Teaching Improvement Model Designed with DEA Method and Management Matrix.

Bernard Montoneri

Abstract

This study uses student evaluation of teachers to design a teaching improvement matrix based on teaching efficiency and performance by combining management matrix and data envelopment analysis. This matrix is designed to formulate suggestions to improve teaching. The research sample consists of 42 classes of freshmen following a course of English in Taiwan. The empirical findings show that proposed model can distribute all the evaluated classes into 4 quadrants depending on their performance and efficiency, identify the importance of each performance indicator, and suggest the improvement direction in different quadrants for all the evaluated classes. A study case of one inefficient class is presented in order to demonstrate the proposed model utility and feasibility.

Keywords: Data envelopment analysis; Teaching improvement matrix; Evaluation of teaching; teaching performance and efficiency.

Introduction

Students' evaluation of teaching (SET) has become, over the years, one the most important measures of teaching quality and performance in universities of Taiwan. An increasing number of higher education institutions (HEIs) use online and anonymous SET to evaluate educators. In the current system in Taiwan, SET typically occurs at the end of the semester and teachers receive the results during the vacation or at the beginning of the following semester. However, the results of SET can provide detailed and applicable suggestions neither to educators who wish to improve their teaching, nor to administrators who want to guide and encourage educators. Therefore, this study addresses the issue of improving classroom teaching from the viewpoint of first line educators and administrators and tries to provide a systematic procedure which combines the well-know concept of matrix in the field of management and data envelopment analysis (DEA), a famous quantitative evaluation method which has already been applied to various fields. Under DEA model, efficiency is relative to the other evaluated units in the same group. Montoneri et al. (2011) showed that some units may obtain a high performance but a low relative efficiency, and vice versa. This ambiguous and disturbing situation needs to be addressed and explained. By consequence, this paper presents a new teaching improvement matrix based on teaching efficiency and performance and develops teaching improvement procedures in order to formulate concrete suggestions. An empirical analysis is conducted to test the applicability of the proposed model.

The research data consist of 42 classes of freshmen following a course of English in a university of Taiwan during the academic year 2004 and 2005. Two inputs (teaching clarity and teaching enthusiasm) and two outputs (students' learning interest and students' satisfaction about grades) are used in this study because they are highly correlated. The four indicators selected here demonstrate the utility and feasibility of the model. Other and more indicators could be selected as long as they show a high degree of correlation. This paper not only conducts a general analysis of the 42 classes, but also presents the case study of one inefficient class, that is, C30.

Literature review

Students' evaluation of teaching

Stronge (1997) and Theall & Franklin (2000) identified SET as the most frequently studied aspect of education. Most HEIs use student ratings as a measure of teaching quality and performance (Cashin, 1999; Zabaleta, 2007). According to Wilson (1997), around 2,000 studies were conducted on SET in the 20^{th} century. Various studies show that SET are used for both formative (for teachers to improve teaching) and summative (for administrators to monitor quality) reasons (Edström, 2008; Arthur, 2009). A large number of studies focused primarily on the validity of SET (Marsh, 1987, 2007; Wachtel, 1998), and even though some scholars identify possible biasing factors (Marsh & Roche, 1987; Huston, 2005; Al-Issa & Sulieman, 2007), most publications agree that SETs are useful to both teachers and administrators. Following Marsh (1987), who states that the central purpose guiding student evaluations of professors should be to provide feedback for the improvement of teaching, we provide, in our study, concrete advices for educators to improve their teaching. Penny (2003) for example considers that SETs research has yet to consider seriously that student ratings are an interaction between the students' own conception of learning and the teaching process. In our study, the selected indicators reflect students' perception of good teaching (for example, students' perception of fairy grading). An increasing number of HEIs use online SET to evaluate teachers. Sorensen and Johnson (2004) edited a publication focusing on how online SET was used to evaluate both on-campus and online classes. Carle (2009) analyzed student perceptions of teaching effectiveness across time for face-to-face and online courses. There is a trend in HEIs to increase the use of the Internet in conducting SET (Achtemeier, Morris, & Finnegan, 2003; Harrington & Reasons, 2005). Many universities tend to prefer online, anonymous and confidential end-of-term course evaluation.

There are obviously many advantages in implementing online ratings: significant cost savings (Bothell & Henderson, 2003), improved turnaround time (Sorenson & Reiner, 2003), and greater convenience for students to respond without using valuable class time (Hmieleski & Champagne, 2000; Sorenson & Johnson, 2004). In order to measure teaching effectiveness, various indicators have been tested and studied. Montoneri et al. (2012) review the literature in detail and list all the indicators used in studies from various countries such as China, Taiwan, the UK, the US, and Spain (notably from page 384 to 387). Some indicators such as communication skills, attitude toward the student, knowledge of the subject, organizational skills, enthusiasm, fairness, flexibility, and encouragement of the student are identified as strongly related to teaching effectiveness (Kim et al., 2000).

Efficiency assessment using data envelopment analysis

Efficiency can be assessed by applying various quantitative evaluation methods such as stochastic frontier analysis (SFA), regression, statistics, ordinary least-squares (OLS), structural equation modeling (SEM), data envelopment analysis (DEA), and multi-level modeling (MLM). Which method is more appropriate depends on the research environment (Ferrier & Knox Lovell, 1990). DEA is an attractive tool because it can measure the performance of educational institutions, departments and courses (Montoneri et al., 2011; Montoneri et al., 2012). DEA model evaluates the relative efficiency of each decision making units (DMUs) within a sample (Samoilenko & Osei-Bryson, 2008) and can receive multiple inputs and produce multiple outputs (Lee, 2009; Lin et al., 2009). There is a large body of literature concerning DEA. Among the most influential studies, Førsund & Sarafoglou (2002) cite Farrell's seminal 1957 paper on concepts of efficiency and the study published by Charnes, Cooper, & Rhodes (1978), which was particularly influential in developing and expanding Farrell (1957). Their model, called the "Charnes-Cooper-Rhodes (CCR) model" or "CCR model", notably includes the function and concept of benchmarking. According to Johnes (2006), the multiple input and output nature of production in HEIs makes DEA rather than SFA the ideal choice of method in this context. Many studies assess the efficiency of universities (Ng & Li, 2000; Abbott & Doucouliagos, 2003; Johnes, 2006) and university departments (Colbert, Levary, & Shaner, 2000; Martin, 2006). Even though many scholars apply statistical analysis or qualitative methods to assess the performance of various courses (Leshem & Bar-Hama, 2008; Richards, 2010), there is a growing trend to use DEA (McGowan & Graham, 2009; Montoneri et al., 2011; Montoneri et al., 2012).

