

The data

We use two original data sets for 96 academic programs at the University of Iowa (UI). The data was compiled by the UI Graduate College for the 1996 Profile of the UI graduate programs. The first data set (DS1) includes 21 outputs with values specified in a discrete domain L = [1, 2, ...10] and a single constant input. The second data set (DS2) is comprised of 14 partial performance outputs (the ratios of the observed performance of the University of Iowa unit over the averages observed by peer programs nationally) with values defined in a continuous domain (0, ∝), and a single constant input. A special characteristic of DS1 and DS2 is that $\{DS2\} \xrightarrow{f} \{DS1\}$, where f is the relation representing the method employed by the UI Graduate College to determine the decile values for the outputs from the partial performances observed for each output. The items shown in Table 1 are the outputs comprising the data sets. They are explained in Table 8C.

Data set DS1 (decile)		Data set DS2 (ratios)
Indicator Output		Indicator Output
EntGPA	MinorityEnr	EntMaScore
Inquiree	Int'lStudentEnr	EntTotalScore
AppMaScore	FemaleEnr	AppMaScore
AppTotalScore	yrs/deg/PhD	AppTotalScore
EntMaScore	yrs/deg/Mast	EntGPA
EntTotalScore	NRC-decile	Inquiree
Selectivity	USNews-rank	Selectivity
Yield		Yield
IowaFellow		IowaFellow
GradFellow		GradFellow
Awd/Publ		Exter$FTE
Cite/fac		FemaleEnr
NRC%Supp		MinorityEnr
Exter$FTE		Int'lStudentEnr

Table 1. List of outputs in data sets

Computational results using DS1 and DS2

We performed a series of analyses, ranging from an analysis involving all 21 outputs in DS1 to analyses of different subsets of DS1. Inclusion and exclusion of the outputs was done arbitrarily, but a randomized procedure can also be employed. We experimented with several subsets of outputs to evaluate the effect on the efficiency rating of the programs of applications of the CCR-OO-CI model to the various subsets with and without restrictions on the weights. Also to gauge the need, if any, for reducing the number of output variables, and to try to establish consistency between the different DEA results depending on the input-output mixes. Twenty-five DEA tests were performed using with and without weight restrictions. One test was performed using DS2 Summary of twenty DEA applications performed with the data sets (DEA1-DEA19 and DEA-DS2) are shown in Table 2. Complete details for some of the tests are shown in Table 3 and Table 4.

When the CCR-OO DEA model was applied directly to multiple subsets of DS1 satisfying the condition $(s+n) \leq 1/3k$, k = 96 and $s = 1$, a sizable proportion of DMUs became efficient even in cases where $(1+n)$ is significantly less that 1/7 of the sample size (k). For example, in the case where all inputs and outputs in DS1 $(1+n=22)$ are used, the efficiencies of all but four DMUs are 100% (test DEA1, Table 2, 3). Although other applications of the model to subsets of DS1, ranging from 11 to 19 outputs (tests DEA2 through DEA19, Table 2), show less loss in efficiency discrimination among the DMUs, over 60% of the DMUs are still ranked efficient. In this case, the loss of discrimination among the DMUs, can be explained by the characteristic of the data rather than by an excessive number of input-output variables. This claim is supported when comparing the results using both discrete and continuous output variables. For example, when DEA was applied to the data set with fourteen continuous outputs in DS2, thirty-two of the DMUs (33.4%) are deemed efficient (DEA-DS2, Table 2, 4). Whereas, when DEA is applied to a subset of outputs in DS1, including the same 14 outputs of DS2 but with their values specified in decile (a closed discrete interval $y_{ij}\, e\, [1,2,...10]$), 72 of the DMUs (76.1%) are rated 100% efficient (DEA22, Table 4). This is almost double the number of DMUs deemed efficient by DEA when using the continuous variables of DS2. These results are

	Summary of applications of DEA to subsets of DS1(DEA1-DEA22), DS1-F7 (DEA23) and DS2										
DEA test ID	DEA1	DEA2	DEA3	DEA4	DEA5	DEA6	DEA7	DEA8	DEA9	DEA10	DEA11
No. Indicators	21	16	15	15	13	13	13	12	12	12	11
Efficient DMUs	88	68	80	76	75	67	72	70	62	61	61
Average eficiency score	99.85	96.97	98.45	97.88	97.72	96.69	95.59	95.67	94.93	95.44	94.29
Median efficiency score	1	1	1	1	1	1	1	1	1	1	1
Minimum efficiency score	0.9438	0.8	66.26	0.8	66.26	0.7	0.4	0.4	0.5	0.6	0.4
DEA test ID	DEA12	DEA13	DEA14	DEA15	DEA16	DEA17	DEA18	DEA19	DEA20	DEA-DS2	DS2-F7
No. Indicators	13	19	17	13	14	19(lb)	21(lb)	14(r)	7 (F)	14	7(F)
Efficient DMUs	68	81	74	73	78	12	12	73	37	32	17
Average efficiency score	96.69	98.66	98.01	95.83	98.11	78.42	78.57	97.51	94.5	95.96	95.14
Median efficiency score	1	1	1	1	1	80.52	80.18	1	0.98	96.44	95.49
Minimum efficiency score	66.21	0.8	0.4	0.4	66.26	50.06	49.57	78.99	0.73	84.01	81.21

(lb) lower bound proportional restriction on the weights of 0.025
(r) The 14 indicator outputs included in DS2
(F) Extracted factors, data set DS1(DEA23), data set DS2 (DS2-

