



Assessing the performance of business colleges in Taiwan using data envelopment analysis and student based value-added performance indicators

Wei-Hsin Kong^a, Tsu-Tan Fu^{b,*}

^a Institute of Economics, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei 115, Taiwan

^b Department of Economics, Soochow University, 56 Kueiyang Street, Section 1, Taipei 100, Taiwan

ARTICLE INFO

Article history:

Received 14 March 2010

Accepted 21 October 2011

Available online 10 November 2011

Keywords:

AHP

AR-DEA

Value added indicator

Higher education performance

ABSTRACT

The purpose of this study is to construct a student-based performance evaluation model for business schools in Taiwan. College graduates' starting wage and their multiple abilities cultivated in school are used as performance indicators. The value-added forms of these indicators are employed to assess the pure impact of school on graduate's performance. To prevent impractical indicator weights, we incorporate job market recruiters' weights from an AHP survey into the assurance region data envelopment analysis (AR-DEA) for empirical analysis. Empirical results show that the public schools on average outperform the private schools. However, there are still some private schools that perform better than the public ones. Besides, in regard to the discriminatory power and the distribution of output weights, the proposed AR-DEA is better than DEA in measuring the performance of the business colleges in Taiwan.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

College and university rankings have been the focus of attention when the performance of colleges is considered. The survey reports presented by the media, such as the US News and World Report of the U.S. and The Times of the U.K. are important references for high school students when applying to enter college. A large number of studies that are concerned with the performance ranking of colleges can be found in the literature. However, most previous studies have evaluated college performance from the perspective of the operators, as in Johnes [1], Coelli et al. [2], and Hashimoto and Cohn [3], who discuss the efficiency of college production; Abbott and Doucouliagos [4], Johnes and Johnes [5], and Cherchye and Abeele [6], who discuss the academic research performance; and Athanassopoulos and Shape [7] and Casu and Thanassoulis [8], who evaluate the efficiency of cost; Kao and Hung [9], who evaluate the efficiency of university departments. As a result, those studies tend to regard performance in terms of teaching and research as an output. In the literature, output items related to teaching include student retention, the graduation rate, enrollments, and so on,

while research-related output items include papers published in academic journals, books, and research funds.

The output items mentioned above are the items with which the school operators and the government are most concerned. However, those items are not necessarily the main concerns of students (the consumers of higher education). Students do not care about whether colleges work efficiently or not, or how good their research is. They care about whether they can get a job easily or not when they graduate from college. They also like to evaluate the schools in terms of the help they can provide in relation to their workplace. Universities are composed of the operators and the students. When evaluating their performance, the needs of each of the two parties should be taken into consideration. However, studies dealing with the evaluation of university performance seem to conduct their research from the perspective of the operators, and seldom do the academic studies consider the universities' performance from the eyes of the students. Neither students nor their parents can really obtain the information with which they are really concerned. For example, when deciding which university to attend, high school students might wonder which university can help them further their abilities so that they will have better rewards in their future career. Thus, our research aims to construct a student-based performance evaluation model for the business schools in Taiwan.

To achieve this aim, the indicators that we use are those that can be representative of the degree of helpfulness in enhancing

* Corresponding author. Tel.: +886 2 23111531; fax: +886 2 23611295.

E-mail addresses: kws@ntu.edu.tw (W.-H. Kong),

tfu@scu.edu.tw (T.-T. Fu).

students' future wages and abilities of all kinds. We also use the starting wage of college graduates as a measure of college graduate's job market performance. These indicators are "value-added indicators", which represent the influence of the colleges only. Influences outside of the colleges (e.g., students' talents) are not included. To insure the starting wage to be a value-added indicator, a hierarchical linear model (HLM) is applied to exclude any possible influences outside the college.

Since a business school within a university is a decision-making unit (DMU) that generates a variety of outputs, the literature makes use of data envelopment analysis (DEA) as the method of evaluation [1,10,11]. The weights used in the DEA model are derived from the data instead of being fixed in advance. Each DMU is assigned a best set of weights that may vary from one DMU to another. Each DMU chooses its own weights. A DMU may achieve efficiency by weighting a single input and a single output, with the remaining inputs and outputs being accorded zero weights. This can happen when the DMU is best at one thing. Doyle and Green [12] described this situation as the "Best engine on the market—pity the car has no wheels". One output with a large weight and the rest with very small weights seem unreasonable when it comes to evaluating a college's performance. Therefore, Thompson et al. [13] developed the "assurance region" (AR in short) approach to avoid an unreasonable distribution of weights. In this research, we conduct a survey to job market recruiters of top 100 business firms in Taiwan to elicit their subjective judgments on the relative importance between outputs. The analytic hierarchy process (AHP) proposed by Saaty [14] was used to quantify their subjective judgments. The weight results derived from this AHP analysis are then used for setting the upper and lower bounds of weight restrictions to be used in the proposed AR-DEA model. Such proposed model could insure a reasonable weight distribution of performance outputs.

The model combining the AHP and AR-DEA (the AHP-AR-DEA model) had been previously applied to some different fields before. For examples, Zhu [15] was the first study using the AHP-AR-DEA to evaluate performance of Nanjing textiles corporation in 1996. Seifert and Zhu [16] in 1998 used the AHP-AR-DEA to investigate excesses and deficits in Chinese industrial productivity for the years 1953–1990. In that paper, the weights of ARs were obtained through expert opinions by the Delphi and AHP approaches. Several applications of the AHP-AR-DEA can also be found for public sector. For examples, in 2003, Takamura and Tone [17] used the AHP-AR-DEA to relocate Japanese government agencies out of Tokyo. Sun [18] assessed the performance of joint maintenance shops in the Taiwanese Army by AHP-AR-non-discretionary DEA in 2004. In 2008, Meng et al. [19] combined AHP, AR, and two-level DEA to evaluate the research performance of the research institutes in the Chinese Academy of Sciences.

In this paper, we apply the AHP-AR-DEA model to evaluate the performance of university business colleges in Taiwan. However, differing from those previous applications, the current study makes a contribution to the literature in the following four aspects. First, we actually carried out a representative survey to business college graduates of Taiwan for eliciting relevant information for our empirical assessment. Second, we processed such survey data to construct the student based value-added performance indicators. Third, we conducted an AHP survey to business recruiters to derive the upper and lower bounds of weight restrictions of the ARs. The last and also the most important contribution of this study is using the proposed AHP-AR-DEA model to assess the performance of business colleges from a new and different perspective. That is, instead of assessing the university performance from the perspective of school operators, we evaluate such performance from the concerns of students.

Empirically, the constructed student-based value-added performance indicators will be used as outputs in the proposed student based AHP-AR-DEA model to assess the performance of business colleges in universities in Taiwan. The results of our research indicate that public schools outperform private schools in general. However, there are still some private schools that perform better than the public ones, indicating that when performances are evaluated from the students' wages and performance, the stereotype that "public schools are better than private ones" will be changed. Finally, we suggest that high school students consider the results of our analysis combined together with traditional enrollment predictions as a reference while choosing which schools to attend.

There are 7 sections in this article. Following this introduction, the evaluation structure is illustrated in Section 2. Section 3 presents the methodology and Section 4 describes the data. The evaluation of the indicators and weight restrictions setting is given in Section 5. The empirical results and concluding remarks are provided in Sections 6 and 7.

