

El-Mahgary, Sami; Rönholm, Petri; Hyypä, Hannu; Haggeén, Henrik;
Koponen, Jenni

Article

Evaluating the performance of university course units using data envelopment analysis

Cogent Economics & Finance

Provided in Cooperation with:

Taylor & Francis Group

Suggested Citation: El-Mahgary, Sami; Rönholm, Petri; Hyypä, Hannu; Haggeén, Henrik; Koponen, Jenni (2014) : Evaluating the performance of university course units using data envelopment analysis, Cogent Economics & Finance, ISSN 2332-2039, Taylor & Francis, Abingdon, Vol. 2, Iss. 1, pp. 2-19,
<http://dx.doi.org/10.1080/23322039.2014.918856>

This Version is available at:

<http://hdl.handle.net/10419/147703>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/3.0/>

COURSE UNITS (Modules)



CS-101



LIT-202



ECO-103



ART-211



PHY-310



SOC-420



BIO-323

RESEARCH ARTICLE

Evaluating the performance of university course units using data envelopment analysis

Sami El-Mahgary, Petri Rönholm, Hannu Hyyppä, Henrik Haggrén and Jenni Koponen

Cogent Economics & Finance (2014), 2: 918856



Received: 19 November 2013
Accepted: 21 April 2014
Published: 23 May 2014

*Corresponding author: Sami El-Mahgary,
Department of Computer Science
and Engineering, Room A145, Aalto
University, P.O. Box 15400, Aalto 00076,
Finland
E-mail: sami.mahgary@aalto.fi

Reviewing Editor:
Steve Cook, Swansea University

Further author and article information is
available at the end of the article

RESEARCH ARTICLE

Evaluating the performance of university course units using data envelopment analysis

Sami El-Mahgary^{1*}, Petri Rönholm², Hannu Hyypä², Henrik Haggrén² and Jenni Koponen³

Abstract: The technique of data envelopment analysis (DEA) for measuring the relative efficiency has been widely used in the higher education sector. However, measuring the performance of a set of course units or modules that are part of a university curriculum has received little attention. In this article, DEA was used in a visual way to measure the performance of twelve course units that are part of a Photogrammetry curriculum taught at Aalto University. The results pinpointed the weakest performing units, i.e. units where the provided teaching efforts might not be adequately reflected in the students' marks in the unit. Based on the results, a single unit was considered to offer poor performance with respect to its teaching resources and was selected as a candidate for revision of its contents. Financial resources were not used as such; instead, the performance of students in previous pre-requisite units was used as the inputs. For clarity, a single output covering the overall student performance in the examined unit was used. The technique should be widely applicable assuming the grade point averages of the students who took the course unit are available along with the marks obtained in the evaluated units and their pre-requisites.

Keywords: data envelopment analysis, performance, efficiency, course units

AUTHOR BIOGRAPHY



Sami El-Mahgary obtained his licentiate degree in Computer Science in 2013 and is now a doctoral student focusing on the retrieval and analysis of data related to the performance of students in the courses they have taken. Petri Rönholm is a Senior University Lecturer and is active in both academic and research work, being the author of over 70 publications. Hannu Hyypä was Research Director at Helsinki University of Technology at the time the research was carried out, and is now Technology Manager at the School of Civil Engineering and Building Services at Helsinki Metropolia University of Applied Sciences and a docent at Aalto University. Henrik Haggrén is a full professor of Photogrammetry. His primary research interest is in developing photogrammetry as a medium for creative and innovative imaging applications. Jenni Koponen received her M.Sc. from Helsinki University of Technology, and is currently working as an educational developer.

PUBLIC INTEREST STATEMENT

This article applies a technique developed for measuring relative efficiency, known as data-envelopment analysis (DEA for short) in a novel application. Rather than measuring the efficiency of universities or their departments, DEA was used to measure the performance of a set of university course units that were all taught at an engineering department in a Finnish university. By taking into consideration the students' standings in the pre-requisites for the examined units as well as their performance in these same units, an efficiency score (0–100%) was obtained for each course unit. Units with a relatively low score were examined closer as they might be proving too difficult for the students and are also likely to make poor use of the teaching resources. We found the results useful as they pinpointed a course unit which required our attention as to how its contents could be improved and its teaching carried out more effectively.

1. Introduction

The performances of major universities are often monitored, assessed and ranked at regular intervals. Such an assessment is usually based on a set of specific criteria, generally known as *performance indicators* or PIs (Barnetson & Cutright, 2000). For instance, two common PIs used to measure university research are the number of publications and the number of supervised Ph.D. theses (Tzeremes & Halkos 2010; Köksal & Nalçacı, 2006; Martín, 2007). Another PI often stressed is the students' job placement upon graduation (Maingot & Zeghal, 2008; Martín, 2007). Job placement may, however, reflect a university's reputation more than teaching quality itself. Nevertheless, the studies by Biggeri and Bini (2001) and Mohamad Ishak, Suhaida, and Yuzainee (2009) confirm that indeed a wealth of PIs have been developed for use in higher education.