Table 2. Summary of DEA applications

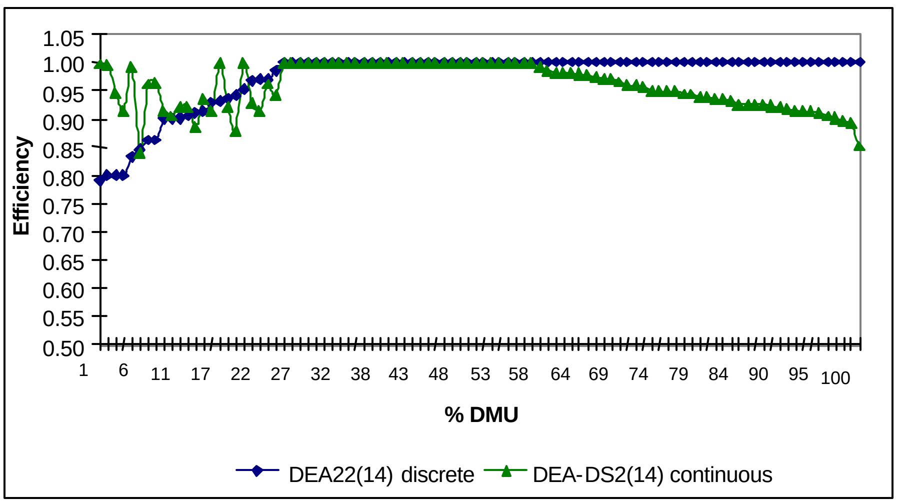

Fig. 1. DEA-DS2 and DEA22 comparison

depicted in Figure 1 where it shows the tendency of applications of the model with discrete output variables to crowd the frontier. That is, with discrete output variables the distribution of the efficiency scores approaches a straight line from the 100% efficiency level as $(s+n)$ approaches $1/3k$. This is clearly a significant drawback.

In all applications of the DEA model with unrestricted weights, independently of the type of the output data, most or all DMUs select optimal weight policies ($\boldsymbol{l}^* \in \Lambda$) that effectively ignore one or more outputs showing less advantages in order to maximize their efficiency scores (f^*). That is, setting $\boldsymbol{n}_i \in \boldsymbol{l}^*$ equal to zero for $y_{mk} \in Y$ showing less advantages. These policies, although technical efficient, violates the input-output mix efficiency condition ($f^* = 1$ & $l^* \succ 0, \forall$), which bounds extreme specialization that hide low performance. They may also violate higher-order policies (i.e., the requirement that all academic programs compete across the range of selected outputs). Thus, specialization can lead to wrong policy decisions, such as wrong target setting. Extreme specialization is a known difficulty of DEA. With the CCR-OO-CI model, however, this problem seems more pervasive in particular across applications with discrete output data, even when $(s+n) << 1/3k$.

In this work, we extend the view that specialization on subsets of outputs is inconsistent with the University of Iowa's objectives of (a) developing graduate programs of national distinction, and (b) assuring high quality and a diverse student body. We argue that the outputs considered in the Profile of the programs were carefully selected from the available data, and consequently no program, when assessing its own efficiency, should be allowed to ignore completely these outputs which were considered by the Graduate College to be relevant. That is, the graduate programs must compete to some extent in *each* of the areas represented by the outputs. We can further argue that the Graduate College intended that all outputs for an academic program must be included in the assessment for that program. The requirement to that all of the outputs be weighted in the DEA analysis such as to ensure compliance with the university's policy, implies restricting their weights to be greater than zero ($l \succ 0$).

Imposing weight restrictions can help overcome some limitations of the CCR-OO DEA model (see for example DEA17 and DEA18, Tables 2, 3), but it poses some new dilemmas as well. This approach, when relevant, incorporate prior views or information regarding the assessment of efficiency of DMUs, as well as consideration to the risk of rendering the model infeasible. These requirements for prior information and knowledge of DMUs policy limit the analyst's choice of the specific process for setting bounds. (For a more thorough explication of weight restrictions the reader is referred to Allen et al., 1997).

An alternative approach to imposing weight restrictions is to use factor analysis, both for reduction of the inputs and outputs set via aggregation, and for specification of discrete data in a continuous domain. This is the approach being follows in this work.

DEA/FA to increase the utility of DEA for policy decision

In addition to using FA for extracting the underlying COs rather than selecting them, FA is also being used in this work for the purpose of transforming the original data set from a discrete domain into a continuous domain.