Various applications of management matrix

A number of studies have designed two-dimensional, categorical conceptions of performance quality collectively called management matrices; these matrices have been found to be useful in understanding and improving a variety of job performances. Management matrix was first implemented in the aerospace industry at the end of the 1950s. In the 1960s, Professor Allan Pred criticized normative location theories and introduced the concept of the behavioral matrix in connection with a theory of behavior and location (Pred, 1967). Davis & Lawrence (1977) showed that a matrix organization could include various organizing principles such as function, product, and area. Selby (1987) proposed to use Pred's behavioral matrix as a tool for the analysis of enterprises in rural areas. The time management matrix, popularized by Covey (1989), divides time into four quadrants: quadrant 1-urgent and important; quadrant 2-important but not urgent; quadrant 3-urgent but not important; quadrant 4-neither urgent nor important. Jung (2005) designed a matrix divided into four quadrants focusing on information and communication technology (ICT). Taylor et al. (2004) analyzed why so few non-credentialed teachers remained in teaching in the Los Angeles Unified School District (half of the new teachers leave after their first year). They proposed a matrix of teaching practice classification based on works by notably Coloroso (1994) and Edwards (2000).

Methodology

The study extends the concept of management matrix to construct a teaching improvement matrix model. This proposed matrix divides classes into quadrants according to their performance indicators' room for improvement. Applicable steps are developed to test model's feasibility and utility in order to formulate concrete suggestions for both administrators and educators.

Teaching improvement matrix model

The average values of teaching performance and teaching efficiency can segment the matrix into four quadrants, named I, II, III, and IV.

- Teaching performance: average value of classes' performance indicators selected from students' ratings to teachers at the end of each semester (in y-axis).
- The teaching efficiency: relative efficiency value calculated by applying DEA model and by using the above-mentioned selected performance indicators as input and output indicators (in x-axis).

Classes are located in quadrant I if their teaching performance and their teaching efficiency are both superior to the average values; on the contrary, the classes having both values inferior to the average are in quadrant III. Classes are in quadrant II if their teaching performance is superior to the average value and their teaching efficiency is inferior to the average value. Classes with a teaching performance inferior and a teaching efficiency superior to the average value are in quadrant IV.

Analysis of classes' improvement direction

Once classes have been located in different quadrants, the improvement direction for each class is explored. We take into consideration indicators' contribution in calculating the relative efficiency and their room for improvement in order to identify the importance of each performance indicator and to suggest the improvement direction. Classes' relative efficiency can be increased by minimizing inputs' value or by maximizing outputs' value. An output orientation evaluates the maximum output performance needed under the current input resources, while an input orientation evaluates the minimum input effort needed to maintain the current output performance. However, Montoneri et al. (2011) indicate that minimizing input effort in order to obtain an efficiency value equal to one can mislead educators, because input orientation means to obtain a higher relative efficiency by reducing teaching efforts. This will probably discourage hard-working teachers from making tremendous efforts to improve their teaching skills. Therefore, we choose an output oriented analysis and we only discuss how to increase the performance of output indicators. Accordingly, classes' relative teaching efficiency can be enhanced by additional improvement effort in output indicators. We define for each class the additional effort needed for any output indicator, O_i, as the ratio of the importance of O_i's improvement needed in calculating relative efficiency to the importance of all the outputs' improvement needed. It is expressed as follows (1):

$$\text{DMU's additional effort needed for } O_i\ (\%) = \frac{O_i\text{'s contribution in efficiency} \times O_i\text{'s room for improvement} \times 100}{\sum_{j=1}^{\text{number of outputs}} O_j\text{'s contribution in efficiency} \times O_j\text{'s room for improvement}}$$

In equation (1), O_i represents selected output indicators; i varies from one to the number of outputs. If indicator O_i's contribution is 100%, it means that the relative efficiency value of this class is totally owing to this indicator. For example, an empirical study chooses only two outputs, such as Ouput1 and Ouput2, to evaluate classes' relative efficiency. Assuming the outputs' contribution in calculating efficiency value are 0% and 100%, and the outputs' rooms for improvement are 3.4% and 1.0%, respectively.

Then, the additional improvement effort needed in

$\text{Ouput1} = \frac{0 \times 3.4 \times 100}{0 \times 3.4 + 100 \times 1.0} = 0\%$ and in $\text{Ouput2} = \frac{100 \times 1.0 \times 100}{0 \times 3.4 + 100 \times 1.0} = 100\%$.

It means that this class only needs to improve the indicator Ouput2 and can neglect Ouput1 in order to increase its efficiency value in the short term. Similarly for all the classes, the equation

(1) suggests that they should concentrate or accentuate more improvement effort on the output indicator needing highest improvement effort value. According to this rule, we are able to segment all the classes into several categories, such as: "100% effort on Ouput1", "Effort on Ouput1> Effort on Ouput2", "Effort on Ouput2> Effort on Ouput1" and "100% effort on Ouput2". This segmentation aims at providing classes with concrete information about the additional effort in what direction they need to concentrate on in order to effectively improve their efficiency and performance.

Construction steps for formulating teaching improvement suggestions

This section presents the detailed steps to apply, from micro angle, the proposed matrix to construct improvement suggestions. This analysis concerns a specific study for an individual inefficient class. It can help teachers to know in what quadrant they and other classes are located and how much effort they should make to improve their efficiency and their performance. This phase consists of two stages: calculating the relative efficiency of each class and applying the proposed teaching improvement matrix. They are described as follows:

Stage one: Relative efficiency calculation

This stage gathers the results of relative efficiency calculated by DEA approach for a specific class in order to support the formulation of improvement suggestions.

Step 1. Calculating all the classes' relative teaching efficiency by applying DEA method in order to identity the inefficient and efficient ones.