Even though the reliability of PIs has been questioned by several sources as shown in a compilation by Harvey (1999), there are simple and widely accepted indicators in higher education such as the *grade-point average* (GPA) that are not particularly prone to bias or results distortion. The GPA, a basic indicator of student performance, is readily available from a student database system as pointed out by Young (1993) and is easy to compute, being simply the arithmetic mean of the grades obtained from the course units (also known as modules) taken by the student. If the GPA is further weighed according to the number of credits or European Credit Transfer System (ECTS) points awarded for completing each course unit (units for short), we get the Weighed Grade-Point Average or *WGPA*, often a slightly more accurate reflection of a student's performance.

The primary purpose of this study is to identify those core curriculum course units with the lowest efficiency scores, as such units will have the weakest return on the allocation of their teaching resources. It is expected that such units would benefit the most from a revision of their contents so as to lead to improved student learning. The relative performance score of each course unit is measured based on the performance of the students who completed the unit, with respect to how well the students were prepared for taking the unit in question.

The technique used here is a simple application of data envelopment analysis (DEA), developed by Charnes, Cooper, and Rhodes (1978) who applied it to measuring the efficiency of similar organizations in the public sector. These organisations whose efficiencies were measured were referred to as *decision-making units* or *units* for short, to stress the fact that a unit can, independently, make decisions to try and improve its performance by reducing expenses, for example. The term "units" in DEA thus refers to the collection of units whose efficiency is being measured and which are subject to similar operating or teaching practices. Furthermore, in the DEA technique, each PI is known as a *factor* and is further classified as an *input* or as an *output*. An input is a PI (or factor) which expresses the consumption of a resource or takes into account some qualitative trait; an output is a factor which expresses the transformation of a resource or describes a qualitative trait of that transformed resource. Inputs and outputs are defined so that an increase in an input value does not result in a decrease in any output value.

The DEA technique has been widely applied to measuring the efficiency of many kinds of different units such as universities and their departments: Beasley (1995), Hanke and Leopoldseder (1998), Johnes (2006a), Bobe (2009), Tzeremes and Halkos (2010), and Alwadood, Noor, and Kamarudin (2011). These six studies, summarised in Table 1, differ basically as to whether financial resources are incorporated into one of the inputs. Instead of yearly budgets, in Johnes (2006a) input is measured through the quality of entering students by including their A-level exam scores while the inputs in Alwadood et al. (2011) take into account the faculty to student ratio and the total credit hours offered to students. Interestingly, the study by Johnes (2006a) is based on the premise that females achieve better scores than males, which means that the DEA analysis in question expects more out of females than males. In the study by Beasley (1995), research income has a dual nature: it acts as an output to approximate the number of publications and as an input to emphasize that it is a resource that should be invested wisely.

Table 1. The Main PIs Used in the Six Mentioned Studies

| Study | Applied on | Inputs based on | Outputs based on |
|-------------------------------|---|---|--|
| Beasley (1995) | Departments in different universities | Academic personnel budget, other expenses budget, research income | No of undergraduates, no. of taught postgraduates, no. of research postgraduates, research income, department rating |
| Hanke and Leopoldseder (1998) | Different universities | No. of students, academic personnel budget, other expenses budget | No of graduates, no. of courses held, no. of publications |
| Johnes (2006a) | Different universities (economics only) | Mean of score in A-level exams, gender, type of secondary school attended | Standings of graduates |
| Bobe (2009) | Faculties within single university | Academic personnel expenses, other expenses | Total teaching load, no. of publications, research grants |
| Tzeremes and Halkos (2010) | Departments within single university | No. of academic personnel, no. of students (incl. post-graduates), total budget funds | No. of graduates (post-graduates weighed more), no. of publications ^a |
| Alwaddood et al. (2011) | Departments within single university | Department utilization ^b , no. of courses offered | Quality of graduates ^c , quality of research ^d , quality of staff service ^e |

^aPapers in foreign journals were given more weight than conference and discussion papers.

^bDepartment utilisation includes faculty to student ratio, mean student credit hours and the number of graduate students.

^cThe quality of graduates takes into account the percentage of employers surveyed who are satisfied with the work of the graduate, the percentage of graduates surveyed who are themselves satisfied with their current position along with the GPA of graduate students.

^dThe quality of research includes the number of publications and conference papers, the number of research grants and the number of supervised graduate students.

^eThe quality of staff takes into account the number of short courses organised for staff, the number of consultancy jobs and the percentage of staff who are taking part in industrial consultancy.

DEA was chosen for this study because unlike traditional PIs which measure quality or effectiveness on an absolute scale, DEA measures *relative efficiency*. That is, the efficiency of each unit is not computed against an ideal level of performance that may never be achievable in practice, but rather against the set of other units in the study. This means that the result set will always contain at least one efficient unit with a 100% relative efficiency score. With DEA, the relative efficiency of each unit is obtained from the ratio of multiple outputs to multiple inputs.