2. Framework for measuring college performance

Before constructing a student-based performance evaluation model for universities and colleges, "how they recognize the performance" is an important question that needs to be answered. For instance, students and parents care about the scores the students obtain when the children are at the elementary and junior high school stage. As a result, test scores are usually regarded as the output of teaching when concerned with the evaluation of junior high school operating performance [20–22]. However, this is not the case with college students. For college students, whether or not they can get a promising job after graduating from that school becomes the first concern. In other words, they are more likely to evaluate the schools in terms of the help they can provide in relation to their work place. In elementary and high schools, the test scores at each stage are indicators of the evaluation, but when it comes to colleges, due to the different courses that the students take and the inconsistent standards that the instructors abide by, test scores become incomparable between graduates from different universities. All these prevent us to regard the test scores of college students as performance indicators for universities and colleges.

This study applies two groups of variables as indicators to represent the help that the colleges provide. The first group consists of "vocational performance" related variables, such as the students' wage for their first job, their current wage and the scale of the wage increases, by which the differences in terms of performance in a job can be observed directly. Job performance is less inconsistent among students in Taiwan since they face the same workplace and market with the same demand for and supply of labor after they graduate from college.

The other group of variables consists of the degrees of helpfulness that characterize the colleges in terms of cultivating the students' multiple abilities. These include the "cognitive", "psychomotor", and "affective" domains that are also seen in the discipline of education [23]. The "cognitive domain" is knowledge- or mind-based with three practical instructional levels that include fact, understanding, and application. The "psychomotor domain" is skill-based and the three practical instructional levels are imitation, practice, and habit. The affective domain is based upon behavioral aspects and may be referred to as beliefs and the three levels in this domain are awareness, distinction, and integration. The helpfulness of colleges in developing the three domains is therefore used to construct the indicators in our college performance evaluation.

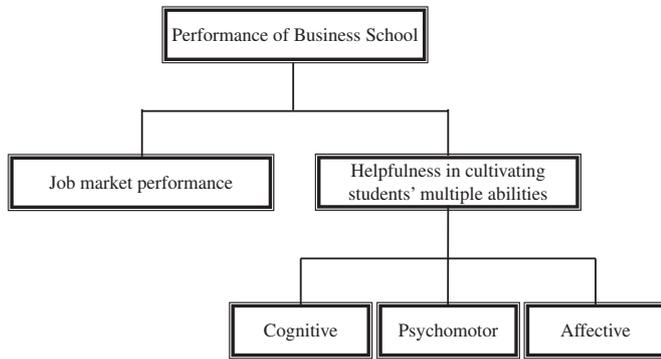


Fig. 1. The evaluation structure of business schools.

The evaluation structure in our analysis is illustrated below (see Fig. 1). The performance of business schools is measured by the “job market performance” and by their “helpfulness in cultivating students’ multiple abilities.” The performance is evaluated from the bottom where the three domains “cognitive”, “psychomotor”, and “affective” form the “helpfulness” indicators in the middle. Finally, the “helpfulness” and “job market performance” variables are further combined to become the overall indicators of business schools. They are also conceptualized as follows:

$$\text{Performance of business school} = \text{students' job market performance} + \text{helpfulness in cultivating students' multiple abilities} \quad (1-1)$$

$$\text{Helpfulness in cultivating students' multiple abilities} = \text{cognitive} + \text{psychomotor} + \text{affective} \quad (1-2)$$

By combining Eqs. (1-2) and (1-1)

$$\text{Performance of business school} = \text{students' job market performance} + \text{cognitive} + \text{psychomotor} + \text{affective} \quad (1-3)$$

3. Methodology: DEA and AR-DEA

The concepts and the application of DEA and AR-DEA are introduced below.

3.1. DEA

The most basic DEA model, the CCR model, was initially proposed by Charnes, Cooper and Rhodes [24]. In the standard DEA formulation, each expression is a ratio of weighted outputs to weighted inputs. However, from the viewpoint of the students, it does not matter how many inputs that college used. So, in this paper, we use a CCR model with a single constant input and M measurements

$$E_i = \max_{w,u} \sum_{j=1}^M w_j Y_{ij} \quad (2)$$

$$\text{s.t.} \quad \sum_{j=1}^M w_j Y_{ij} \leq u X_i$$

$$u X_i = 1, \quad i = 1, \dots, N; \quad j = 1, \dots, M$$

where Y_{ij} is the j -th measurement of the i -th college, X_i is a single constant input and the performance score of the i -th college, E_i , M is the number of measurements, and N is the number of colleges. In this model, the variable w_j can be regarded as the weight to be assigned to the corresponding measurement and u as the weight of X_i . Each college selects the weights which will result in it obtaining the highest possible score.

3.2. AR-DEA

In the real world, reviewers do not always allow colleges to choose their best weights. This is because such best weight choice could result in extreme weight distribution which was pointed out by Doyle and Green [12]. Such a course of action is also not usually employed in the real college performance assessment practices. Therefore, we employ the “assurance region” (AR in short) approach developed by Thompson et al. [13] to avoid such an unreasonable distribution of weights. The AR model, however, could allow weights varying within a region by imposing constraints on the relative magnitudes of the weights for special items. For every pair (Y_{ij1}, Y_{ij2}) of measurement, the ratio W_{j1}/W_{j2} must be bounded by L_{j1j2} and U_{j1j2} . Here, Y_{ij1} and Y_{ij2} are the $j1$ -th and $j2$ -th measurement; W_{j1} and W_{j2} present the weight of Y_{ij1} , Y_{ij2} and the L_{j1j2} and U_{j1j2} present the lower and upper bounds of the ratio. This constraint limits the region of weights to some special area. The linear inequality form is

$$L_{j1,j2} w_{j2} \leq w_{j1} \leq U_{j1,j2} w_{j2} \quad j1 \neq j2 \quad (3)$$

Thus, by adding constraints in Eq. (3) into the CCR model of Eq. (2), we obtain the AR-CCR model.

However, it needs to be asked how one determines the lower and upper bounds. Some studies determine the lower and upper bounds based on the relative prices of inputs (outputs) [13]. When the relative prices of inputs (outputs) are not available, some studies determine them based on expert opinions [15,16,17,18,19,25], a notable example of this being the analytic hierarchy process (AHP) which is also the approach we taken in this research.

4. Data sources

The data used in this study are obtained from two sources. The first source comprises the survey data provided by the “What undergraduates learned during their time at university and their job market performance” survey. This was a questionnaire survey that investigated recent college graduates in business schools of 21 universities as the target sample, with male alumni graduating in 1998 and female alumni graduating in 2000. The basic information (gender, school, department, when they got their job, and working experiences), courses taken (how many courses they took, what kinds of computer skills they learned, whether they took double-majors, etc.), as well as their satisfaction with the school’s helpfulness (whether they were satisfied with the way in which their school cultivated them) were included. The survey to recent college graduates was conducted in 2004.