Step 2. Finding out each inefficient class benchmark reference classes in order to define its role models.

Step 3. Listing each reference class contributions to the inputs'/outputs' optimal values in order to suggest a better choice of role models' output or input indicators.

Step 4. Ranking reference classes' contributions for each indicator in order to know their impact order, since the highest contribution does not always come from the same efficient class.

Step 5. Listing the rooms for improvement of each input/output indicator in order to provide suggestions to inefficient classes.

Step 6. Listing each input/output indicator's contribution in calculating classes relative efficiency in order to provide some clues in finding indicators' importance.

Stage two: Teaching improvement matrix application

This stage applies the proposed teaching improvement matrix in order to design applicable improvement suggestions.

Step 7. Calculating the average value of all the classes' relative teaching efficiencies and teaching performance (the average value of all the classes' indicators) in order to draw a teaching improvement matrix.

Step 8. Comparing each class relative teaching efficiency and indicators' average value with the average values obtained in Step 7 in order to locate them in the quadrants of the previously defined matrix.

Step 9. Engaging the analysis of indicators' improvement effort in order to identify the importance of each performance indicator and indirectly to indicate the improvement direction.

Step 10. Formulating improvement suggestions for the inefficient classes.

The data source

The study case is a private university established in 1956 in Taiwan. The data comes from the university's online student rating system, which provides student feedback to teachers at the end of each semester. The characteristics of the data source and research object are as follows:

1. 42 classes are selected during the academic years 2004 and 2005. They are the decision making units (classes), that is, the evaluated units, named from D1 to D42. There is an average of 35 students per class and 42 classes. Therefore, the data consists of around 1470 students. The sample is big enough to draw reliable conclusions.
2. Freshmen students in a university of Taiwan are chosen as a research object; Students are all freshmen, so they are 18-19 years old. Earlier data were used to protect undergraduate students' privacy.
3. Because of major modifications in the questionnaires in 2007, this paper uses data prior to this date for the sake of consistency.
4. The English course is a two-credit course (two hours/week).
5. Each teacher teaches only one class, that is, the 42 classes are taught by 42 different teachers.
6. The data are based on questionnaires (10 questions) filled out by the students at the end of each semester for each class. Each question is rated from one (very unsatisfied) to five (very satisfied).
7. All the students are required to fill out the questionnaires online if they want their grades to be validated. So it is assumed they all did it.

Selection of input and output indicators

Two inputs and two outputs are selected for the empirical study based on the focus of this study, that is, to find indicators having a significant impact on students' motivation and satisfaction in taking a course of English language. This paper aims at demonstrating the importance of teaching clarity and teacher enthusiam and their impact on students' learning interest and perception of their scores. A correlation analysis is performed to test the reliability of the selected indicators. All the 10 questions in the questionnaires have been tested, however, only the four questions selected below are very positively correlated (with scores higher than 0.9; please see Table 1 below) and are reliable enough to draw scientific and useful conclusions (the highest correlations for the other indicators is 0.882, between Q7 and Q10; however, it seems odd to correlate teachers' attendance with students' perception about their scores).

The results of Pearson correlation coefficients between input and output indicators are summarized in Table 1. The inputs and outputs are all significantly positively correlated, reaching a statistically significant level of 1%, which is in line with the principle of equal expansion and means that the increase in inputs will result in the increase in outputs. The four indicators abbreviated by I1, I2 and O1, O2 respectively are presented as follows:

Input indicators:

I1. Teaching clarity (Q3: "Teachers explain clearly, make the content is easy to assimilate"): it refers to the degree of assimilation by the students in relation with teachers' professional knowledge and preparation of teaching materials.

I2. Teaching enthusiasm (Q6: "Teachers show enthusiasm for the course taught"): it indicates whether teachers can actively answer students' queries and clear their doubts. It signifies whether teachers can positively respond to students' questions and the maturity of teachers' teaching skills and communication skills.

Output indicators:

O1. Students' learning interest (Q5: "Teachers can increase your interest in this course"): students' interest and motivation are generally proportional to their learning performance.

O2. Students' satisfaction about grades (Q10: "Teachers give a very fair assessment of student achievement"): It does not mean that students are happy to have good grades or upset to have bad grades, but that they consider they have been graded fairly and objectively.

Table 1. Pearson correlation coefficients between input and output indicators. [a]

Inputs / Outputs	I1 (Teaching clarity)	I2 (Teaching enthusiasm)
O1 (Students' learning interest)	0.965***	0.905***
O2 (Students' satisfaction about grades)	0.934***	0.953***

*** Significant levels at 1% and *p* value<0.001. [a] The number of observations is 42.

Empirical study

The empirical study illustrates the feasibility of the proposed teaching improvement matrix model. We first calculate the 42 classes' overall relative efficiency.

Overall relative efficiency

Stage one: Relative efficiency calculation

The selected input and output indicators data is fed into the software Frontier Analyst to calculate relative teaching efficiency values and relevant efficiency factors of the selected classes. The results under CCR model of DEA are listed in Table 2:

The column "Teaching efficiency" ranks classes by descending order. Classes with an efficiency value equal to 1 are efficient and constitute "reference sets" of efficiency benchmark for inefficient classes (classes with efficiency value inferior to one). These efficient classes form efficiency frontier curves; the efficiency value of each class is calculated by the distance between their location and these efficiency frontier curves. Eight classes (C33, C13, C3, C22, C29, C27, C5, and C25) are efficient and represent about 19% of all the classes. They do not need any improvement in the input and output indicators. The average efficiency of all the classes is 0.978.

The column “Reference DMUs” includes only efficient classes. The classes with relative teaching efficiency do not have to refer to other classes; but each inefficient class has its proper reference classes and can emulate their features and take them as role models. As a result, inefficient classes can approach to their efficiency frontier curves and by consequence enhance their relative teaching efficiency.

The column “Reference times” indicates the number of times an efficient class acts as a peer. By finding out the most popular reference class, it helps to identify a benchmark class and to formulate improvement suggestions for inefficient classes. Table 2 shows that C13 is the most popular reference class (24 times). Most of the inefficient classes refer to two or three efficient classes which constitute their efficient frontier curves and become their reference classes. C16 has only one reference class, C5. This means that C16’s efficiency frontier curve is only constituted by C5. Thus, all the efficiency factors concerning C16 are calculated based on C5’s values.