When a unit is relatively efficient, then its output/input ratio is optimal among the examined units, requiring no further increase in any of its output values or decrease in any of its input values. An inefficient unit on the other hand, will have a relative efficiency score of less than 100% and will have to reduce all its input values (while keeping its outputs constant) by a factor equal to the efficiency score so as to become efficient. This is illustrated in detail in the section dealing with results.

2. DEA Studies on the Efficiency of Teaching

Despite the many relative efficiency assessments in higher education, most studies have focused either on an inter-comparison of different universities or in comparing different departments within the same university as attested by the studies compiled in Bobe (2009) and Johnes (2006b). Even though the use of DEA in higher education was mentioned already in the late 1970s by Lindsay (1982), few studies have hitherto been devoted to an assessment of the syllabus or the curriculum

of course units in a university department. This cannot be merely due to the lack of suitable analytical tools for conducting such a study.

Until the late 1990s, the use of DEA within a classroom and teaching context in higher education was still practically non-existent. As Becker (2004) put it, “DEA could be used to determine whether the teacher and/or student exhibits best practises ... unfortunately, no one ... in education research has yet used DEA in a meaningful way for classroom teaching practises”. According to Ekstrand (2006), it is only following the new millennium that higher education efficiency studies shifted their focus on the efficiency of modules or units within a particular university. The stochastic frontier analysis by Ekstrand (2006) made use of a group of 94 students at a Swedish university who took a macroeconomics unit to test the hypothesis whether this macroeconomics unit exhibited any inefficiency. The score on the final exam was used as the output, and three inputs were used to estimate a student’s preparation for the exam. These were: (1) the student’s knowledge of the course unit’s material at the onset of taking the unit, (2) the time spent by the student in studying the material for the unit, and (3) the student’s attendance record for the unit.

To our knowledge, DEA has been used only once in measuring the effectiveness of teaching at the course unit level in the concise study by Sarkis and Seol (2010). However, the study did not compare the effectiveness of different courses, but rather measured the teaching effectiveness of a single instructor using student evaluation data such as the students’ perception of the instructor’s ability to clearly present the teaching material and organise the unit. This study on the other hand, examines several course units and avoids subjective data by relying solely on the academic achievements of the involved students.

2.1. Some Considerations about a DEA Assessment

It was mentioned earlier that DEA measures relative efficiency through a ratio of the outputs to inputs. Actually, this efficiency ratio is a sum of the *weighted outputs* over the sum of the *weighted inputs*. In fact, the good news about DEA is that these weights are the unknowns, which means each unit can put more emphasis on those inputs or outputs in which it has fared better, thus allowing that unit to appear in the best possible light (Boussofiane, Dyson, & Thanassoulis, 1991). By letting the DEA technique determine the weights separately and optimally for each unit under consideration, we are guaranteed that the efficiency score obtained for a particular unit is indeed the best possible score given the selected factors and the set of all examined units.

The not so good news is that there is no guarantee that all factors will be taken into account when computing a unit’s relative efficiency score. In other words, a unit may obtain a high efficiency score at the expense of overlooking certain factors, that is, by assigning zero weights or minimal, near-zero weights to one or more factors. Such a unit’s efficiency score, as Sarrico and Dyson (2000) aptly put it, is due to a “judicious choice of weights rather than good performance”. There is also the concept of inefficient *odd units*, or units that are inefficient but are not homogenous with the rest of the units. An odd unit could be due to incorrect data, which as pointed out by Metters, Vargas, and Whybark (2001), can easily distort the DEA results. The low score of an inefficient unit may thus be in part attributable to random data errors rather than to authentic inefficiency (Barros & Weber, 2009). Finally, an odd inefficient unit may also simply imply that the unit is too different from the rest of the units and cannot be thus accurately rated. This is illustrated in the section entitled “Results”.

To help avoid such situations, this work restricts the number of factors to three, which minimizes the chances that a certain factor is overlooked, and more importantly, it allows for visualising the results in a simple two-dimensional graph as shown later in the article. Being able to visualise the results using simple two-dimensional charts as shown in El-Mahgary and Lahdelma (1995) improves managerial understanding of DEA and avoids the use of DEA extensions such as those mentioned in Emrouznejad, Parker, and Tavares (2008) that are used to restrict the number of factors that can be overlooked, i.e. assigned a near-zero weight.

The reader desiring a lucid introduction to DEA is referred to Boussofiane et al. (1991) and to the pragmatic book by Norman and Stoker (1991), while a glossary of key DEA terms is given in El-Mahgary (1995). A rigorous analysis of DEA in the framework of economic efficiency can be found in Murillo-Zamorano (2004).

The rest of this article is organised as follows: first we discuss the aims of teaching, then focus on what it is exactly that we aim to measure, followed by an in-depth look at how the technique was applied to our sample. Finally, the results are presented along with conclusions.