DMU	DEA1 21-O	DEA18 21-O lb 2.5%	Mean Efficiency	DEA24 7-COs	DMU	DEA1 21-O	DEA18 21-O lb 2.5%	Mean Efficiency	DEA24 7-COs
Art	1	0.7830	1	1	EdPsy/Meas/St	1	0.8943	0.9882	0.9749
Art History	1	0.9317	1	1	Higher Ed	1	0.6317	0.8773	0.9133
Dance	1	0.5745	0.9883	1	InstrDes/Tech	1	0.7040	0.7254	0.8227
Music	1	0.6695	0.9588	1	Science Ed	1	0.5268	0.7176	0.7830
Theatre Arts	1	0.9372	0.9857	1	Secondary Ed	1	0.6544	0.8955	0.8119
Af-Am Wld St	1	0.5105	1	0.9956	Social Foundtns	1	0.6432	0.8533	0.9118
American St	1	0.8914	1	1	Social Studies	1	0.5754	0.7921	0.7925
Asian Civiliz	1	0.6021	0.9765	0.9798	Special Ed	1	0.6709	0.9286	0.7958
Communicat St	1	1	1	1	Biomedical Eng	1	0.8646	1	1
Film/Video	1	0.9563	0.9941	1	Cheml/Bioch	1	0.5826	0.9954	0.9310
Comp Lit	1	0.9563	1	1	Civil/Environ	1	0.8346	0.9805	0.9785
English Lit/Lang	1	1	1	1	Elect/Cmptr	1	0.8816	0.9997	0.8813
Creative Writing	1	1	1	1	Industrial Eng	1	0.7446	0.9787	0.9794
Jour/MassComm	0.9699	0.8444	0.9242	0.9433	Mechanical Eng	1	0.8451	0.9784	0.9720
Lib/Info Sci	1	0.8653	1	0.9552	Accounting	0.9770	0.8488	0.8518	0.8440
Philosophy	0.9924	0.5455	0.9654	0.9320	Finance	1	0.6560	1	0.8779
Religion	1	0.6880	0.9749	1	Mgmt/Orgnztn	1	0.8823	1	0.9015
Classics	1	0.7833	1	1	Mgmt Sciences	0.9875	0.6991	0.9124	0.8566
French	0.9854	0.4947	0.9108	1	Marketing	1	0.7265	1	0.8477
German	1	0.8639	0.9856	1	Anat'y/Cell Bio	1	0.7308	1	0.9386
Russian	0.9438	0.7894	0.9052	0.8350	Biochemistry	1	0.9975	0.9910	0.9262
Spanish	1	0.5138	0.9844	0.9239	Biological Sci	1	1	0.9984	0.9830
Anthropology	1	0.7893	0.9873	1	Genetics	1	1	1	1
Economics	1	0.8887	0.9727	0.9956	Immunology	1	0.7908	1	0.9812
Geography	1	1	1	1	Microbiology	1	1	1	1
History	1	1	1	1	Molecular Biol	1	0.9000	1	1
Linguistics	1	0.6923	0.9807	0.8910	Neuroscience	1	0.8927	1	1
Spt/Hlt/Lei/Ph	1	0.5821	1	0.9808	Pathology	1	0.7615	0.9835	0.9087
Political Sci	1	0.8553	1	1	Pharmacology	1	0.8658	1	1
Psychology	1	1	1	1	Physiol/Biophys	1	1	1	1
Social Work	1	0.9443	0.9845	0.8788	Radiation Biol	1	0.6486	0.9864	0.9920
Sociology	1	0.8087	0.9733	0.8526	Dietetic Interns	1	0.7162	0.9765	0.8912
Thrd Wld Dvpt	1	0.6399	0.9647	0.9983	Hsp/Hlth Adm	1	0.8819	1	1
Urban/Reg Plng	1	0.9242	0.9941	0.9285	PhysicianAssist	1	0.8510	1	1
Chemistry	1	0.9129	0.9994	1	PhysiclTherapy	1	1	1	1
Exercise Sci	1	0.8377	0.9687	0.9640	PrvMed/EnvHl	1	0.7550	0.9904	0.9280
Geology	1	0.5741	1	0.8633	Nursing	1	0.6915	0.9827	0.9089
Physics/Astron	1	0.6076	0.9394	1	Pharmacy	1	0.7737	1	1
SpeechPath/Aud	1	1	1	0.9782	Oral Science	1	0.5407	0.9941	1
Appl. Math Sci	1	0.7327	1	0.9104	Dent Pub Hlth	1	0.8476	0.9961	0.8422
Computer Sci	1	0.7948	0.9504	0.8774	Endodontics	1	0.8578	0.9922	1
Mathematics	1	0.7023	0.9629	0.8968	Oral/MaxilSurg	1	0.7102	0.9898	0.9617
Statistics	1	0.5647	0.9908	0.9840	Operative Dent	1	0.8908	0.9816	0.8857
Actuarial Scienc	1	0.8191	0.9633	0.9192	Orthodontics	1	0.8997	1	1
QualMgmt/Prod	1	0.5481	0.9647	0.8724	Pediatric Dent	1	0.6881	1	0.9935
Counselor Ed	1	0.8713	0.9667	0.9109	Periodontology	1	0.7558	1	1
Erly Chld/Elem	1	0.5861	0.9028	0.8426	Prosthodontics	1	0.8963	1	0.9403
Ed Admin	1	0.5804	0.7412	0.7316	Stomatology	1	0.8305	0.9796	0.9217

Table 3. Comparison of results attained with DEA1, DEA15, mean efficiency and DEA24

Using both DS1 and DS2, two standardized orthogonal data sets of 7 underlying factors (DS1-F7, DS2-F7) were determined by employing FA. These factors are presented in Tables 5 and 6. They represent the underlying constructed outputs (COs) targeted by the DMU's policy, and are expressed as linear combinations of the original data plus a random component (or unique factor). For each DMU a CO is defined as:

$$CO_{fj} = \left(\sum_{i=1}^{n} u_i \hat{y}_{ij} + E \right) + C_f, \qquad f = 1,2,\ldots 7;\ i = 1,2,\ldots n;\ j = 1,2,\ldots 96$$

where f are the indices of the COs of DMU j, u_{ij} is the factor coefficients or factor loadings for the standardized output $\hat{y}_{ij}$, and for each *f*, C_f is a positive constant which is identical for all DMU. This constant has been added to ensure positive values of the COs. These CO scores form the new data sets DS1-F7 and DS2-F7, which are then used in the DEA computations. (For conservation of space, these scores are not included here, but can be furnished upon request). The modified CCR-OO DEA model is presented in $M2$, where CO^T is the transpose *(n ´F)* matrix of COs and the rest are defined in $M1$. When $M2$ was applied directly using the 7 COs in DS1-F7, 37 of the DMUs were deemed 100% efficient, because these COs effectively constrain the feasible region. This ranking of the programs is contrasted in Figure 2 with the ranking obtained using the

$$\max_{f,l} \ \boldsymbol{f} \qquad\qquad (M2)$$
$$s.t.\ \ CO^T \boldsymbol{l} \geq CO_0^T,$$
$$\bar{X}^T \boldsymbol{l} \leq 1,$$
$$\boldsymbol{l} \geq 0,$$

mean efficiency, which is the average efficiency of a DMU computed over all applications of DEA using various subsets of DS1 and without weight restrictions.

$$m_j = Mean\_efficiency = \frac{1}{T} \sum_{t=1}^{T} e_{tj}, \qquad t = 1,2,\ldots 22;\ j = 1,2,\ldots 96$$