The column “Room for improvement” indicates the additional effort needed to become an efficient class. The calculation of the room for improvement of inefficient classes is based on their reference classes. An increase or decrease of the inputs or outputs may increase classes’ efficiency value. Under the output oriented model, an increase of outputs’ performance under current input resources can enhance the relative efficiency of classes until they become efficient. This explains why the values of inputs’ room for improvement are always zero or negative.

Table 2 shows that the room for improvement in outputs for all the inefficient classes’ varies from 0.3% to 9.2%. Inefficient classes have to pay different effort to O1 and O2 according to each output’s room for improvement. C28 and C37 need to improve O1 more than O2; however, C12, C16, and C20 need to improve O2 more than O1.

The column “Contribution in calculating CCR efficiency” can provide useful information concerning the importance of each input and output indicator in designing improvement suggestions for inefficient classes. Averagely speaking, O1 is the most important factor in determining classes' relative teaching efficiency (63.3%); the next most important factor is I2 (59.0%). However, for the efficient classes, O1 is the most important factor (69.5%), followed by I1 (52.8%).

Table 2. Teaching efficiency and efficiency factors [a] of evaluated classes.

DMU name [b]	Teaching efficiency	Rank	Reference DMUs	Reference times	Quadrant in efficiency and performance matrix	Room for improvement (%)				Contribution in calculating CCR efficiency (%)			
						O1	O2	I1	I2	O1	O2	I1	I2

C33	1.000	1	C33	7	I	0.0	0.0	0.0	0.0	100.0	0.0	65.1	34.9
C13	1.000	1	C13	24	I	0.0	0.0	0.0	0.0	100.0	0.0	60.9	39.1
C3	1.000	1	C3	0	I	0.0	0.0	0.0	0.0	100.0	0.0	28.8	71.2
C22	1.000	1	C22	6	I	0.0	0.0	0.0	0.0	88.4	11.6	58.1	41.9
C29	1.000	1	C29	19	IV	0.0	0.0	0.0	0.0	67.8	32.2	42.8	57.2
C27	1.000	1	C27	1	IV	0.0	0.0	0.0	0.0	0.0	100.0	66.4	33.6
C5	1.000	1	C5	16	IV	0.0	0.0	0.0	0.0	100.0	0.0	100.0	0.0
C25	1.000	1	C25	19	IV	0.0	0.0	0.0	0.0	0.0	100.0	0.0	100.0
C8	0.997	9	C13, C25	0	I	0.3	0.3	-0.3	0.0	45.7	54.3	0.0	100.0
C42	0.995	10	C13, C25, C29	0	I	0.5	0.5	0.0	0.0	48.5	51.5	23.4	76.6
C14	0.995	11	C13, C25, C29	0	I	0.6	0.6	0.0	0.0	48.7	51.3	23.3	76.7
C28	0.991	12	C5, C25	0	IV	4.4	1.0	0.0	0.0	0.0	100.0	40.1	59.9
C15	0.990	13	C5, C22, C33	0	I	1.0	1.0	0.0	0.0	98.3	1.7	65.2	34.8
C34	0.990	14	C13, C25	0	IV	1.0	1.0	-0.1	0.0	45.2	54.8	0.0	100.0
C26	0.988	15	C5, C13, C29	0	I	1.2	1.2	0.0	0.0	68.5	31.5	43.4	56.6
C17	0.987	16	C13, C25, C29	0	I	1.3	1.3	0.0	0.0	48.6	51.4	23.3	76.7
C11	0.987	17	C5, C13, C29	0	I	1.4	1.4	0.0	0.0	68.8	31.2	43.8	56.2
C10	0.985	18	C5, C22, C33	0	I	1.5	1.5	0.0	0.0	98.3	1.7	64.5	35.5
C35	0.983	19	C13, C25	0	IV	1.7	1.7	0.0	0.0	45.9	54.1	0.0	100.0
C31	0.981	20	C13, C25, C29	0	I	2.0	2.0	0.0	0.0	47.9	52.1	23.1	76.9
C19	0.980	21	C13, C25, C29	0	I	2.0	2.0	0.0	0.0	49.1	50.9	23.5	76.5
C36	0.979	22	C13, C25, C29	0	I	2.1	2.1	0.0	0.0	48.5	51.5	23.2	76.8
C41	0.978	23	C5, C33	0	III	2.2	2.2	0.0	0.0	100.0	0.0	67.0	33.0
C18	0.977	24	C5, C13, C29	0	III	2.3	2.3	0.0	0.0	68.5	31.5	43.2	56.8
C37	0.977	25	C5, C27	0	III	4.9	2.4	0.0	0.0	0.0	100.0	66.6	33.4
C21	0.977	26	C13, C25, C29	0	II	2.4	2.4	0.0	0.0	48.4	51.6	23.1	76.9
C4	0.975	27	C5, C22, C33	0	II	2.5	2.5	0.0	0.0	98.3	1.7	64.7	35.3
C39	0.975	28	C13, C25, C29	0	II	2.5	2.5	0.0	0.0	48.7	51.3	23.2	76.8
C12	0.974	29	C5, C33	0	II	2.6	4.8	0.0	0.0	100.0	0.0	66.4	33.6
C7	0.966	30	C13, C25, C29	0	III	3.6	3.6	0.0	0.0	48.1	51.9	22.7	77.3
C23	0.965	31	C13, C25, C29	0	II	3.7	3.7	0.0	0.0	48.2	51.8	22.9	77.1
C24	0.963	32	C13, C25, C29	0	III	3.9	3.9	0.0	0.0	47.3	52.7	22.9	77.1
C40	0.962	33	C13, C25, C29	0	II	3.9	3.9	0.0	0.0	46.9	53.1	23.2	76.8