3. Teaching and Its Assessment

While Ramsden (1991) stresses the difficulties in defining good teaching, an immediate aim of teaching is, of course, to provide the student with a deeper understanding of the contents in a course. However, a poorly planned or carried out curriculum can render learning an inefficient process. Possible negative factors include a difficult subject area, students' or teachers' lack of motivation, poor or out-dated teaching methods or facilities, poorly-written literature or lecture material, or a generally unsatisfactory atmosphere that is not conducive to learning. An immediate consequence of poor teaching is generally poor student performance, leading to weak academic integration, which, as pointed out by Longden (2006), increases the likelihood that the student will eventually interrupt his/her studies.

An interesting Australian study by Taylor (2001) shows how pressure emanating from outside the university (in this case the Ministry of Education) can also force instructors to lower their grading standards without changing any of the teaching methods. As a participant in the study from university X relates: "... the government basically said either pass these [students] or ... get a penalty". In response, the failure rate at university X dropped from 23% to 11% in a single year among first year students (Taylor, 2001).

The characteristics of students taking the course units can also affect overall performance. As mentioned earlier, Johnes (2006a) referred to evidence that females would outperform males in their studies for a university degree. However, because our study focuses on a curriculum of units that are mostly taken by the same set of students, factors such as student gender, age and marital status can be expected to play a minor role.

Typically, study programmes or curriculums contain units that can vary significantly in their degree of difficulty as well as in the way the units are conducted. When one examines several units taught by different instructors, one wonders whether it is, in fact, accurate to make an inter-comparison of such modules. As attested by Young (1993), there are different grading standards not just between college institutions but downright at the unit's instructor level also. There are of course, simple ways in which the instructor can, at his/her discretion, increase the average of the students' grades taking the unit, such as:

- Lowering the requirements for the unit.
- Increasing the number of learning activities other than formal lectures (such as excursions or in-class use of multimedia) and awarding more points for attendance.
- Administering final exams that are easier than expected.

To avoid situations where instructors can improve the scores of PIs without actually intervening in the pedagogy of the course, this study uses a DEA-based indicator to estimate the performance of students in a unit relative to other units of the same curriculum. The use of DEA also avoids the inherent bias that may be present in instruments that are solely based on student evaluation of teaching (SET). Beran and Violato (2005) surveyed some 371,000 student ratings over a three-year period and noticed that with SET, laboratory-type units are rated higher than lectures or tutorials. The comprehensive work by Crumbley, Henry, and Kratchman (2001) unearthed how students may

resent and “punish” those teachers who require more efforts from the students. The study by Brockx, Spooren, and Mortelmans (2011) suggests that there is a correlation between the grade received by the student and how the student evaluates his/her teacher. In light of this, our study seeks to answer the following two fundamental questions using an objective approach:

- (1) How well do the units match the abilities of the students?
- (2) Is a particular unit proving to be too difficult/too easy for students?

If a unit is found to have a weak performance score, then it is likely that students are not truly gaining the required knowledge and skills from the unit, the importance of which was raised by Maingot and Zeghal (2008). It is expected that if a unit is a pre-requisite for a more advanced one, then a successful completion of the pre-requisite unit would be reflected in the student’s performance on the more advanced units. In other words, the teaching resources are incorporated into the study as measures of the overall student performance in the pre-requisite course units. These resources are then expected to reflect the overall student performance for the unit under evaluation. The better students fare in the pre-requisites of the unit under measurement, the better final marks they are expected to achieve in the unit being evaluated. A constant returns to scale is assumed to exist between the inputs (resources) and the output (overall student performance in the unit).

4. Measuring Performance in a Course Unit

We need a way of measuring the performance of students in a single unit that is part of the curriculum, before we can say anything about the suitability of the curriculum. A straightforward way to measure this would be to take the average of the students’ grades in the course unit. This average is referred to here as the *Mean of the Grades in a Unit* for a particular course unit. In this article, we adopt a grading system which is prevalent in Nordic universities, based on a six point scale, where the value “5” denotes outstanding achievement and “0” indicates a failure in a course as summarised in Table 2. The interested reader will find more on the equivalence of grading schemes in a report by Duke University (2011).