DMU	DEA23(14r) Discrete	DEA-DS2 Continuous	DS2-F7 COs	DMU	DEA23(14r) Discrete	DEA-DS2 Continuous	DS2-F7 COs
Art	1	0.9204	0.8685	EdPsy/Meas/St	1	0.8954	0.9620
Art History	1	0.9158	0.9428	Higher Ed	0.9138	0.9358	0.9210
Dance	1	0.9141	0.8121	InstrDes/Tech	0.8000	0.9953	0.9487
Music	0.8466	0.8401	0.9426	Science Ed	0.8000	0.9128	0.8951
Theatre Arts	1	0.9320	0.9137	Secondary Ed	0.9700	0.9137	0.9381
Af-Am Wld St	1	1	0.8966	Social Foundtns	0.9057	0.9205	0.8609
American St	1	1	0.9301	Social Studies	0.9000	0.9149	0.9090
Asian Civiliz	1	0.8528	0.8818	Special Ed	1	0.9258	0.9359
Communicat St	1	1	0.9836	Biomedical Eng	1	0.9442	0.9394
Film/Video	1	1	0.9545	Cheml/Bioch	1	0.9803	0.9573
Comp Lit	1	1	0.9446	Civil/Environ	0.8623	0.9635	1
English Lit/Lang	1	0.9778	0.9860	Elect/Cmptr	1	1	0.9377
Creative Writing	1	1	0.9507	Industrial Eng	0.9857	0.9436	0.9583
Jour/MassComm	0.9668	0.9273	0.9281	Mechanical Eng	0.8333	0.9918	0.9334
Lib/Info Sci	1	0.9653	0.9732	Accounting	0.8623	0.9626	1
Philosophy	0.9362	0.9217	0.9391	Finance	1	0.9493	1
Religion	1	0.9502	1	Mgmt/Orgnztn	1	0.9149	1
Classics	1	1	1	Mgmt Sciences	0.9288	0.9143	0.9534
French	0.9115	0.8846	0.9018	Marketing	1	1	1
German	1	0.9014	0.9532	Anat'y/Cell Bio	1	0.9816	0.9572
Russian	0.9416	0.8778	0.8965	Biochemistry	1	0.9053	1
Spanish	1	1	0.9594	Biological Sci	1	1	0.9849
Anthropology	1	0.9811	0.9637	Genetics	1	0.9918	0.9730
Economics	0.9010	0.9214	0.9548	Immunology	1	0.9864	1
Geography	1	0.9723	0.9551	Microbiology	1	0.9734	0.9681
History	1	0.9589	0.9604	Molecular Biol	1	1	0.9828
Linguistics	0.9714	0.9628	0.9334	Neuroscience	1	1	0.9810
Spt/Hlt/Lei/Ph	1	0.9483	0.8719	Pathology	1	1	0.9426
Political Sci	1	0.9256	0.9582	Pharmacology	1	1	0.9714
Psychology	1	0.9485	1	Physiol/Biophys	1	1	1
Social Work	1	1	0.9491	Radiation Biol	1	1	0.9215
Sociology	0.9000	0.9059	0.9456	Dietetic Interns	1	1	0.9496
Thrd Wld Dvpt	1	0.9814	0.9005	Hsp/Hlth Adm	1	0.9612	0.9563
Urban/Reg Plng	1	0.9151	0.9579	PhysicianAssist	1	1	0.9835
Chemistry	1	0.9254	0.9656	PhysiclTherapy	1	0.9693	0.9622
Exercise Sci	1	1	0.9316	PrvMed/EnvHl	1	1	0.9492
Geology	1	0.8915	0.9245	Nursing	1	0.9377	0.9116
Physics/Astron	1	0.9119	0.9689	Pharmacy	1	1	1
SpeechPath/Aud	1	1	1	Oral Science	1	1	0.9380
Appl. Math Sci	1	1	0.9855	Dent Pub Hlth	1	0.9438	0.8348
Computer Sci	0.9310	1	0.9802	Endodontics	1	0.9251	0.9474
Mathematics	0.7899	0.9974	0.9476	Oral/MaxilSurg	1	1	1
Statistics	1	0.9999	0.9492	Operative Dent	1	0.9567	0.9864
Actuarial Scienc	1	1	1	Orthodontics	1	1	1
QualMgmt/Prod	1	1	0.9715	Pediatric Dent	1	0.9348	0.9564
Counselor Ed	1	0.9234	0.9623	Periodontology	1	1	0.9747
Erly Chld/Elem	0.9516	1	0.9234	Prosthodontics	1	0.9350	0.9497
Ed Admin	0.8000	0.9461	0.8857	Stomatology	1	0.9391	1

Table 4. Ranking of the programs attained with DEA-DS2, DEA23 and DS2-F7

If one considers Figure 2, it is clear that these two methods render consistent results. We interpret m_j as a best approximation of the DMU's efficiency for that data set. Based upon the use of m_j, 38 of the DMUs are rated efficient. This number is very close to 37 but considerable lower than 88, the number of DMUs with 100% efficiency when DEA was applied to all inputs and outputs in DS1 (DEA1). It also represents less than 50% of the number of DMUs deemed efficient in each of the remaining 21 applications of the model to subsets of DS1. Furthermore, 63.2% of the DMUs rated efficient based upon m_j are also efficient with the DEA/FA method. From this we infer that the efficiency frontiers determined by applications without weight restrictions were grossly populated with poor performers and consequently were unrealistic.

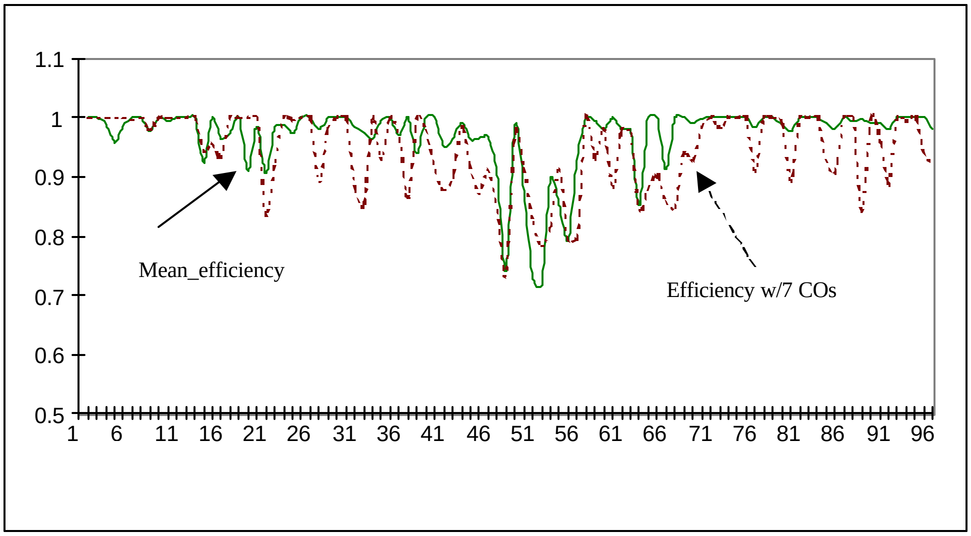

Fig. 2. Mean efficiency DEA24 comparison

Although the mean efficiency method and the DEA/FA method produce similar results, the former has some drawbacks: (1) computational cost increases with the number of runs and the sample size; (2) it is not possible to take advantage of the policy-planning features of DEA such as benchmarking, target setting, and output and input reduction because the slacks and reference sets cannot be computed; and (3) an optimal pricing policy is not easily computable.