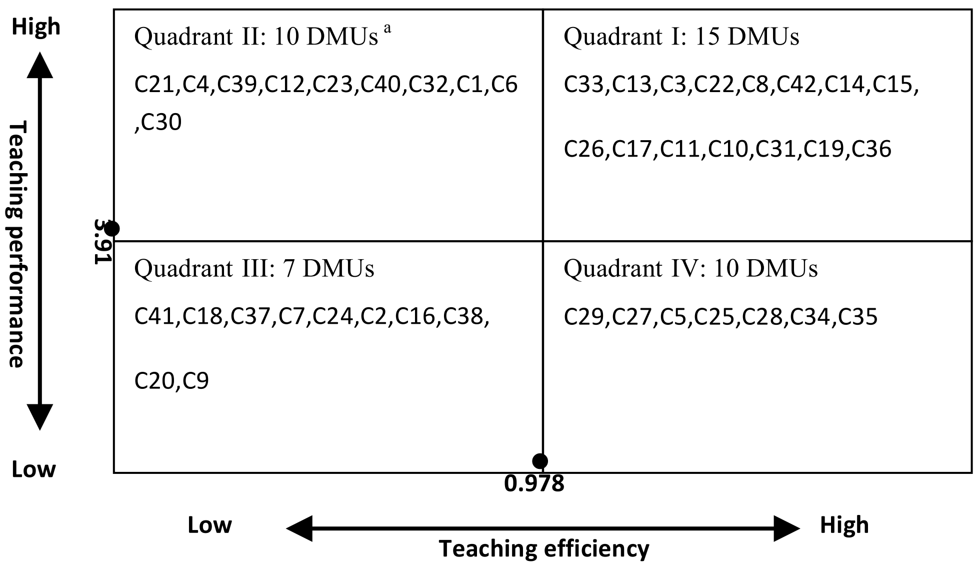
Average of all the DMUs	0.978					2.4	2.4	0.0	-0.1		63.3	36.7	41.0	59.0
Average of the efficient DMUs	1.000					0.0	0.0	0.0	0.0		69.5	30.5	52.8	47.2
Average of the inefficient DMUs	0.973					3.0	3.0	0.0	-0.1		61.9	38.1	38.3	61.7

Notes. [a] I1 indicates teacher's teaching clarity; I2 indicates teacher's teaching enthusiasm; O1 indicates students' learning interest; O2 indicates students' satisfaction about grades.

[b] DMU denotes the evaluated class. The number of observations is 42.

Stage two: Teaching improvement matrix application

The average values of teaching efficiency and teaching performance of all the classes are 0.978 and 3.91; they segment the matrix into four quadrants and divide all the classes into different locations in the matrix, as shown in Figure 1. There are 15 (representing 35.7%), 10 (23.8%), 10 (23.8%), and 7 (16.7%) classes in quadrant I, II, III, and IV, respectively.



Note. [a] DMUs located in each quadrant are in descending order of relative teaching efficiency.

Figure 1. Distribution of classes in the teaching improvement matrix.

In order to figure out the improvement direction for classes in different quadrants, we take into consideration input and output indicators' contribution in calculating the relative efficiency and their room for improvement to identify the importance of each indicator. Since this study uses an output oriented model to engage the analysis, the inefficient classes' relative teaching efficiency can be increased by making an additional effort in output indicators, as defined in the equation (1). The results are listed in Table 3.

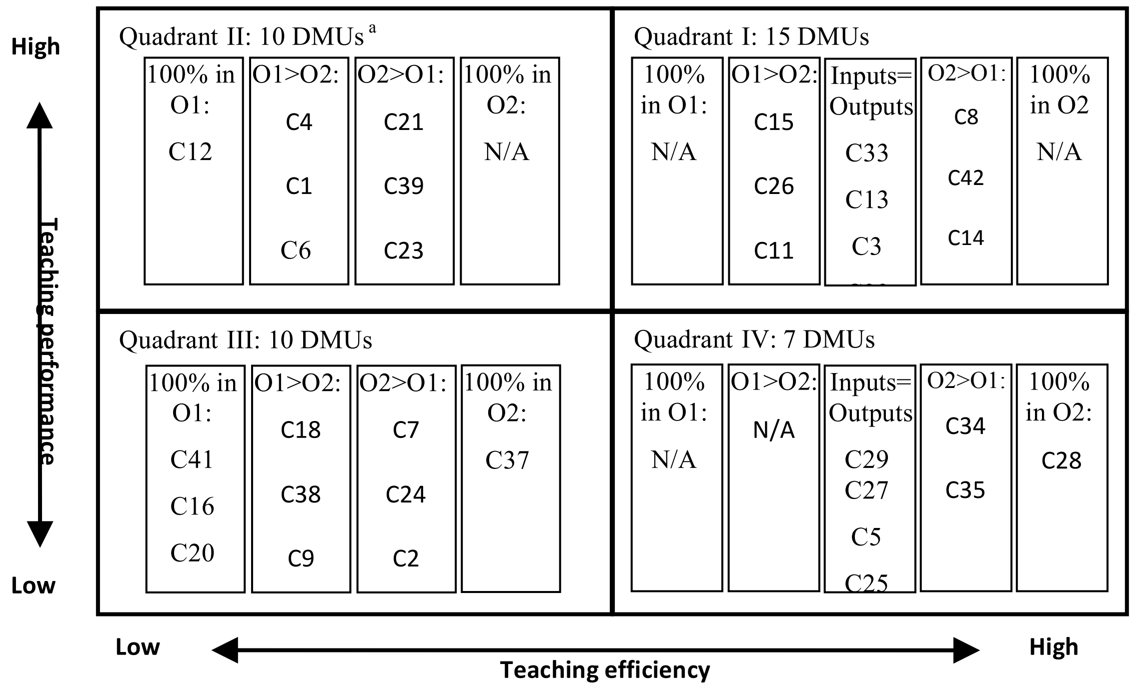
We observe that there are 11, 10, 10 and 3 inefficient classes located in quadrants I, II, III and IV, and they are identified in red, blue, green and violet colors, respectively in Figures 1-2 and in Tables 2-3. Two classes, C28 (in IV) and C37 (in III), are suggested to improve indicator's performance only in O2 (students' satisfaction about grades); four classes (C12 in II and C41, C16, C20 in III) are suggested to improve indicator's performance only in O1 (students' learning interest). Most of other inefficient classes (18 classes) are suggested to improve O2 more than O1. Moreover, for the more efficient classes (those located in quadrants I or IV), none is suggested to improve only O1; for the classes with better performance (those located in quadrants I or IV), none is suggested to improve only O2.