To help one identify the areas of the curriculum that are proving difficult to students and thus require immediate attention, we need a simple set of analytical tools that nevertheless provide reliable results. Since a direct comparison of the mean of the grades of two different units, even when taught at the same institution, might lead to erroneous conclusions, we use a more robust performance index, *PCOURSE*, that is based on the *WGPA*. The value for *PCOURSE*, for a student *i* who took course unit *c*, is obtained as follows: (1) subtract the student’s *WGPA* from his/her grade (*G*) for unit *c*, and finally, to avoid negative values, (2) add “5” to the obtained difference. This is summarised in the following simple equation:

$$PCOURSE = G_{ic} - WGPA_i + 5 \tag{1}$$

As “5” is the maximum awarded grade, *PCOURSE* will have the range [0, 10]. The performance index *PCOURSE* reaches the maximum theoretical value of “10” when a student with a very low *WGPA* (we assume a *WGPA* of zero for practical purposes) obtains the best grade, “5”, in a given unit. In such a case, the unit can be considered to be extremely easy with respect to the abilities of the student and there is thus no evidence of difficulties. In the other extreme, when a student with a very

Table 2. The Mappings of the Six-scale Grading to Their Letter Equivalents

| | 5 | 4 | 3 | 2 | 1 | 0 |
|-------------------------|-----------------------|-----------------------|-----------------|-------------------------|---------------------|-----------------------|
| Letter grade equivalent | A excellent (91–100%) | A– very good (81–90%) | B good (71–80%) | C satisfactory (61–70%) | D adequate (60–52%) | F failure (below 52%) |

Table 3. An Example in Calculating the Performance of Some Imaginary Course Units

| Student ID | Course C ₁ grade | Course C ₂ grade | Course C ₃ grade | Student's WGPA | PCOURSE for course C ₁ | PCOURSE for course C ₂ | PCOURSE for course C ₃ |
|------------|-----------------------------|-----------------------------|-----------------------------|----------------|-----------------------------------|-----------------------------------|-----------------------------------|
| 901 | 3 | 2 | 1 | 2.3 | 0.7+5=5.7 | -0.3+5=4.7 | -1.3+5=3.7 |
| 952 | 5 | 3 | 2 | 3.3 | 1.7+5=6.7 | -0.3+5=4.7 | -1.3+5=3.7 |
| 973 | 5 | 4 | 5 | 4.1 | 0.9+5=5.9 | -0.1+5=4.9 | 0.9+5=5.9 |
| Mean | 4.3 | 3.3 | 2.7 | 3.7 | 6.1 | 4.7 | 3.7 |

high WGPA (now we assume a WGPA of “5”) obtains the lowest possible grade, “0” in a given unit, PCOURSE will be zero, indicating maximum difficulty in that particular course unit.

The use of PCOURSE can be illustrated with a simple hypothetical example for three course units C₁–C₃ (each worth 4 ECTS) taken by only three students as shown in Table 3. The general academic abilities of students are reflected in their WGPA measure. In this case, student “973” is considered to have the best abilities of the group (WGPA=4.1). Note that in the last row, the mean of the grades for C₂=3.3, suggesting that students generally performed well in unit C₂. However, its PCOURSE=4.7 which is below average (below 5), indicating a small level of performance difficulty in course unit C₂.

In practice, however, we take the mean of PCOURSE of all students who took the course unit during a certain period. The mean is denoted with PCOURSE* and describes the level of difficulty for the whole unit. The measure PCOURSE* can be additionally improved by taking into account the number of credits awarded for each unit. Assuming that the average credits awarded for a unit is 4 ECTS, and that larger course units are on the whole, harder for the student, as typically there is a wider range of material to be covered, the following weighting scheme can be used:

$$WPCOURSE^* = PCOURSE^* \times k \tag{2}$$

where for units with ECTS values of 6 and 5, $k=1.05$ and $k=1.025$ respectively, and for ECTS values of 3 and 2, $k=0.975$ and $k=0.95$ respectively. A course that awards 4 ECTS will have $k=1$. The number of ECTS points seemed as the most straightforward way to adjust the variable k .

4.1. The Factors Used

As the introduced measure WPCOURSE* is about the students’ achievement in a unit, it is clearly an output. Additionally, we will need two inputs so that the total number of factors amount to three. With DEA, inputs are factors that get transformed into outputs by undergoing a transformation process that is usually value-adding. In this case, the added value depends on how well students are prepared for taking a particular unit. Prior to taking a certain unit, students should have completed certain pre-requisite units to help them get the most out of teaching and succeed in the unit in question. The better the students complete their pre-requisites, the better the results that can be expected. To approximate how well students are prepared for taking a certain unit, two measures will be used. The pre-requisites for each of the examined units are listed in Table 8 in the Appendix.

The first measure relating to the pre-requisites, called PCMATHS, computes the percentage of students who have completed the required mathematics pre-requisites on time, that is, before taking the unit. The other measure, PTGRADE3, computes the percentage of students who have obtained the grade of “3” (“good”) or better in the pre-requisites.

Since these two pre-requisite measures relate to the use of resources, they can be classified as inputs. The factor WPCOURSE* on the other hand, is clearly an output since it measures student per-

Table 4. The Two Inputs and the Output for the Efficiency Study

| Factor | Type | |
|-----------------------|--------|---|
| WPCOURSE ¹ | Output | Level of easiness for course unit (10 = easiest) |
| PCMATHS | Input | Percentage completing math pre-requisites before taking unit |
| PTGRADE3 | Input | Percentage who obtained grade '3' or better in pre-requisites |

formance. These three factors (summarised in Table 4) will be used to estimate the teaching efficiency of units.