When DEA was applied to DS2-F7, seventeen DMUs were deemed efficient, which represents a decrease of about 50% relative to when the model was applied to all inputs and outputs in DS2 (test DS2-DEA). These results are displayed in Table 4 Consider Table 4, where it shows that using the COs produce a significantly different set of efficient DMUs than does using the continuous data set (DS2) directly. Table 4 also shows that only eight programs (Classis, SpeechPath/Aud, Actuarial Science, Marketing, Physiol/Biophys, Pharmacy, Oral/MaxilSurg and Orthodontics) remain efficient across the applications with discrete (DEA22) and continuous (DS2-DEA) data and COs scores

(DS2-F7). Since the average and median efficiency in DS2-DEA and DS2-F are not significantly different (see Table 2), the robustness of the model is not disputable.

Original variables	Candidate Academic Capability F1	Faculty quality productivity F2	University support F3	Information cost/effect F4	Gender mix F5	Visibility F6	Ethnic/ racial mix F7
EntMaScore	0.878						
EntTotalScore	0.856						
AppMaScore	0.796						
AppTotalScore	0.876						
EntGPA							
NRC%Supp		0.790					
NRC-Faculty		0.774					
NRC-Awd/Publ		0.729					
NRC-Cite/fac		0.658					
Inquiree						0.584	
Selectivity				0.551			
Yield			-0.602				
IowaFellow			0.581				
GradFellow			0.758				
Exter$FTE			-0.560				
yrs/deg-PhD			-0.659			0.475	
yrs/deg-Master						0.813	
USNews-rank				-0.826			
FemaleEnr					0.748		
MinorityEnr							0.884
Int'lStudentEnr				0.651			

Table 5. Factors (COs) extracted from DS1 (DS1-F7)

Original variables	Candidate Academic Capability F1	University support F2	Admission boundaries F3	Potential Demand F4	Selectivity (quality) F5	Gender mix F6	Int'l diversity F7
EntMaScore	0.954						
EntTotalScore			0.659	-0.543			
AppMaScore	0.946						
AppTotalScore	0.936						
EntGPA			0.641				
Inquiree				0.877			
Selectivity					0.846		
Yield					0.769		
IowaFellow		0.949					
GradFellow		0.947					
Exter$FTE						-0.727	
FemaleEnr						0.767	
MinorityEnr			-0.798				
Int'lStudentEnr							0.985

Table 6. Factors (COs) extracted from DS2 (DS2-F7)

An examination of the pricing policies for DS2-DEA and DS2-F7 (Table 7A) also reveals some consistencies. The weight of an output reflects the relative importance that a DMUs assigns to that output. Table 7A shows the average, median and maximum weights that are assigned to each output in these applications of DEA. We use these values instead for our assessment. In DS2-DEA the outputs contributing most to the efficiency scores based upon their average, median and maximum weighted values are "AppTotalScore" (39.66%), "AppMaScore" (17.54%), "EntTotalScore" (12.29%), and "GPA" (9.76%). These policies are similar to those in DS2-F7. In this case, the "*Candidate academic capability*" CO, which is determined essentially by "AppMaScore", "AppTotalScore" and "EntMaScore", has an average weight of 30.38%, and is assigned a positive weight by 91% of DMUs. Likewise, the *"internal funding'* CO, which is essentially determined by "IowaFellow" and "GradFellow", has an average weighted value of 44.66% and is assigned a positive weight by most (92%) of the DMUs. The constituent outputs, "IowaFellow" and "GradFellow", however, don't receive high weights in DS2-DEA. *'Selectivity'* formed by "Selectivity" and "Yield" is another relatively important COs in DS2-F7, but only "Selectivity" seems to be somewhat important in DS2-DEA.

In summary, *'Candidate academic capability'*, *'Internal funding'*, and *'Selectivity'* are important COs for most academic departments. Their components (the original outputs) are not all given congruent importance. We believe that this apparent lack of congruency in the policies between the two applications of DEA, rather than weakening our claim about the existence of underlying COs which are targeted by departmental policies, in fact strengthens it. That is because in DS2-DEA the DMUs are able to exploit variations in performance in the individual outputs by specializing in the most advantageous outputs, whereas in DS2-F7 they are unable to apply a similar strategy. Choosing a CO implies including all of its component outputs, some of which may not show a DMU's performance to best advantage. In other words, the reduced dimension of the output space restricts the feasible region and as such it also precludes extreme specialization, which reduces the number of DMUs that, although ratio efficient, are mix-inefficient. Controlling for crowdedness in this manner can improve policy effectiveness by diminishing the risk of imitating a poor role model. Furthermore, the pricing policies observed are congruent with the decision-making process of the academic program because they are comparing each other with respect to COs targeted by their internal policies.

Since the extracted factors do not explain 100% of the variance, it is wise to use the DEA results cautiously. However, if the unexplained variance can be attributed to errors, then we can be more confident with these results. The analyst will have to form some theories about the data and the underlying factors in order to explain the results. Exploratory data analysis (EDA) could prove useful for this purpose.

Conclusions

Although the classification of the DMUs as efficient or inefficient using the CCR output-oriented model with a single constant input may not be meaningless, this model should be employed with care. The efficiency frontier thus determined could be significantly distorted, both in the case of discrete-valued as well as continuous-valued data sets. Regardless of the type of data, the need for discriminatory power of the model imposes strict limitations on the number of input and output variables relative to sample size, which must be significantly below the more generally accepted upper bound of 1/3 of the DMUs. Even with a very small number of variables, extreme specialization is the most likely policy, which raises a serious concern with regard to the adequacy of the efficiency frontier. A crowded frontier likely includes poor performers masquerading as efficient. This poses consequences for the development of improvement policies, which might encourage an inefficient DMU to imitate poor performers posing as role models. These limitations of the CCR output-oriented model can be improved by combining DEA with factor analysis.