A total of 10,515 questionnaires were sent out by regular mails during the period of Chinese New Year vocation in February of 2004 to receive a better response rate. A total of 1587 questionnaires were returned in the following 3 months after sending out 2 times of reminders mails. The overall return rate is 15.09%, which is considered to be satisfactory because alumni’s mailing addresses were home addresses of their parents, not their current contact addresses. By screening out those defective questionnaires which were responded not by the alumni himself or herself, or answered with serious missing or incompleteness in requested information, 1250 were useable among the 1587 returned questionnaires. These 1250 observations cover 21 business schools of universities in Taiwan.

By conducting representative tests to 1250 usable samples, we find our sample is no difference to the population in gender distribution, whereas it is different from the population in school distribution. In fact, we find that the percentage of college graduates from public universities in the sample is higher than

that in population. On the contrary, the share of college graduates of private universities in the sample is lower than that in population. The plausible reason for such difference is that the starting wages of public school graduates are usually higher than those of private school graduates in job market of Taiwan. Therefore, public school graduates may be more willing to respond the questionnaire than private school graduates due to such wage advantage in job market, which results in difference in school distribution between sample and population.

The second source of the data was based on investigations that were directed toward the job market recruiters in human resource department of top 100 business firms in Taiwan. The survey was also conducted via mail, and questions were asked regarding the relative importance of each indicator. The subjective output weights of these recruiters reflect the recognition of relative importance among performance indicators from the perspective of market demand side. We randomly split 100 firms into two equal size sub-samples to elicit recruiters' opinions toward two different types of schools, universities, and universities of Sciences & Technology. Of the 50 questionnaires sent out for assessing universities, 24 were returned and 10 consistent respondents' data were used in the AHP analysis for generating relative importance between output weights and region of weights in this study.

5. Measurement of performance indicators and the setting of lower and upper bounds of weights

The performances of business schools were measured by the "job market performance" and by their "helpfulness in cultivating students' multiple abilities."

5.1. Measurement of the job market performance: application of hierarchical linear model (HLM)

In the survey "What undergraduates learned during their time at university and their job market performance," the questions regarding job market performance included those related to "starting wage" and "current wage". Because the "current wage" is more complicated than the "starting wage", since not only the school will affect the current wage but also experience and work performance, we therefore measure the "job market performance" based on the "starting wage."

Although the "starting wage" is less complicated than the "current wage", the "starting wage" may be affected by the student's talent and the abilities that he or she had already accumulated before entering college. In order to measure the marginal effect of the school on the educational outcomes, we use the "starting wage" as the relevant educational outcome and decompose it into the predicted "starting wage" and the marginal school effects based on the value-added techniques described in Meyer [26].

The value-added model that was developed by Meyer [26] is a HLM. The HLM assumes that the data to be analyzed are hierarchical. Ignoring the hierarchical structure of the data can lead to problems such as aggregation bias, misestimated parameter estimates and standard errors, resulting in an increase in the probability of committing a Type I error [27]. The HLM, by explicitly modeling the hierarchical structure of the data, avoids such a problem.

In his value added model, Meyer [26] used individual and family characteristics variables such as race, ethnicity, gender, special education status, and eligibility for free lunch, whether a family received welfare benefits, parental education and income and family attitudes toward education to explain student

achievement growth. Johnes [1] also employed a similar HLM to rank the universities' teaching efficiency. She used gender, age, marital status, country of origin, type of degree, and type of living accommodation whilst at university as variables to explain students' achievement at university. In addition to student achievement, a growing body of research has been explored the effect of college characteristics or quality on graduates' earnings using the HLM (examples as Zhang [28], Rumberger and Thomas [29]). Almost without exception, these studies assume individual earnings to be a function of individual and school levels variables. While the individual-level variables consist of demographic, family background, education, and labor market variables, the main school-level variables include types of institutional control, mean SAT scores of entering freshmen and other college selectivity variables.

In this paper, we also employ a two-level HLM where students (level 1) are nested within universities (level 2). Similar to previous researches on wage determination, we include student's individual and family characteristics as well as education background as variables to explain college graduates' wage difference. Given the data we collected, we use gender, mother's education status, father's education status, and whether live in capital city (Taipei) as variables for student's characteristics. The dependent variable is the starting wage of college graduates. Previous studies indicated that students' ability would significantly affect graduates' earnings [30,31]. To measure the pure college effect on starting wage, we include the student's college entrance examination score (CEES) as a control variable to exclude pre-college ability impact on wage, since the CEES can be regarded as a proxy for the graduate's talent and accumulated ability before entering college.¹ The educational background variables consist of student's education status and graduate point average (GPA) at college. As for the specification in level 2, we follow the simple two-level HLM of Johnes [1].

Thus, our HLM can be specified as

$$Wage_{ik} = \beta_{0ik} + \beta_1 CEES_{ik} + \beta_2 StuChar_{ik} + \beta_3 EduBack_{ik} \quad (\text{level 1})$$

$$\beta_{0ik} = \beta_0 + \eta_i + \varepsilon_{ik} \quad (\text{level 2}) \quad (5)$$

where $Wage_{ik}$ is the starting wage for the i -th college and k -th graduate. $CEES_{ik}$ is the college entrance examination score for the i -th college and k -th graduate. $StuChar_{ik}$ is the student's characteristics for the i -th college and k -th graduate. $StuChar$ consists of four variables: gender, whether reside at the capital city (Taipei), mother's education status, and father's education status. Generally speaking, the starting wage of a female is less than that of a male [32]. For this reason, we include the gender variable in Eq. (6). $Gender_{ik}$ is a dummy variable, where male=1 and female=0. $ResCap_{ik}$ is a dummy variable for whether reside at capital city, where yes=1 and no=0. The cost of living at capital city is relatively high, thus those who residing at capital city ($ResCap_{ik}=1$) could be paid more by employers. $Mother_edu_{ik}$ and $Father_edu_{ik}$, family characteristics, are defined to be mother's and father's schooling years, respectively. Previous studies have found that family background tends to have a greater impact on how much schooling individuals received than on earnings of college graduates [33].

¹ In Taiwan, all high school graduates must take a comprehensive college entrance exam right after their college graduation if they want to enter college. There are two kinds of universal college entrance exams, one for students who want to enter nature and sciences departments, and the other one for students who intend to enroll humanity, social and management sciences departments. The prospective students for business colleges are required to take the latter one. Since all our sample students were taking the same exam, therefore, their CEE scores can be regarded as a proxy for the graduate's talent and accumulated ability before entering college.

The EduBack_{ik} consists of two variables: Degree_{ik} represents educational background during the first job. It is a dummy variable, and if the individual has obtained a master degree, then Degree = 1, and if not, then Degree = 0.² Another educational background variable is GPA of graduate at college. Since GPA may reflect cognitive ability cultivated at school, thus positive relationship between student's GPA and starting wage is expected.

In Eq. (6) above, β₁, β₂, and β₃ are the coefficients of CEES, StuChar, and EduBack, respectively. Intuitively, the higher the pre-college ability that a graduate has, the higher the starting wage that he (or she) should obtain. Graduates with better student's characteristics or being males are expected to have higher starting wage, whereas graduates with better educational background also earn relatively high starting wage. For the above reasons, we expect β₁, β₂, and β₃ to all will be positive.