5. The Data Set

To conduct this study, we selected a set of course units from the Department of Surveying at Aalto University's School of Engineering that are part of the curriculum for Photogrammetry. Only courses taken by degree students admitted between 2003 and 2010 were considered and units with a class size of less than ten students were excluded. Non-degree students were omitted from the study because they typically do not complete any specific curriculum. All in all, the sample set consisted of 343 students of which nearly a third (31.5%) were female as shown in Table 5. The coefficient of variation was also computed due to its usefulness as pointed out by Mahmoudvand and Hassani (2009), and it showed little difference in the relative dispersion between females ($cv=20\%$) and males ($cv=22\%$).

In passing, we note that we also tested the premise that females outperform males in their academic results. This was done using the GPA of another sample (not shown in Table 5) that consisted of only those students who had finished their studies and graduated with a major in Geodesy and Photogrammetry or in Real Estate Economics and who had been admitted between 2003 and 2010, for a total of 276 students. There were 112 female graduates with a mean GPA of 3.34 ($skew=-0.16$, $kurtosis=-0.53$) and 164 male graduates whose mean GPA was 3.14 ($skew=0.29$, $kurtosis=-0.21$). Normality was checked in each test group through plotting and confirmed with both a Shapiro-Wilk test and a Kolmogorov-Smirnov test. A 2-tailed independent parametric-test supported the hypothesis that the GPA for female students indeed outperforms that of male students ($t=3.32$ and $p<0.005$, Cohen's $d=0.40$).

A total of 12 course units were examined, denoted with the abbreviations GED-101A, GED-101B, GED-101AB, GED-102, GED-103, GED-204, GED-205, GED-206, GED-310, GED-310*, GED-311 and GED-318 as shown in Table 6. For comparison purposes, a mandatory module, GED-101AB, was split into two parts, with unit GED-101A consisting of Geodesy and Photogrammetry majors and unit GED-101B consisting only of students majoring in Real Estate Economics.

Finally, unit GED-310 was included in two forms: the basic unit GED-310, which included students who had passed an older, but nearly equivalent course GED-220, and a separate unit GED-310* which included only students who had taken a new revised GED-310 unit.

These 12 units generated a total of 697 records excluding any units taken as a pre-requisite or as a substitute. Each record was made up of the student's id, the course id, the date when the course unit was passed, the name of the instructor, as well as the grade obtained. In addition to these 697 records

Table 5. Descriptive Statistics for the 343 Students Who Took the Modules

| Student count (all students: mean GPA, its standard deviation and coefficient of variation) | Female student count (female students only: mean GPA, its standard deviation and coefficient of variation) | Male student count (male students only: mean GPA, its standard deviation and coefficient of variation) |
|---|--|--|
| $n=343$ (mean = 2.89, std. dev. = 0.63, $cv=22\%$) | 108 (mean = 3.05, std. dev. = 0.60, $cv=20\%$) | 235 (mean = 2.81, std. dev. = 0.63, $cv=22\%$) |

Table 6. The Data along with the Obtained Efficiency

| Unit | PCMATHS (%) | PTGRADE3 (%) | WPCOURSE' | Teaching efficacy (efficiency, %) |
|-----------|-------------|--------------|-----------|-----------------------------------|
| GED-101A | 56.6 | 63.2 | 4.98 | 60.8 |
| GED-101B | 69.3 | 77.8 | 4.46 | 44.4 |
| GED-101AB | 64.3 | 72.3 | 4.70 | 50.4 |
| GED-102 | 53.3 | 63.0 | 5.21 | 66.8 |
| GED-103 | 64.9 | 63.5 | 4.90 | 53.3 |
| GED-204 | 53.1 | 58.6 | 3.41 | 44.4 |
| GED-205 | 53.3 | 58.6 | 4.97 | 64.5 |
| GED-206 | 46.0 | 60.6 | 4.26 | 62.1 |
| GED-310 | 42.1 | 78.6 | 6.81 | 100 |
| GED-310' | 48.5 | 82.1 | 6.18 | 80.7 |
| GED-311 | 38.5 | 23.1 | 5.12 | 100 |
| GED-318 | 34.8 | 60.0 | 5.0 | 90.7 |
| Mean | 52.1 | 63.4 | 5.0 | 68.2 |

making up the core curriculum units, the data also contained the pre-requisites for each unit for determining whether a student had completed the necessary pre-requisites on time. The data-set was collected from the university's student database through a special report generator, (El-Mahgary & Soisalon-Soininen, 2007), and was then filtered and re-arranged into a suitable spreadsheet format. In Table 6, the values for the inputs are shown in the second and third columns and the output *WPCOURSE** is shown in the fourth column. To compute the value of *WPCOURSE** for each of the 12 units, the following procedure was used: first, using Equation 1, the value for *PCOURSE* was computed for each student who took the unit in question. The mean of the values, *PCOURSE**, was then converted into *WPCOURSE** using Equation 2 and the ECTS of the unit. The values for the inputs *PCMATHS* and *PTGRADE3* for each unit were computed as explained previously using data for the pre-requisites of the unit question.