Table 3. Additional effort needed to increase teaching efficiency.

Inefficient classes' name	Located [a] quadrant	additional effort needed in [b] O1 (%)	O2 (%)	Inefficient classes' name	Located [a] quadrant	additional effort needed in [b] O1 (%)	O2 (%)
C8	I	45.7	54.3	C41	III	100.0	0.0
C42	I	48.5	51.5	C18	III	68.5	31.5
C14	I	48.7	51.3	C37	III	0.0	100.0
C28	IV	0.0	100.0	C21	II	48.4	51.6
C15	I	98.3	1.7	C4	II	98.3	1.7
C34	IV	45.2	54.8	C39	II	48.7	51.3
C26	I	68.5	31.5	C12	II	100.0	0.0
C17	I	48.6	51.4	C7	III	48.1	51.9
C11	I	68.8	31.2	C23	II	48.2	51.8
C10	I	98.3	1.7	C24	III	47.3	52.7
C35	IV	45.9	54.1	C40	II	46.9	53.1
C31	I	47.9	52.1	C2	III	48.3	51.7
C19	I	49.1	50.9	C16	III	100.0	0.0
C36	I	48.5	51.5	C32	II	47.7	52.3
				C38	III	88.3	11.7
				C1	II	98.3	1.7
				C20	III	100.0	0.0
				C9	III	68.3	31.7
				C6	II	88.0	12.0
				C30	II	47.7	52.3
Average of	I & IV	54.4%	45.6%	Average of	II & III	67.1%	33.0%
Average of I, II, III, IV: 61.9% in O1, 38.1% in O2							

Notes. [a] The located quadrant indicates the area where the classes are located in the teaching improvement matrix.

[b] O1 indicates students' learning interest; O2 indicates students' satisfaction about grades.



Note. [a] N/A means that there is no DMU located in this area. O1 indicates students' learning interest; O2 indicates students' satisfaction about grades.

Figure 2. Indicator's improvement priority in four quadrants.

Individual analysis: case of C30

C30 is ranked last in relative teaching efficiency according to DEA model. However, C30 has teaching performance score higher than the average value of all the classes and is located in quadrant II of the matrix. Therefore, C30 is a good example to demonstrate how to formulate improvement suggestions for the classes having good teaching performance but low relative teaching efficiency. The analysis procedure is divided into two stages:

Stage one: Relative efficiency calculation

Step 1. The results of relative teaching efficiency analysis show that only eight classes are efficient (C33, C13, C3, C22, C29, C27, C5 and C25).

Step 2. C30's relative efficiency value is 0.916. C30's reference classes are C13, C25 and C29.

Step 3. C13, C25 and C29's contributions to C30's inputs'/outputs' benchmark values are listed in Table 4.

Step 4. C13, C25 and C29's contributions ranking for each input/output indicator are listed in Table 4 below.

Step 5. The rooms for improvement for C30's input/output indicators are listed in Table 4.

Step 6. C30's input/output indicators' contributions in calculating C30's relative efficiency are listed in Table 4.

Table 4. Efficiency improvement analysis for the inefficient DMU C30.

		Outputs [a]		**Inputs**	
		O1	**O2**	**I1**	**I2**
Stage 1: Relative efficiency calculation					
Reference set's contributions to indicators' benchmark values (%)	C13	28.9(3) [b]	27.0(3)	28.6(3)	27.7(3)
	C25	33.9(2)	35.4(2)	35.2(2)	34.5(2)
	C29	37.2(1)	37.6(1)	36.2(1)	37.8(1)
Room for improvement (%)		9.17	9.17	0	0
Outputs/Inputs contribution (%) in calculating relative efficiency		47.7	52.3	23.0	77.0
Stage 2: Teaching improvement matrix application					
Analysis of indicators' improvement effort (%)		47.7	52.3	-	-

Notes. [a] I1 indicates teacher's teaching clarity; I2 indicates teacher's teaching enthusiasm; O1 indicates students' learning interest; O2 indicates students' satisfaction about grades.

[b] Numbers in the parentheses indicate the contributions ranking for each input/output indicator.

Stage two: Teaching improvement matrix application

Step 7. The average value of all the classes' relative teaching efficiencies and teaching performance are 0.978 and 3.91. The teaching improvement matrix is drawn.

Step 8. C30's relative teaching efficiency (0.916) is inferior to the average value of all the classes (0.978) and teaching performance (4.04) is superior to the average value of all the classes (3.91). Therefore, C30 is located in quadrant II of the matrix.

Step 9. The analysis of indicators' improvement effort shows that C30 needs an additional effort of 47.7% in O1 and 52.3% in O2.

Step 10. Teaching improvement suggestions for C30:

1. C30's two output indicators should be improved equally to 9.17%.

2. C30's two inputs, I1 and I2, can be maintained at the same level. It means that the teacher of C30 does not need to improve teaching clarity and teaching enthusiasm in order to improve his/her global evaluation.

3. All the input and output indicators have contribution in calculating C30's relative efficiency. O1 represents 47.7% and O2 represents 52.3% for the output indicators; I1 represents 23.0% and I2 represents 77.0% for the input indicators. Accordingly, the priority of indicators for C30 is I2>O2>O1>I1. (I1: teaching clarity; I2: teaching enthusiasm; O1: students' learning interest; O2: students' satisfaction about grades.)

4. Taking into account both the room for improvement in inputs and outputs and their contribution in calculating efficiency, the indicators with values not equals to zero at the same time should be improved in a priority in order to increase the class relative efficiency. Concretely speaking, it means that C30 needs to make efforts only on students' learning interest (O1) and students' satisfaction about grades (O2), and concentrate on improving more O2 than O1. It means that the teacher of C30 should, according to students, give the impression that the way the educator grades them is fair and objective. It implies giving feedback to students just after they receive their score (individually, not in public, as most students will feel unconfortable about it, making things even worse). In our experience, students appreciate when teachers give them a feedback and an explanation after the exam, even if they fail, or problably we should say, especially if they have a low score.