Note that β₀ + η_i is the college-specific contribution to the starting wage and ε_{ik} is the individual's deviation from the college's contribution. Furthermore, the β₀ is the mean value across all colleges, and η_i is the deviation from the mean. The term of interest in this business colleges' performance evaluation study is η_i which is defined as the college effect. Because the colleges are assumed to be a random sample from the population of universities, the η_i are therefore distributed among colleges and are assumed to be normally distributed with mean 0 and variance σ_η². The student residuals (ε_{ik}) are also normally distributed with mean 0 and variance σ_ε². The η_i can be estimated as follows:

$$\hat{\eta}_i = \frac{n_i \sigma_\eta^2}{n_i \sigma_\eta^2 + \sigma_\epsilon^2} \frac{\sum_k (Wage_{ik} - \hat{W}age_{ik})}{n_i} \quad (6)$$

Where n_i represents number of sample graduates in the i-th college, $\hat{W}age$ is the predicted value of Wage using the estimated Eq. (6).

Some colleges that add less value than the sample average have a negative output measure. Since DEA cannot handle negative outputs, we follow Meyer [26] and transform the value-added into tractable per-graduate output measures by adding the mean of the starting wage to the value-added. Therefore,

$$Wage_adj_i = \overline{Wage} + \eta_i \quad (7)$$

is our proxy for the output of the i-th college.

We present the estimates of Eq. (6) in Models 1 and 2 of Table 1. Results of Model 1 show that except for CEES, gender and degree, all other explanatory variables in Eq. (6) are not significant at 5% level. Using the likelihood ratio test for model selection between Models 1 and 2, we find that Model 2 is the better one. Results of Model 2 show that both estimated coefficients of CEES (β₁) and degree (β₃) are positive and significant under α < 0.01, while gender (β₂) is also positive and significant under α < 0.05. The sign of each coefficient is consistent with our expectations. β₁ = 30.37 indicates that, other things being equal, on average a 10 points more of CEES (higher pre-college ability) that a graduate has, a NT\$303.7 more of the starting wage that he (or she) should receive. β₃ = 4,974.7 indicates that people who obtain a master's degree will receive a NT\$4974.7 higher starting wage than those who only complete a bachelor's degree; estimated coefficient of gender, β₂ = 504.9 indicates that males will receive a NT\$504.9 higher starting wage than females, others being equal. These estimated parameters are then used to compute predicted wage and adjusted starting wage as in Eqs. (7) and (8).

By comparing the adjusted starting wage and unadjusted starting wage³ for each college (see Table 2), we find that each college's rank changes. In the case of the unadjusted starting

Table 1
The HLM estimation of wage function.

| Variable | Parameter | Model 1 | Model 2 |
|--------------------------------------|----------------|----------------------|-----------------------|
| CEES | β ₁ | 28.906** (5.05) | 30.371** (4.83) |
| StuChar: | β ₂ | | |
| Gender | | 549.758* (2.49) | 504.912* (2.30) |
| ResCap | | 622.41 (1.64) | |
| Father_edu | | -5.753 (-0.08) | |
| Mother_edu | | 86.481 (1.16) | |
| EduBack: | β ₃ | | |
| Degree | | 4928.57** (10.24) | 4974.747** (10.40) |
| GPA | | 10.697 (0.38) | |
| Constant | β ₀ | 18573.76** (6.67) | 20079.478** (9.64) |
| Likelihood ratio test Models 1 vs. 2 | | | |
| Chi-square | | 5.59 | - |
| p-Value | | 0.23 | - |

Note: “***” and “**” denote significance at the 1% and 5% significance levels, respectively.

wage, 5 public colleges and 2 private colleges make up the top 7, so that on average the public colleges' performance is better than that of the private colleges. However, in the case of the adjusted wage, there are only 2 public colleges and 5 private colleges among the top 7, in which case so that on average the private colleges' performance is better than that of the public colleges. Because the public college students' CEE scores are higher than those of the private college students, in regard to the unadjusted starting wage, the public colleges' performance is better than that of the private colleges. From this result we can conclude that the differences in average starting wages among colleges may be the result of their differences in individual characteristics. Therefore, if we do not exclude the graduate's individual characteristics when we measure a college's performance, are results may be misleading.

5.2. Measurement of helpfulness in cultivating students' multiple abilities

For this dimension, we use the questionnaire designed to elicit students' satisfaction on the help college can provide in cultivating multiple abilities. In the questionnaire, the question “How do you feel about the helpfulness of your school in cultivating your abilities?” is asked in an attempt to understand the helpfulness of the school. This question includes 9 types of abilities such as: (1) oral expression, (2) data analysis, (3) problem solving, (4) knowledge application in the major field, (5) knowledge application in the non-major field, (6) foreign language skill, (7) computer operation, (8) personal relationships, and (9) leadership. Respondents are evaluated based on a 5-point scale indicating “of no help”, “of little help”, “of moderate help”, “helpful”, and “of great help”. Since the degree of help is a value-added practice that corresponds to this study, it does not need any adjustment.

The helpfulness of schools is measured as follows. We calculate the percentage of alumni who consider the school where they graduated from to be either “helpful” or “of great help”. Supposing that 20% of the alumni of school A think their school is “helpful” or “of great help” in developing their oral expression ability, then the helpfulness of school A in terms of oral

² There is no PhD graduate in the sample.

³ The unadjusted starting wage for each college is the average value of the graduates' starting wage for the respective college.

Table 2
Unadjusted and adjusted starting wages of the graduates of 21 business colleges.

| Classification | Business college | Number of observations | Unadjusted starting wage ^a | Rank | Adjusted starting wage | Rank |
|-----------------|------------------|------------------------|---------------------------------------|------|------------------------|------|
| Public college | Taiwan | 99 | 34,462.62 | 3 | 31,238.69 | 12 |
| | Chung-Hsing | 10 | 35,750.00 | 1 | 31,734.27 | 4 |
| | Cheng-Kung | 10 | 31,805.56 | 8 | 31,266.22 | 10 |
| | Chengchi | 129 | 35,215.16 | 2 | 31,946.89 | 1 |
| | Central | 28 | 32,234.85 | 6 | 31,189.65 | 14 |
| | Sun Yat-Sen | 17 | 31,488.10 | 10 | 31,214.89 | 13 |
| | Taipei | 57 | 32,056.45 | 7 | 31,053.14 | 16 |
| Private college | Soochow | 89 | 32,560.24 | 4 | 31,592.62 | 5 |
| | Chung-Yuan | 56 | 29,910.71 | 13 | 31,021.45 | 17 |
| | Tamkang | 127 | 31,425.78 | 11 | 31,381.14 | 9 |
| | Tung-Hai | 63 | 30,496.03 | 12 | 31,119.03 | 15 |
| | Feng-Chia | 155 | 28,571.43 | 19 | 30,360.57 | 21 |
| | Chinese Culture | 40 | 28,881.58 | 17 | 30,882.54 | 20 |
| | Providence | 55 | 28,750.00 | 18 | 30,932.87 | 19 |
| | Fu-Jen | 57 | 31,527.78 | 9 | 31,543.95 | 6 |
| | Yuan-Ze | 14 | 32,343.75 | 5 | 31,740.53 | 3 |
| | I-Shou | 18 | 26,637.93 | 21 | 31,257.56 | 11 |
| | Ming-Chuan | 130 | 29,710.82 | 14 | 31,780.91 | 2 |
| | Shih-Chien | 62 | 29,233.87 | 16 | 31,531.96 | 7 |
| | Chung-Hua | 12 | 29,619.57 | 15 | 31,530.31 | 8 |
| | Aletheia | 22 | 28,238.64 | 20 | 30,940.40 | 18 |

Note: The unadjusted starting wage of each college is the average value of the graduates' starting wage for that respective college.