It might appear that since we are examining different units taught by different instructors, these units might not be homogenous, thus increasing the risk for introducing bias into results. However, it should be borne in mind that we are not so much contrasting different course units as measuring how a particular set of students can assimilate the contents of different units that are taught in a single curriculum. Even though different course units are being measured, the set of students taking the units is basically the same, and since nearly all students who took a particular unit took it with the same instructor, we believe that homogeneity is maintained. This issue will be addressed further while examining the results.

6. Results

A DEA analysis was performed on the data sample of 343 students for whom the values of the three factors (*PTGRADE3*, *PCMATHS*, *WPCOURSE'*) were computed and the results are reported in the last column of Table 6. While special DEA software is generally used for finding the set of optimal weights for each factor and calculating the efficiency scores, an accurately drawn two-dimensional chart can be used to get accurate enough values for the efficiency scores as will be shown next.

Since the factors are made up of a single output and two inputs, we are going to draw the two-dimensional graph by using ratios of inputs over the output. In this way, the single output acts as a denominator for both ratios. Equation 3 gives for each unit *c*, its *x*-coordinate through the ratio of the unit's value for input *PCMATHS* over the value of the output *WPCOURSE'*. Similarly, equation (4) gives for each unit *c*, its *y*-coordinate through the ratio of the unit's value for the input *PTGRADE3* over the value of the output *WPCOURSE'*. The result is multiplied in both equations by 100 for scaling purposes.

Using the data from Table 6, each course unit *c* has been plotted into the graph in Figure 1 by computing its (*x_c*, *y_c*) coordinates according to Equations 3 and 4 and then multiplying the obtained

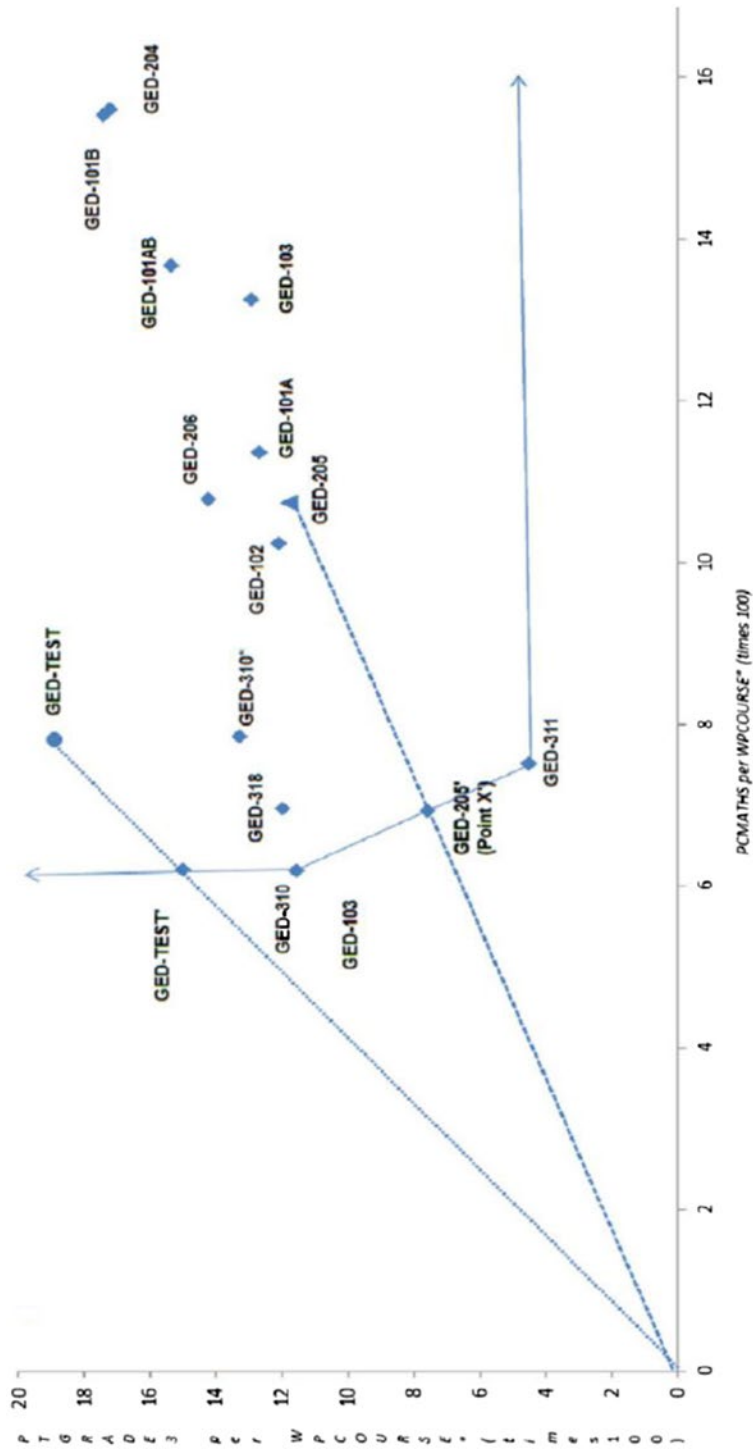


Figure 1. Visualization the performance of the twelve units.

value by 100. So for instance, the coordinates for unit GED-205 are obtained through the following: $x_c = (53.3\%/4.97) \times 100$, $y_c = (58.6\%/4.97) \times 100$ thus placing the unit into the coordinates (10.72, 11.79) in Figure 1.