DEA/FA effectively constrains *extreme specialization* behavior while still allowing unconstrained flexibility on the choice of optimal weights. The result is a more realistic or adequate frontier. Specialization is also more meaningful because it is based on the output constructs maximized by the DMUs rather than on individual components of the constructs. The observed pricing policies are congruent with the decision-making process of the academic units because they are making comparisons with each other with respect to their relevant output constructs. Employing DEA/FA, the advantages of using DEA are still maintained. For example, it is widely accepted that DEA can be used to test for returns to scale. DEA also determines a reference set for each inefficient DMUs, which can provides the inefficient DMUs with important information for the purpose of target setting and efficiency improvement. In this sense, DEA/FA can be a valuable tool for improving the educational policy development process. These advantages of DEA cannot be enjoyed when using FA or principal component analysis alone to produce an alternative ranking of the DMUs.

.

References

Allen, R., Athanassopoulos, A., Dayson, R.G., Thanassoulis, E., 1997. Weight restrictions and value judgements in Data Envelopment Analysis: Evolution, development and future directions. *Annals of Operations Research* 73, 13-34.

Beasley, J. E., 1990. Comparing University Departments. *OMEGA* **18**(2), 171-183.

Beasley, J.E., 1995. Determining Teaching and Research Efficiencies. *Journal of the Operational Research Society* 46 (4), 441-452.

Bessent, A.M., Bessent, E.W., Charnes, A., Cooper, W.W., Thorogood, N.C., 1983. Evaluation of Educational Program Proposals by Means of DEA. *Educational Administration Quarterly* 19(2), 82-107.

Boussofiance, A., Dayson, R.G., Thanassoulis, E., 1991. Applied Data Envelopment Analysis. *European Journal of Operational Research* 52, 1-15.

Charenes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research* 2, 429-444.

Charnes, A, Cooper, W.A., Lewin, A.Y., Morey, R.C., Rousseau, J.,1985. Sensitivity and stability analysis in DEA. *Annals of Operations Research*. 2, 139-156.

Charnes, A., Cooper, W.W., Lewin, a.Y., Seiford, L.M., 1997. Data Envelopment Analysis: Theory, Methodology and Applications. Academic Publishers (Eds.), Boston. pp. 3-62.

Charnes, A., Cooper, W.W.,Thrall, R.M., 1991. A Structure for Classifying and Characterizing Efficiency and Inefficiency in Data Envelopment Analysis. *Journal of Productivity Analysis* 2, 197-237.

Cooper, W. W., Thompson, R.G., Thrall, R.M., 1996. Extensions and New Developments in DEA. *Annals of Operations Research* 66, 3-45.

Dyson, R.G, Thanassoulis, E., Boussofiane, A., 1990. Data envelopment analysis. In *Operational Research Tutorial Papers.* (L.C. Hendry and R. Eglese, Eds). *The Operational Research Society, UK*, 13-27.

Dyson, R.G., Thanassoulis, E., 1988. Reducing weight flexibility in data envelopment analysis. *Journal of the Operational Research Society* 39(6), 563-576.

Ehreth, J. L., 1994. The Development and Evaluation of Hospital Performance Measures for Policy Analysis. *Medical Care* 32(6), 568-587.

Knox Lovell, C.A., Pastor, J.T., 1999. Radial DEA models without inputs or without outputs. *European Journal of Operational Research,* 118, 46-51.

Rhode E., Southwick L. Jr., 1993. Variations in public and private university performance. Applic. Mgmt. Sci. 7, 145-170.

Roll, Y., Cook, W.D., Golany, B., 1991. Controlling Factor Weights in Data Envelopment Analysis. *IIE Transactions* 23(1), 2-9.

Sinauny-Stern, S., Mehrez, A., Barboy, A., 1994. Academic Departments Efficiency via DEA. *Computers and Operations Research* 21(5), 543-556.

Sinuany-Stern, Z., Friedman, L., 1998. Data envelopment analysis and discriminant analyses of ratios for ranking units. *European Journal of Operational Research* 11, 470-478.

Smith, P., 1997. Model misspecification in Data Envelopment Analysis. *Annals of Operations Research* 73, 233-252.

Tacq, J., 1997, Multivariate analysis Techniques in Social Science Research. Sage Publications, Thousand Oaks.

Thanassoulis, E., Dustan, P., 1994. Guiding schools to improved performance using data envelopment analysis: an illustration with data from a local education authority. *Journal of the Operational Research Society* 45(11), 1247-1262.

Tomkins, C., Green, R., 1988. An Experiment in the Use of Data Envelopment Analysis for Evaluating the Efficiency of UK University Departments of Accounting. *Financial Accountability & Management* 4(2), 147-164.

Ueda, T., Hoshiai,Y.,1997. Application of Principal Component Analysis for Parsimonious Summation of DEA Inputs and/or Outputs. *Journal of the Operational Research Society of Japan* 40(4), 466-478.

University of Iowa Graduate College, 1997. 1996 Profile of Graduate Programs. Iowa City, IA.

Wong, Y-H. B., Beasley, J.E., 1990. Restricting weight flexibility in data envelopment. *Journal of the Operational Research Society* 41(9), 829-835.

Zhu, J., 1998. Data envelopment analysis vs. Principal component analysis: An illustrative study of economic performance of Chines cities. *European Journal of Operational Research* 111, 50-61.