Table 3
Performance scores in the cognitive, psychomotor, and affective domains.

| Classification | Business college | Cognitive | Rank | Psychomotor | Rank | Affective | Rank |
|-----------------|------------------|-----------|--------|-------------|--------|-----------|--------|
| Public colleges | Taiwan | 1.0000 | 1 | 0.7192 | 5 | 0.7874 | 6 |
| | Chung-Hsing | 0.8538 | 7 | 0.1312 | 21 | 0.6333 | 16 |
| | Cheng-Kung | 0.8818 | 5 | 0.4668 | 17 | 1.0000 | 1 |
| | Chengchi | 0.8328 | 8 | 0.6456 | 11 | 0.6485 | 15 |
| | Central | 0.8036 | 10 | 0.6176 | 14 | 0.9314 | 2 |
| | Sun Yat-Sen | 1.0000 | 1 | 1.0000 | 1 | 0.8444 | 3 |
| | Taipei | 0.6311 | 18 | 0.4565 | 18 | 0.5651 | 19 |
| | Mean | 0.8576 | | 0.5767 | | 0.7729 | |
| | Private colleges | Soochow | 0.7205 | 13 | 0.5936 | 15 | 0.6974 |
| Chung-Yuan | | 0.8094 | 9 | 0.7154 | 8 | 0.8444 | 3 |
| Tamkang | | 0.7186 | 14 | 0.7914 | 4 | 0.7451 | 9 |
| Tung-Hai | | 0.7234 | 12 | 0.5798 | 16 | 0.7559 | 8 |
| Feng-Chia | | 0.6377 | 17 | 0.624 | 13 | 0.7261 | 11 |
| Chinese Culture | | 0.5731 | 20 | 0.2168 | 20 | 0.5521 | 20 |
| Providence | | 0.7145 | 15 | 0.6574 | 10 | 0.6333 | 16 |
| Fu-Jen | | 0.8559 | 6 | 0.7166 | 7 | 0.6988 | 13 |
| Yuan-Ze | | 0.8857 | 4 | 0.9794 | 3 | 0.7015 | 12 |
| I-Shou | | 0.9482 | 3 | 0.6716 | 9 | 0.7862 | 7 |
| Ming-Chuan | | 0.6277 | 19 | 1.0000 | 1 | 0.5868 | 18 |
| Shih-Chien | | 0.6641 | 16 | 0.6262 | 12 | 0.7355 | 10 |
| Chung-Hua | | 0.7937 | 11 | 0.7185 | 6 | 0.8444 | 3 |
| Aletheia | | 0.3565 | 21 | 0.3594 | 19 | 0.4318 | 21 |
| Mean | 0.7164 | | 0.6607 | | 0.6957 | | |

expression is labeled as 0.2. By following this procedure, the helpfulness of each school in terms of the above 9 abilities can be obtained.

If there are too many outputs or inputs in DEA, the results may be less discriminating, and hence we use AHP to determine the weight restrictions for output. Asking experts to make a pairwise comparison of the relative importance of 9 items is too complicated. To avoid this problem, we follow Chen [34]. First, we classify the 9 items according to the "cognitive", "psychomotor", and "affective" domains. In the "cognitive" domain, (1) oral expression, (2) data analysis, (3) problem solving, (4) knowledge application in the major field, and (5) knowledge application in the non-major field are included. In the "psychomotor" domain,

(6) foreign language skill and (7) computer operation are included. In the "affective" domain, (8) personal relationships, and (9) leadership are included. Second, we use values in items (1)–(5) as outputs to calculate the performance score for "cognitive" using the CCR model. In this way, we can also obtain the performance scores for domains of "psychomotor" and "affective".

We summarize the performance scores for the three domains in Table 3. On average, the public colleges' performance scores in the "cognitive" and "affective" domains (0.8576 and 0.7729, respectively) are higher than those of the private colleges (0.7164 and 0.6957). On the contrary, the private colleges' performance score in the "psychomotor" domain (0.6607) is higher than that for the public colleges (0.5767). It is shown that

the graduates of public colleges are more satisfied than the graduates of private colleges with their colleges' performance in the "cognitive" and "affective" domains. On the other hand, the private colleges' graduates are more satisfied than public colleges' graduates with their colleges' performance in the "psychomotor" domain.

In the following section, we will use four outputs to measure the business colleges' performance in the proposed AR-CCR model, namely, the adjusted starting wage ($W_{\text{age_adj}}$), the performance scores of the cognitive domain (A_1), the psychomotor domain (A_2) and the affective domain (A_3). In the conventional DEA, it is difficult to compare the relative importance of output items with their weights, because the weights largely depend on the unit of measurement in the original data. To generate unit free weights, we normalize the given output data based on one targeted college (DMU_0) so that the output weights represent the degree of importance of the items. The normalized j -th output data of the i -th college are obtained as Y_{ij}/Y_{0j} . Therefore, we use the proposed AR-CCR model with the normalized output data to evaluate each sample college's performance.

5.3. The weights restriction of outputs from AHP

To insure reasonable weight regions for each output in the proposed AR-CCR model, we conducted a survey to key recruiters in human resource department of top 100 business firms in Taiwan and elicit their subjective judgments on the relative importance between outputs. The analytic hierarchy process (AHP) proposed by Saaty [14] was used to quantify their subject judgments. The results derived from this AHP analysis then served as a guideline for setting the upper and lower bounds of (4). The recruiters' derived weights attached to the outputs of the four types of performance are shown in Table 4.⁴

In Table 4, $W_{\text{age_adj}}$ denotes the weight of adjusted starting wage ($W_{\text{age_adj}}$); W_{A_1} denotes the weight of cognitive ability (A_1); W_{A_2} denotes the weight of psychomotor ability (A_2); and W_{A_3} denotes the weight of affective ability (A_3). The summation of W_{A_1} , W_{A_2} , and W_{A_3} represents the weight of helpfulness in cultivating students' multiple abilities, W_{MA} . Weight values in Table 4 reflect relative importance of performance indicators from demanders of job market. Results of weights in Table 4 showed that the mean of $W_{\text{age_adj}}$ and W_{MA} were 0.533 and 0.467, respectively. On average, in recruiters' minds, vocational performance is more important than helpfulness in cultivating students' multiple abilities. By comparing the importance of the "cognitive", "psychomotor", and "affective" domains, the means of "cognitive", "psychomotor", and "affective" were 0.150, 0.107, and 0.202, which implies that in recruiters' minds helpfulness in cultivating students' affective abilities is more important than enhancing the other domains. Table 4 also indicates that individual recruiters' weights may differ from other recruiters'. To consider such weight variations, the weight ranges of recruiters will be used for further AR analysis.