5. If C30 hopes to increase its relative efficiency in the short term, it should mainly refer to C29's students' satisfaction about grades (O2) up to 37.6%, to C25's O2 to 35.4% and to C13's O2 to 27.0%; then refer to C29's students' learning interest (O1) up to 37.2%, to C25's O1 to 33.9% and to C13's O1 to 28.9%.

6. If C30 hopes to increase its overall performance in each input and output indicators in the long term, its performance improvement measures can not merely refer to one single efficient class, even though C29 is the major model for C30. C30 is suggested to mainly refer to all of C29's input and output indicators around 37%, then refer to all of C25's indicators around 35%, and refer to all of C13's indicators around 28%. Concretely, one way to improve teaching is to ask other teachers who obtain higher evaluation for advice. It is also recommended to attend classes taught by some colleagues (if they agree) to benefit from their experience. It is also advised to ask students why they appreciate one teacher's class. In our experience for example, some teachers are severe, grade students relatively low and still receive a high score because students feel they were fairly graded. As a result, they don't blame their teacher for the results.

7. C30, currently located in quadrant II of the matrix, might make progress in teaching efficiency and upgrade to quadrant I, through the above-mentioned suggestions.

Conclusion

Contribution

The present paper addresses the issue of improving classroom teaching by using online students' ratings of teachers at the end of each semester for the academic year 2004 and 2005. The three main contributions of this study are:

1. Combining the concept of management matrix and a quantitative evaluation method to build a teaching improvement matrix based on teaching efficiency and teaching performance;
2. Developing teaching improvement procedures in order to formulate concrete suggestions to improve teaching;
3. Conducting an empirical study to demonstrate the proposed model's feasibility and utility.

These contributions may help educational administrators to have an overview of classes' efficiency and to obtain information concerning the number and the proportion of classes from the viewpoint of performance and efficiency. Applying this matrix every year may allow administrators to assess the progression or regression of classes' efficiency and performance trend for each academic year.

Main findings

An empirical study is conducted and provides an overall analysis of all the evaluated classes and an individual analysis in order to construct a teaching improvement matrix. This matrix is drawn according to the classification of classes under the CCR model of DEA. Classes are distributed in the four quadrants depending on their performance and efficiency. Once classes have been located in different quadrants, we take into consideration indicators' contribution in calculating the relative efficiency and their room for improvement in order to identify the importance of each performance indicator and indirectly to indicate the improvement direction in different quadrants for all the efficient and inefficient classes. The results of the overall analysis show that the average values of teaching efficiency and teaching performance of all the classes are 0.978 and 3.91, respectively; they segment the matrix into four quadrants and divide all the classes into different quadrants in the matrix. There are 15, 10, 10, and 7 classes in quadrant I, II, III, and IV, respectively. The results of the individual analysis are based on the case C30; because this class is located in quadrant II, it demonstrates how to formulate improvement suggestions for the classes having good teaching performance but low relative teaching efficiency. C30's relative efficiency value is 0.916; it needs to make efforts only on O1 (students' learning interest) and O2 (students' satisfaction about grades), and concentrate on improving more O2 than O1. Then, we formulate improvement directions and suggestions in the short term and in the long term.

Future directions

In the current system in Taiwan, SET typically occurs at the end of the semester and teachers receive the results during the vacation or at the beginning of the following semester. As a result, no matter whether educators are willing to improve their classroom teaching, they can only apply changes to the students of the next year. Teachers and students often find little use for the evaluation process because there are no real time improvement suggestions during the semester. Since students have become courted customers, universities lacking of financial support and depending on student tuition fees for survival should pay more attention to students' opinions and satisfy their demands.

References

Abbott, M., & Doucouliagos, C. (2003), The efficiency of Australian universities: A data envelopment analysis. *Economics of Education Review, 22*(1), 89–97.

Achtemeier, S., Morris, L., & Finnegan, C. (2003), Considerations for developing evaluations on online courses. *Journal of Asynchronous Learning Networks, 7*(1), 1-13.

Arthur, L. (2009), From performativity to professionalism: lecturer's responses to student feedback. *Teaching in Higher Education*, *14*, 441-454.

Al-Issa, A., & Sulieman, H. (2007), Student evaluations of teaching: perceptions and biasing factors. *Quality Assurance in Education, 15*(3), 302-317.

Bothell, T. W., & Henderson, T. (2003), Do online ratings of instruction make Sense? In D. L. Sorenson & T. D. Johnson (Eds.), Online student ratings of instruction [Special issue]. *New Directions for Teaching and Learning,* 96, 69-80. Hoboken, NJ: Jossey-Bass.

Carle, A. C. (2009), Evaluating college students' evaluations of a professor's teaching effectiveness across time and instruction mode (online vs. face-to-face) using a multilevel growth modeling approach. *Computers & Education*, *53*(2), 429–435.

Cashin, W. (1999), Student Rating of Teaching: Uses and Misuses. In P. Seldin, *Changing Practices in Evaluating Teaching: A Practical Guide to Faculty Performance and Promotion/Tenure Decisions* (pp. 25-44). Boston: Anker Publishing Company, Inc.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978), Measuring the Inefficiency of Decision Making Units. *European Journal of Operational Research, 2*(6), 429–444.

Colbert, A., Levary, R., & Shaner, M. (2000), Determining the Relative Efficiency of MBA Programs using DEA. *European Journal of Operational Research, 125*(3), 656–660.

Coloroso, B. (1994), *Kids Are Worth It: Giving Your Child the Gift of Inner Discipline*. New York: William Morrow.

Covey, S. R. (1989), *The Seven Habits of Highly Effective People*. New York: Free Press.

Davis, S. M., & Lawrence, P. R. (1977), *Matrix*. Reading, MA: Addison-Wesley Publishing Co.

Edström, K. (2008), Doing course evaluation as if learning matters most. *Higher Education Research & Development*, *27*, 95-106.

Edwards, C. (2000), *Effective Classroom Management*. New York: Allyn Bacon.

Farrell, M. J. (1957), The measurement of productive efficiency. *Journal of the Royal Statistical Society Series A. General, 120*(3), 253-282.