Appendix A

Table 7A

	DEA22(14)				DS2-DEA			
	Average Weight	Median Weight	Assigned Weight = 0	Max Weight	Average Weight	Median Weight	Assigned Weight = 0	Max Weight
EntMaScore	0.0325	0.0000	72%	1.0000	0.0325	0.0000	67%	0.7182
EntTotalScore	0.0665	0.0000	67%	1.0000	0.1229	0.0000	70%	0.9701
AppMaScore	0.0268	0.0000	61%	0.4353	0.1753	0.0300	42%	0.9423
AppTotalScore	0.0382	0.0000	70%	0.4082	0.3966	0.3270	41%	0.9968
EntGPA	0.5278	0.0000	53%	0.0477	0.0976	0.0000	66%	0.9430
Inquiree	0.0826	0.0060	25%	0.8079	0.0173	0.0000	47%	0.7150
Selectivity	0.0291	0.0000	76%	0.7748	0.0157	0.0000	68%	0.0307
Yield	0.1877	0.0159	28%	1.0000	0.0243	0.0050	41%	0.6159
Fellowship	0.0300	0.0000	47%	1.0000	0.0119	0.0000	63%	0.5708
Grand	0.0534	0.0110	36%	0.3767	0.0041	0.0000	89%	0.1320
Exter$FTE	0.0670	0.0000	18%	0.6926	0.0169	0.0000	38%	0.9446
FemaleEnr	0.1088	0.0570	23%	0.7279	0.0439	0.0000	55%	0.8032
MinorityEnr	0.0764	0.0070	44%	0.5929	0.0306	0.0000	51%	0.9597
Int'lStudentEnr	0.1483	0.0440	32%	0.8105	0.0104	0.0008	36%	0.9340
No. efficient DMUs	73				32			
Average Efficiency	0.9751				0.9596			
Min Efficiency	0.7899				0.8401			
	DS2-F7(Vx)				DS2-F6(Vx)			
Factor 1	0.3038	0.2956	9%	0.7471	0.0373	0.0161	16%	0.4385
Factor 2	0.4466	0.0366	8%	0.9411	0.6581	0.6525	6%	0.9577
Factor 3	0.0118	0.0000	72%	0.7455	0.2294	0.2039	31%	0.9540
Factor 4	0.0160	0.0000	54%	0.1385	0.0353	0.0000	55%	0.3892
Factor 5	0.1030	0.0050	42%	0.8095	0.0158	0.0000	91%	0.4839
Factor 6	0.0412	0.0047	50%	0.2862	0.0242	0.0179	17%	0.4116
Factor 7	0.0775	0.0262	42%	1.0000				
No. efficient DMUs	17				11			
Average Efficiency	0.9514				0.9175			
Min Efficiency	0.8121				0.8005			

Appendix B: Factor Analysis using Principal Component Analysis

Factor analysis (FA) is a statistical technique used to identify a relatively small number of factors that can be used to represent relationships among sets of many interrelated variables. For example, variables such as scores in a battery of college entrant tests may be expressed as a linear combination of factors that represents the innate academic abilities of the examinees, such as verbal skills, mathematical aptitude, and comprehension speed. These underlying dimensions or constructs have to be identified. The basic assumption of factor analysis is that underlying dimensions, or factors, can be used to explain complex phenomena. Observed correlation between the observable variables (herein outputs) result from their sharing these underlying factors. For example, correlations between GRE test scores might be attributed to such shared factors as general intelligence, abstract reasoning skill, and reading comprehension. The goal of factor analysis is to identify the not-directly-observable factors based on a set of observable outputs. In the model,

$$F = YU + E$$

Y is the original (*k x n*) data matrix, albeit in standardized form, U is the (*n x n*) matrix of factor coefficients u_{ij}, which are to be determined, E denotes the matrix of scores of the unique factors (i.e., noise) and F denotes the matrix of scores of the common factors shared by the observable data.

Alternatively, each observable outputs can be expressed in term of some common factors and unique factors, $Y = FA' + E$, where A' is the matrix of coefficients, and the matrices Y and E are as defined above. Two main assumptions are : (1) E_i are mutually uncorrelated, which implies that $E'E/(n-1)$ is a diagonal matrix. (2) Unique and common factors are uncorrelated, which implies that $E'F$ is a null matrix.

FA can be considered a technique of latent structure analysis or recognition of hidden structure in the material by the statistical analysis. The goal is to reduce the dimension of the space and the complexity of the original observable structure. This structure is revealed by clustering patterns of the data shown in the correlation matrix. Two forms of factor analysis are principal component analysis (PCA) and principal factor analysis (PFA). These two have different underlying models, but the method for the calculation of the factor solution is the same, namely the principal axis method or examination of the eigenstructure, according to which the eigenvalue of the first factor extracts a maximum of variance from the variables, the second factor extracts a maximum from the remaining variance and is orthogonal to the first factor, and each successive factor extracts a maximum from the remaining variance and is orthogonal to all other factors. In PCA, $Y = CA'$, the eigenvalues explain proportions of the total variance, whereas in PFA, $Y = FA' + E$, the eigenvalues represent only the estimated common variances.

In PCA the model specifies the components as linear combinations of the observable outputs... Let y_{ij} be the individual indicator measures i, $i = (i,....,n)$, for each DMU$_j$ $j = (i,....,k)$. The $k \times n$ data matrix Y is the matrix of original indicator measures:

$$Y = (y_{1,} y_{2,...} y_n)_{k\ x\ n}$$

with each row represents the n individual outputs (or output-input ratios where the input is a unique constant value) y_{ij} for each DMUs and each column represents a specific indicator for all DMUs. (We should note that with multiple input levels, Y is an $(n \times m)$ matrix where m is the number of input levels.)

The independent latent structures (principal components) or measures to be extracted are linear combinations of $z_1,.....,z_n$, where the z_n are standardized y_n. In matrix form this latent structure is:

$$C_{k\,x\,n} = Z_{k\,x\,n}\, x U_{n\,x\,n} \quad \text{or} \quad C = ZU$$

where C is the matrix of component scores, Z is the matrix of original indicator measures , albeit in standardized measures, and U is the matrix of factor or component coefficients u_{ij} . Then $C_p^j = \sum_{i=1}^{n} u_{pi}\, z_i^j$ represents the components as linear combinations of the standardized indicator measures, where $i = 1,....,n$, $p= i,....,n$, and j is equal to the number of observations for each component. The C_p^j are defined under two restrictive conditions: (1) that they are perpendicular or uncorrelated (orthogonality), and (2) that $\sigma_{C_1}^2 > \sigma_{C_2}^2 >,....,> \sigma_{C\,n}^2$, where $\sigma_{C_p}^2 = \lambda_p$, the eigenvalues , and $p = 1,....,n$. The second condition implies that the first component has to extract as much variance as possible from the original outputs, the second component as much as possible from the remaining variance, and so forth.. The number of components retained may satisfy the condition that the cumulative variance extracted be greater than some θ, $(i.e., 50\% \prec \theta \leq 100\%)$, or the condition that the eigenvalue be greater than 1, $\lambda_p \geq 1$ (Kaiser, 1959). [In addition to the Kaiser criterion the Scree test developed by Cattel (1966) can also be used. This criterion implies plotting the eigenvalues against the factors or components and retaining the number with values greater or equal to the value of the factor at the kink.] The retained components, therefore, imply a reduced *q* dimensional space. The interpretation is that the *p-q* components not retained are redundant. Their marginal contributions to the cumulative variance are small. Additionally, the retained components will have to be given an appropriate substantial interpretation indicating the decision or policy variable that they represent. The resulting matrix of component scores C can be used as the new data set for the DEA analysis.