To incorporate these weights in the AR-CCR model, we made pairwise divisions between weights. We then find the largest and smallest values of each weight ratio for all recruiters and construct the upper and lower bounds values of such weight ratio. For example, the ratio W_{A_1}/W_{A_2} takes on a value of $0.319/0.129=2.47$ for recruiter 1 and $0.154/0.039=4$ for recruiter 2. We can also calculate the ratio W_{A_1}/W_{A_2} for rest of 8 recruiters. Thus,

Table 4
AHP weights of outputs of recruiters.

| | $W_{\text{age_adj}}$ | W_{MA} | W_{A_1} | W_{A_2} | W_{A_3} | C.I. |
|---------------|-----------------------|----------|-----------|-----------|-----------|------|
| Recruiter 1 | 0.500 | 0.500 | 0.319 | 0.129 | 0.053 | 0.04 |
| Recruiter 2 | 0.500 | 0.500 | 0.154 | 0.039 | 0.308 | 0 |
| Recruiter 3 | 0.900 | 0.100 | 0.013 | 0.010 | 0.077 | 0.04 |
| Recruiter 4 | 0.500 | 0.500 | 0.167 | 0.167 | 0.167 | 0 |
| Recruiter 5 | 0.875 | 0.125 | 0.058 | 0.008 | 0.058 | 0 |
| Recruiter 6 | 0.800 | 0.200 | 0.097 | 0.016 | 0.087 | 0.01 |
| Recruiter 7 | 0.167 | 0.833 | 0.277 | 0.277 | 0.277 | 0 |
| Recruiter 8 | 0.667 | 0.333 | 0.009 | 0.212 | 0.035 | 0.04 |
| Recruiter 9 | 0.167 | 0.833 | 0.087 | 0.106 | 0.641 | 0.04 |
| Recruiter 10 | 0.250 | 0.750 | 0.322 | 0.107 | 0.322 | 0 |
| Mean Standard | 0.533 | 0.467 | 0.150 | 0.107 | 0.202 | |
| Deviation | 0.277 | 0.277 | 0.119 | 0.092 | 0.190 | |
| Maximum | 0.900 | 0.833 | 0.322 | 0.277 | 0.641 | |
| Minimum | 0.167 | 0.100 | 0.009 | 0.008 | 0.035 | |

Note: $W_{\text{age_adj}}$ denotes the weight of adjusted starting wage, W_{MA} denotes the weight of multiple abilities, W_{A_1} denotes the weight of cognitive, W_{A_2} denotes the weight of psychomotor, W_{A_3} denotes the weight of affective, and C.I. denotes consistency ratio.

Table 5
Upper and lower bounds of output weight ratios.

| Output weight ratio | Upper bound | Lower bound |
|-------------------------------|-------------|-------------|
| W_{A_1}/W_{A_2} | 7.000 | 0.041 |
| W_{A_1}/W_{A_3} | 6.067 | 0.135 |
| W_{A_2}/W_{A_3} | 6.067 | 0.125 |
| $W_{\text{age_adj}}/W_{A_1}$ | 77.636 | 0.602 |
| $W_{\text{age_adj}}/W_{A_2}$ | 104.478 | 0.602 |
| $W_{\text{age_adj}}/W_{A_3}$ | 19.076 | 0.261 |

we have used the highest $W_{A_1}/W_{A_2}=7$ from recruiter 5 as the upper bound of the ratio W_{A_1}/W_{A_2} , and the smallest $W_{A_1}/W_{A_2}=0.041$ from recruiter 8 as the lower bound. Therefore, the range of W_{A_1}/W_{A_2} is $0.041 \leq W_{A_1}/W_{A_2} \leq 7$. This ratio weight inequality constraint will be incorporated in the AR-CCR model. Other ranges (or upper and lower bounds) of ratio weights can be found in Table 5.

6. Empirical results

To compare the model performances of the CCR and the proposed AR-CCR, we employ both models to measure the colleges' performances. The outputs used in both models are the same, and include adjusted starting wage ($W_{\text{age_adj}}$) and three kinds of ability cultivation, cognitive(A_1), psychomotor(A_2), and affective(A_3) abilities. We summarize the results of the CCR and the results of the proposed AR-CCR in Table 6.

The results of performance score and rank for the CCR in Table 6 show that there are 7 colleges with performance scores of 1 (rank the first), which can be compared to those results from the proposed AR-CCR in Table 6 where there are only 3 colleges with the performance score=1. In addition, the coefficient of variation (CV) of performance score for the whole sample, calculated by dividing mean score by its standard deviation, in the CCR is 0.014, whereas such coefficient is 0.017 in the proposed AR-CCR. Therefore, both the performance scores and CV indicate the proposed AR-CCR possesses better discriminatory power than the CCR.

The benefit of using the proposed AR-CCR is to avoid extreme weight distribution. The weight distribution of outputs for the CCR is shown in Table 6. One would easily find that there are many zero output weights for selected outputs, which is

⁴ We use the consistency ratio which is provided by Saaty [14] to identify whether the recruiters' response is inconsistent or not, for if the consistency ratio < 0.1, then his/her responses will be consistent. Among 24 returned recruiters, 10 of them have passed the consistency test. The data of these 10 recruiters were used to derive weights in Table 4.

Table 6
Output weights and performance score of business colleges calculated by the CCR and the proposed AR-CCR model.