Ferrier, G. D., & Lovell, C. A. K. (1990), Measuring cost eficiency in banking: Econometric and Linear Programming Evidence. *Journal of Econometrics, 46*(2), 229-245.

Førsund, F. R., & Sarafoglou, N. (2002), On the Origins of Data Envelopment Analysis. *Journal of Productivity Analysis, 17*(1-2), 23–40.

Harrington, C. F., & Reasons, S. G. (2005), Online student evaluation of teaching for distance education: A perfect match? *The Journal of Educators Online, 2*(1), 1-12.

Hmieleski, K., & Champagne, M. V. (2000), Plugging into course evaluation. The Technology Source. Retrieved February 04, 2014, from http://technologysource.org/article/plugging_in_to_course_evaluation

Huston, T. (2005), Research report: Race and gender bias in student evaluations of teaching. Retrieved February 04, 2014, from Seattle University, Center for Excellence in Teaching and Learning Web site: http://www.seattleu.edu/cetl/cetl_006.htm

Johnes, J. (2006), Measuring teaching efficiency in higher education: an application of data envelopment analysis to economics graduates for UK universities. *European Journal of Operational Research*, *174*(1), 443-456.

Jung, I. (2005), ICT-Pedagogy Integration in Teacher Training: Application Cases Worldwide. *Educational Technology and Society, 8*(2), 94–101.

Kim, C, Damewood, E., & Hodge, N. (2000), Professor Attitude: Its Effect on Teaching Evaluations. *Journal of Management Education, 24*(4), 458-473

Lee, C. C. (2009), Analysis of overall technical efficiency, pure technical efficiency and scale efficiency in the medium-sized audit firms. *Expert Systems with Applications*, *36*(8), 11156-11171.

Leshem, S., & Bar-Hama, R. (2008), Evaluating teaching practice. *ELT Journal, 62*(3), 257-265.

Lin, T. T., Lee, C. C., & Chiu, T. F. (2009). Application of DEA in analyzing a bank's operating performance. *Expert Systems with Applications*, *36*(5), 8883-8891.

McGowan, W. R., & Graham, C. R. (2009), Factors contributing to improved teaching performance. *Innovative Higher Education, 34*(3), 161-171.

Martin, E. (2006), Efficiency and Quality in the Current Higher Education Context in Europe: an application of the data envelopment analysis methodology to performance assessment of departments within the University of Zaragoza. *Quality in Higher Education, 12*(1), 57-79.

Marsh, H. W. (1987), Student's evaluations of university teaching: Research findings, methodological issues, and directions for further research. *International Journal of Educational Research, 11*(3), 253-388.

Marsh, H. W. , & Roche, L. A. (1987), Making Students' Evaluations of Teaching Effectiveness Effective. The Critical Issues of Validity, Bias, and Utility. *American Psychologist, 52*(11), 1187-1197.

Marsh, H. W. (2007), Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319–383). New York: Springer.

Montoneri, B., Lee, C. C., Lin, T. T., &. Huang, S. L. (2011), A Learning Performance Evaluation with Benchmarking Concept for English Writing Courses. *Expert Systems with Applications*, 38(12), 14542-14549.

Montoneri, B., Lin, T. T., Lee, C. C., & Huang, S. L. (2012), Application of data envelopment analysis on the indicators contributing to learning and teaching performance. *Teaching and Teacher Education*, 28(3), 382-395.

Ng, Y. C., & Li, S. K. (2000), Measuring the research performance of Chinese higher education institutions: an application of data envelopment analysis. *Education Economics, 8*(2), 139–56.

Penny, A. R. (2003), Changing the agenda for research into student's views about university teaching: four shortcomings of SRT research. *Teaching in Higher Education, 8*(3), 399-411.

Pred, A. (1967), Behavior and location: Foundations for a geographic and dynamic location theory. Part I. *Lund Studies in Geography*, Series B, 27.

Richards, J. C. (2010), Competence and performance in language teaching. *RELC Journal, 41*(2), 101–122.

Samoilenko, S., & Osei-Bryson, K. M. (2008), Increasing the discriminatory power of DEA in the presence of the sample heterogeneity with cluster analysis and decision trees. *Expert Systems with Applications, 34*(2), 1568-1581.

Selby, J. A. (1987), On the Operationalization of Pred's Behavioral Matrix. *Geografiska Annaler*. Series B, Human Geography, *69*(1), 81–90.

Sorenson, D. L., & Reiner, C. (2003), Charting the uncharted seas of online student ratings of instruction. In D.L. Sorenson & T.D. Johnson (Eds.), Online student ratings of instruction [Special issue]. *New Directions for Teaching and Learning, 96*, 1-24. Hoboken, NJ: Jossey-Bass.

Sorenson, D. L., & Johnson, T. D. (2004), Online student ratings of instruction [Special issue]. *New Directions for Teaching and Learning,* 96. Hoboken, NJ: Jossey-Bass.

Stronge, J. H. (1997), Improving schools through teacher evaluation. In J. H. Stronge, *Evaluating teaching: A guide to current thinking and best practice*, (pp. 1-23). Thousand Oaks, CA: Corwin Press.

Theall, M., & Franklin, J. (2000), Creating responsive student ratings systems to improve evaluation practice. In K. E. Ryan, Evaluating teaching in higher education: A vision for the future [Special issue]. *New Directions for Teaching and Learning, 83* (pp. 95-107). San Francisco: Jossey-Bass.

Taylor, C., Jones, A., Shindler, J., & Cadenas, H. (2004), "Architect and Steward: Shaping a Vision of Learning". Examining the Role of the Principal in the Immersion of New Teachers into Existing Urban School Climates. Charter College of Education, California State University, Los Angeles.

Wachtel, H. K. (1998), Student evaluation of college teaching effectiveness: a brief review. *Assessment and Evaluation in Higher Education, 23*(2), 191-210.

Wilson, V. (1997), Focus groups: a useful qualitative method for educational research? *British Educational Research Journal, 23*(2), 209-224.

Zabaleta, F. (2007), The use and misuse of student evaluation of teaching. *Teaching in Higher Education*, *12*, 55-76.