The procedures for determining the latent structures is as follows:

First: Calculate the sample mean vector and covariance matrix:

$\bar{y} = (\bar{y}_1, ... \bar{y}_n)_{1 x n}$ in which $\bar{y}_n = \frac{1}{k}\sum_{j=1}^{k} y_{nj}$, and $S = \frac{1}{n-1}(Y - \bar{y})^T (Y - \bar{y})$. Next, determine *Z*, the original *(k × n)* data matrix, albeit in standardized measures.

Second: Determine the eigenstructure of $R = \frac{Z'Z}{(n-1)} = UVU'$, the correlation matrix with *(n-1)* degrees of freedom [*R* is a square symmetric matrix. Finding the eigenstructure of *Z* amounts to the same as finding the eigenstructure of *R*], where *U* and *V* are the *(n × n)* matrices of eigenvectors and eigenvalues of *R* , respectively. The matrix *V* is also the matrix of eigenvalues of *Z'Z* divided by *n-1*. The eigenstructure is then determined from the systems of equation $(R - \lambda I)u = 0$, where *I* is the $n \times n$ identity matrix and *u* is a vector (or matrix) of factor score coefficients. Determine first *R* then solve the characteristic equation $\left| R - \lambda I \right| = 0$. The characteristic equation yields the eigenvalues λ_i *(i = 1,....n)* of *R* satisfying the conditions $\lambda_1 \geq \lambda_2 \geq .... \geq \lambda_n$ and $\sum_{i=1}^{n} \lambda_i = n$, which is the sum of variances of the standardized variables z_{ij} *(i = 1,....n)* or the sum of elements on the principal diagonal of *R*. Substituting the λ_i into the equation $(R - \lambda I)u = 0$ yields the eigenvectors or characteristic vectors u_i *(i=1,....n)* which in turn determine the principal components. These vectors are not unique because the system of equations are not linearly independent. The matrix:

$$U = [u_1, u_2, ... u_n]_{(n x n)}$$

that is formed with the eigenvectors is orthogonal and rotates the original *n* dimensional axes around the origin toward the new component axes. [The C_i *(i=1,....n)* components are orthogonal and their means are equal to zero.] The matrix of standardized component scores $C = ZUV^{-1/2}$ can now be obtained, where $V^{-1/2}$ is the *(n × n)* diagonal matrix of eigenvalues, and $UV^{-1/2}$ is the matrix of factor score coefficient or factor structure $\left\lfloor m_{ij} \right\rfloor$. This matrix contains the correlation between components and variables. The square of a correlation coefficient is the proportion of explained variance (the proportion of the variance of the original variables explained by the component). The sum of squared coefficients in a row of $\lfloor m_{ij} \rfloor$ is the *communality*, which is equal to the variance of the standardized original variable (1) minus the proportion of this variance explained by an underlying unique factor (i.e. error or noise factor).

Third: Apply the criterion for reduction of the component or factor space. For example the Kaiser (1959) criterion retains only the components for which the eigenvalues is greater than 1. Another criterion is that of Cattel (1966), which consists of plotting the eigenvalues, and retaining the components determined by the kink (elbow) in the curve. The Bartlett (1950) inferential approach can also be employed.

The components not retained are the redundant components. The measure of redundancy of a component can also be calculated by partitioning UVU' as $u_i \, l_i u'_i$ *(i=1,....n)*, where l_i denotes the eigenvalues and u_i denotes the eigenvectors, and reproducing the matrix R step by step. The process ends when the reproduced $\hat{R}_{(\hat{n}\, x\, \hat{n})}$ matches $R_{(n\, x\, n)}$ reasonably well, where $\hat{n} \leq n$.

Fourth: Determine $\hat{C} = Z\hat{U}\hat{V}^{-1/2}$, the extracted $(k\, x\, \hat{n})$ matrix of standardized component scores, and $\lfloor \hat{m}_{ij} \rfloor$, the extracted $(n\, x\, \hat{n})$ matrix of factor coefficients, where $\hat{n}$ is the number of retained components. The extracted matrix of standardized component scores forms a new data set, which can be substituted for the original data set for comparison or improvement of results. The new data set effectively reduces the dimension of the original data set and the extracted components or factors form an orthogonal basis of a vector space. These extracted components will then have to be given an appropriate interpretation. Because PCA does not always offer the most elucidating solution. to aid interpretation, it is better to rotate the component space to a new simpler space, which results in better interpretation. Varimax, Equimax or Quartimax rotation methods can be used.

Appendix C

Table 8C

Indicator outputs explained	
EntGPA	Grade Point Average of entering graduate students
Inquiree	% in national discipline pool sending GRE scores to UI
AppMaScore	Applicant most applicable GRE score
AppTotalScore	Applicant total GRE score
EntMaScore	Entrant most applicable GRE score
EntTotalScore	Entrant total GRE score
Selectivity	Ratio of applicants to admitted students
Yield	Ratio of admitted to entering students
IowaFellow	Iowa fellowship awarded per entering students
GradFellow	Graduate Fellowship awarded per entering student
Awd/Publ	NRC award publication
Cite/fac	Citation/FTE faculty
NRC%Supp	% faculty with support relative to discipline avg (NRC)
Exter$FTE	External funding awarded per FTE
MinorityEnr	% minority continuous enrollment
Int'lStudentEnr	% int'l students continuous enrollment
FemaleEnr	% female continuous enrollment
yrs/deg/PhD	Years to degree (PhD)
yrs/deg/Mast	Years to degree (Master Degree)
NRC-decile	NRC ranking of faculty
USNews-rank	USNews ranking of the program