| Ownership | Business college | CCR (without weight restriction) | | | | | | AR-CCR (weight restricted by recruiter) | | | | | |
|------------------|------------------|----------------------------------|----------|----------|----------|-------------|-------|---|----------|----------|----------|-------------|------|
| | | Output weights | | | | Performance | | Output weights (recruiters) | | | | Performance | |
| | | W_{wage_adj} | W_{A1} | W_{A2} | W_{A3} | Score | Rank | W_{wage_adj} | W_{A1} | W_{A2} | W_{A3} | Score | Rank |
| Public colleges | Taiwan | 0.877 | 0.123 | 0 | 0 | 1.000 | 1 | 0.881 | 0.059 | 0.008 | 0.046 | 0.994 | 8 |
| | Chung-Hsing | 0.888 | 0.111 | 0 | 0 | 0.996 | 8 | 0.920 | 0.012 | 0.009 | 0.048 | 0.985 | 14 |
| | Cheng-Kung | 0 | 0 | 0 | 0.787 | 1.000 | 1 | 0.151 | 0.078 | 0.072 | 0.578 | 1.000 | 1 |
| | Chengchi | 0.978 | 0 | 0 | 0 | 1.000 | 1 | 0.920 | 0.012 | 0.009 | 0.048 | 0.998 | 5 |
| | Central | 0.884 | 0 | 0.023 | 0.079 | 0.996 | 8 | 0.885 | 0.011 | 0.019 | 0.072 | 0.995 | 7 |
| | Sun Yat-Sen | 0 | 0 | 0.719 | 0 | 1.000 | 1 | 0.662 | 0.009 | 0.211 | 0.035 | 1.000 | 1 |
| | Taipei | 0.978 | 0 | 0 | 0 | 0.972 | 17 | 0.920 | 0.012 | 0.009 | 0.048 | 0.962 | 18 |
| | Mean | 0.658 | 0.033 | 0.106 | 0.124 | 0.995 | | 0.763 | 0.028 | 0.048 | 0.125 | 0.991 | |
| Private Colleges | Soochow | 0.941 | 0 | 0 | 0.046 | 0.992 | 12 | 0.916 | 0.012 | 0.009 | 0.053 | 0.989 | 11 |
| | Chung-Yuan | 0.884 | 0 | 0.023 | 0.079 | 0.985 | 15 | 0.885 | 0.011 | 0.019 | 0.072 | 0.985 | 14 |
| | Tamkang | 0.934 | 0 | 0.006 | 0.048 | 0.990 | 14 | 0.916 | 0.012 | 0.009 | 0.053 | 0.988 | 13 |
| | Tung-Hai | 0.935 | 0 | 0.005 | 0.048 | 0.981 | 16 | 0.916 | 0.012 | 0.009 | 0.053 | 0.979 | 16 |
| | Feng-Chia | 0.935 | 0 | 0.005 | 0.048 | 0.957 | 21 | 0.916 | 0.012 | 0.009 | 0.053 | 0.954 | 19 |
| | Chinese Culture | 0.978 | 0 | 0 | 0 | 0.967 | 20 | 0.920 | 0.012 | 0.009 | 0.048 | 0.953 | 20 |
| | Providence | 0.958 | 0 | 0.012 | 0.011 | 0.969 | 18 | 0.920 | 0.012 | 0.009 | 0.048 | 0.966 | 17 |
| | Fu-Jen | 0.895 | 0.065 | 0 | 0.037 | 0.992 | 12 | 0.916 | 0.012 | 0.009 | 0.053 | 0.991 | 9 |
| | Yuan-Ze | 0.961 | 0.006 | 0.013 | 0 | 1.000 | 1 | 0.920 | 0.012 | 0.009 | 0.048 | 1.000 | 1 |
| | I-Shou | 0.888 | 0.074 | 0 | 0.036 | 0.995 | 10 | 0.887 | 0.046 | 0.008 | 0.052 | 0.991 | 9 |
| | Ming-Chuan | 0.969 | 0 | 0.010 | 0 | 1.000 | 1 | 0.664 | 0.009 | 0.209 | 0.035 | 0.997 | 6 |
| | Shih-Chien | 0.935 | 0 | 0.005 | 0.048 | 0.993 | 11 | 0.916 | 0.012 | 0.009 | 0.053 | 0.989 | 11 |
| | Chung-Hua | 0.934 | 0 | 0.006 | 0.048 | 1.000 | 1 | 0.916 | 0.012 | 0.009 | 0.053 | 0.999 | 4 |
| | Aletheia | 0.978 | 0 | 0 | 0 | 0.969 | 18 | 0.920 | 0.012 | 0.009 | 0.048 | 0.946 | 21 |
| Mean | 0.938 | 0.01 | 0.006 | 0.032 | 0.985 | | 0.895 | 0.014 | 0.024 | 0.052 | 0.981 | | |
| Whole Sample | Mean | 0.844 | 0.018 | 0.039 | 0.063 | 0.988 | | 0.851 | 0.019 | 0.032 | 0.076 | 0.984 | |
| | (SD) | (0.283) | (0.039) | (0.156) | (0.168) | (0.014) | | (0.177) | (0.018) | (0.061) | (0.115) | (0.017) | |

Note: SD denotes standard deviation.

unreasonable when evaluating college's peer performance. Such an unreasonable situation is not found to exist in the proposed AR-CCR, in which all of the output weights are larger than zero. In regard to the discriminatory power and the distribution of output weights, the proposed AR-CCR model makes the DEA results more accurately to reflect the real decision-making situation of college performance evaluation. Consequently, we have decided to analyze the business colleges' performance using the latter model.

Results of the proposed AR-CCR model have shown that the mean performance score for the whole sample colleges is 0.984 (Table 6), whereas mean performance scores for public colleges and private colleges, are 0.991 and 0.981, respectively. In addition, 4 of 7 public colleges are among the top 7 colleges, and 3 of 14 private colleges are among the top 7. Generally speaking, the performance of the public colleges is better than that of the private ones. However, it is worth noting that some private colleges are performing well, for example Yuan-Ze, Chung-Hua, and Ming-Chuan. These private colleges perform better than many public colleges. Although many people in Taiwan think that the public colleges must be better than their private counterparts, and the CEE scores of public colleges are almost always higher than those of private colleges, when performance is evaluated based on the students' wages and the helpfulness of the college, the stereotype that "all of public colleges are better than private colleges" is strongly challenged.

The variables used to measure the helpfulness in cultivating students multiple abilities are self-reported variables. In this study, respondents are asked to evaluate the degree of helpfulness of school to each of these abilities based on a 5-point scale indicating "of no help", "of little help", "of moderate help", "helpful", and "of great help"⁶. Since each respondent will evaluate such degree by comparing level of his/her ability after college graduation with level of his/her ability before entering the

college, therefore, such measures of helpfulness can be regarded as value-added indicators. But respondents' answers to such survey questions are subjective in nature. Given the same amount of school effort received, student A may find it "of little help", whereas student B may respond "of great help". In addition to such subjectivity problem, the users must be aware of potential problems implicitly embedded when using such variables on university rankings. For example, high quality universities would attract high quality students while low quality universities would attract low quality students. One possibility is that high-quality students may have higher standards in judging the helpfulness while low quality students may have lower standards in the judgment. Then, the helpfulness for the high-quality universities would be underrated, while the helpfulness of low-quality universities would be overrated. This may be one reason why private universities tend to have higher ranking when the helpfulness variables are used to rank universities. On the contrary, there is the opposite possibility. For example, high-quality students may be "reasonable" judges while low-quality students may turn out to be "complainers". In this scenario, low quality universities may be unfairly underrated. In either case, self-reported measures seem to have unwanted consequences.^{5,6} Despite the above mentioned problems, using average figures of these self-reported variables in cross school comparisons is still meaningful if we assume such effects from "mis-judging" or "complainers" for high and low quality schools to be almost similar.

At last, by examining the weight structure between wage and multiple abilities from the proposed AR-CCR model in Table 6, we

⁵ Similar design can be found in the paper of Colbert et al. [35], where they used student satisfaction with teaching, curriculum, and placement as outputs in evaluating relative efficiency of MBA programs.

⁶ We appreciate one referee who provides this valuable comment.

find that the weight values of “Wage_adj” in public schools are lower than those in private schools, whereas the weight values of “multiple abilities (cognitive, psychomotor, and affective)” in public schools are found to be higher than private schools. Such results seem to indicate that public schools have comparative advantage on “students’ multiple abilities cultivation”, while private schools have comparative advantage over public schools on “value-added starting wage”. In addition, Table 6 shows that mean weight values of affective ability (W_{A3}) for both public and private schools are higher than those of cognitive and psychomotor abilities, (W_{A1} and W_{A2}). Therefore, performance of a school may be improved more effectively by promoting student’s affective ability than by promoting the other two abilities.

7. Concluding remarks

Our study seeks to construct a student-based performance evaluation model for the business schools in Taiwan. In order to avoid unreasonable measures of college performance, we measure the performance of the colleges using added-value outputs that are combined with the proposed AR-CCR approach. The starting wage representing graduate job market performance and the satisfaction to 9 types of abilities representing helpfulness of school in cultivating student multiple abilities are used as performance indicators. Through the use of the HLM, the starting wage has been adjusted to be a value added indicator as those multiple abilities indicators. The empirical findings of this research have shown if we do not exclude the factors that are due to the graduates’ individual characteristics when we measure the colleges’ performance, our results may be misleading. The performance scores calculated when using the proposed AR-CCR are more discriminatory than the performance scores calculated when using the CCR, and the output weights calculated using the proposed AR-CCR are more reasonable than the output weights which are calculated using the CCR. Therefore, in evaluating Taiwanese college performance we should use the AR-CCR to improve standard CCR results.

When measuring the Taiwanese business colleges’ performance using the proposed AR-CCR, we also find that on average the public colleges perform better than the private ones. However, it is worth noting that some private colleges are performing well. These private colleges perform better than many public colleges. Our results suggest that the government should also use “value-added” student based performance indicators to assess the performance of colleges. For these private universities with good performance in value added indicators, they can emphasize such value-added contribution in their freshmen recruiting activities. Empirical results from this university performance analysis combined with traditional enrollment predictions could be very helpful to prospective college students and their parents who are in the process of choosing a college to attend.

At last, both the AHP weights expressed by recruiters and optimal weights solved by the AR-CCR model have indicated that affective ability is most appreciated by job market recruiters and most effective in improving school performance, as compared to the other two (cognitive and psychomotor) abilities. Therefore, cultivating student’s affective ability such as developing personal relationships, leadership, and team work skills in school should be well appreciated and prioritized by both students and school administrators. School administrators should try hard to offer more training and courses related to affective ability cultivation or to create better extracurricular environment for students in schools, whereas students are encouraged to attend such courses and activities to improve their employability at the time of their graduation.

References

- [1] Johnes J. Measuring efficiency: A comparison of multilevel modelling and data envelopment analysis in the context of higher education. *Bulletin of Economic Research* 2006;58(2):75–104.
- [2] Coelli T, Rao DSP, Battese GE. An introduction to efficiency and productivity analysis. Boston: Kluwer Academic Publisher; 1998.
- [3] Hashimoto K, Cohn E. Economics of scale and scope in Japanese private universities. *Education Economics* 1997;5(2):107–15.
- [4] Abbott M, Doucouliagos C. The efficiency of Australian universities: A data envelopment analysis. *Economics of Education Review* 2003;22(1):89–97.
- [5] Johnes J, Johnes G. Measuring the research performance of UK economics departments: An application of data envelopment analysis. *Oxford Economic Papers* 1993;45(2):332–47.
- [6] Cherchye L, Abeeel PV. On research efficiency: A micro-analysis of Dutch university research in economics and business management. *Research Policy* 2005;34(4):495–516.
- [7] Athanassopoulos AD, Shape E. Assessing the comparative efficiency of higher education institutions in the UK by means of data envelopment analysis. *Education Economics* 2005;5(2):119–34.
- [8] Casu B, Thanassoulis E. Evaluating cost efficiency in central administrative services in UK Universities. *Omega* 2006;34:417–26.
- [9] Kao C, Hung HT. Efficiency analysis of university departments: An empirical study. *Omega* 2008;36(4):653–64.
- [10] Kao C. Evaluation of junior colleges of technology: The Taiwan case. *European Journal of Operational Research* 1994;72:43–51.
- [11] Zheng H.Y., Stewart A.A. Assessing the effectiveness of public research universities: Using NSF/NCES data and data envelopment analysis techniques. AIR 2000 Annual Forum Paper 2000.
- [12] Doyle J, Green R. Efficiency and cross-efficiency in DEA: Derivations, meanings and uses. *Journal of the Operational Research Society* 1994;45(5):567–78.
- [13] Thompson RG, Singleton FD, Thrall RM, Smith BA. Comparative site evaluations for locating a high-energy physics lab in Texas. *Interfaces* 1986;16:35–49.
- [14] Saaty TL. *The analytic hierarchy process*. New York: McGraw-Hill; 1980.
- [15] Zhu J. DEA/AR analysis of the 1988–1989 performance of the Nanjing textiles corporation. *Annals of Operations Research* 1996;66:311–35.
- [16] Seifert LM, Zhu J. Identifying excesses and deficits in Chinese industrial productivity (1953–1990): A weighted data envelopment analysis approach. *Omega* 1998;26(2):279–96.
- [17] Takamura Y, Tone KA. Comparative site evaluation study for relocating Japanese government agencies out of Tokyo. *Socio-Economic Planning Sciences* 2003;37(2):85–102.
- [18] Sun S. Assessing joint maintenance shops in the Taiwanese army using data envelopment analysis. *Journal of Operations Management* 2004;22:233–45.
- [19] Meng W, Zhang D, Qi L, Liu W. Two-level DEA approaches in research evaluation. *Omega* 2008;36(6):950–7.
- [20] Grosskopf S, Hayes KJ, Taylor LL, Weber WL. Anticipating the consequences of school reform: A new use of DEA. *Management Science* 1999;45(4):608–20.
- [21] Bessent AM, Bessent EW. Determining the comparative efficiency of schools through data envelopment analysis. *Educational Administration Quarterly* 1980;16(2):57–75.
- [22] Ruggiero J. Nonparametric estimation of returns to scale in the public sector with an application to the provision of educational services. *Journal of the Operational Research Society* 2000;51:906–12.
- [23] Bloom B.S. Taxonomy of educational objectives: The Classification of Education Goal. Handbook 1: Cognitive Domain, Mckey, New York, 1956.
- [24] Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. *European Journal of Operational Research* 1978;2:429–44.
- [25] Jeremy AL, Lovell CAK, Murray CJL, Evans DB. World health system performance revisited: The impact of varying the relative importance of health system goals. *BMC Health Services Research* 2004;4(19):1–8.
- [26] Meyer RH. Value-added indicators of school performance: A Primer. *Economics of Education Review* 1997;16(3):283–301.
- [27] Hill P, Rowe K. Multilevel modeling in school effectiveness research. *School Effectiveness and School Improvement* 1996;7:1–34.
- [28] Zhang L. Do measures of college quality matter? The effect of college quality on graduates’ earnings The review of Higher Education 2005;28(4):571–96.
- [29] Rumberger RW, Thomas SL. The economic RETURNS to college major, quality and performance: A multilevel analysis of recent graduates. *Economics of Education Review* 1993;12(1):1–19.
- [30] Hunt SJ. Income determinants for college graduates and the return to educational investment. *Yale Economic Essays* 1963;3:305–57.
- [31] Solmon LC. The definition and impact of college quality. In: Solmon LC, Taubman PJ, editors. *Does College Matter?* New York: Academic Press; 1973. p. 77–105.
- [32] Gerhart B. Gender differences in current and starting salaries: The role of performance, college major, and job title. *Industrial and Labor Relations Review* 1990;43(4):418–43.
- [33] Bretz RD. College grade point average as a predictor of adult success: A meta-analytic review and some additional evidence. *Public Personnel Management* 1989;18:11–22.
- [34] Chen TY. Measuring firm performance with DEA and prior information in Taiwan’s banks. *Applied Economics Letters* 2002;9(3):201–4.
- [35] Colbert A, Levary RR, Shaner MC. Determining the relative efficiency of MBA programs using DEA. *European Journal of Operational Research* 2000:656–69.