Because we are using the ratio of input(s) over output(s), the efficient units are going to be the ones lying nearest to the origin, since they will have the smallest ratio either for x_c and/or for y_c . In other words, the smaller these two ratios for a given unit are, the less resources will have been used up per output in the given unit, and hence, the higher the efficiency (or teaching efficacy) score of the unit must be.

$$x_c = (PCMATHS_c / WPCOURSE_c^*) \times 100 \quad (3)$$

$$y_c = (PTGRADE3_c / WPCOURSE3_c^*) \times 100 \quad (4)$$

As can be seen from Figure 1, there are two units, FED-310 and GED-311, that can be said to lie nearest to the origin. These units are thus relatively efficient and will have an efficiency score of 100%. It is no surprise that the number of efficient units turns out to be two, for as explained by Bessent, Bessent, Elam, and Clark (1988), when there are three factors as in our case, there should be at least two efficient units, because one of the efficient units (unit GED-310 in this case) has the best ratio for x_c while the other unit (GED-311 in this case) has the best ratio for y_c . Stated differently, unit GED-310 shows the best performance for students while taking into account their completion of their math pre-requisites, and unit GED-311 has the best performance for students with respect to how well they completed their pre-requisites for the unit in question. It should be noted that unit GED-311 is relatively efficient mainly due to the low values in its inputs (Table 6), that is, students come to the unit with poor standings in their pre-requisites, yet as they manage to obtain a relatively high score in the unit itself ($WPCOURSE_c = 5.12$), there is no evidence of teaching resources being misused.

These two efficient units are then connected by drawing a piecewise linear connection between them. Moreover, it is customary to extend the line segment connecting the two efficient units through two lines that extend in parallel along each of the two axes. That is, we draw from unit GED-310 a line that extends in parallel along the y -axis and from GED-311 a line that extends in parallel along the x -axis. These extensions parallel to the axes underline the fact that there are no further efficient units beyond units GED-310 and GED-311 and so an extension along the same level is a basic and arguably fair way of extrapolation. These line segments along with the extensions that have been just drawn constitute what is known as the *efficiency frontier*. The efficiency frontier for our data-set based on twelve units and three factors is portrayed in Figure 1. The term DEA is due to the fact that this efficiency frontier envelopes all the inefficient units within since only units lying on the frontier itself (units GED-310 and GED-311) are relatively efficient. The further a unit lies from the frontier, the greater that unit's relative inefficiency is (implicating a lower efficiency score) since the more it will have consumed inputs per output. The units with the lowest efficiency scores must therefore be units GED-101B and GED-204, the ones lying furthest from the efficiency frontier.

Given then that with three factors we need just two efficient units to determine the efficiency frontier, there will only be one so-called *reference set* (also known as a *peer group*). A reference set contains only 100% relatively efficient units and it determines at least a part of the efficiency frontier. In this case, the reference set is the set {GED-310, GED-311} and it makes up the basis of the efficiency frontier. Note that it would be perfectly possible to have more than two efficient units in the same reference set, in fact, any unit located on the segment between GED-310 and GED-311 would be efficient.

We can now use the efficiency frontier to compute the scores for the inefficient units. Consider unit GED-205, which is marked in Figure 1 as an upright triangle and denoted as point X . To become relatively efficient, GED-205 would have to lie on the hypothetical point GED-205' on the frontier, denoted as point X' . This Unit GED-205' is therefore what is known as a *hypothetical composite efficient unit* (HCE unit) for the following two reasons: First, it is a hypothetical unit since it does not actu-

ally exist. Second, it is an efficient composite unit because its output/input values can be obtained by combining the inputs and outputs of its reference set that is, efficient units GED-310 and GED-311. As to how this is done is, however, beyond the scope of this article, the interested reader is referred to Boussofiane et al. (1991). For our purposes, it suffices to note that point GED-205' (point X'), is where the HCE for unit GED-205 is located and represents therefore an efficient point.

Since we denoted the location of unit GED-205 as point X and the location of its HCE unit GED-205' as point X' , the efficiency score for unit GED-205 can be obtained from Figure 1 through the proportion of two distances, that is, from the lengths of segments OX' and OX where OX is a ray emanating from the origin up to point X (the unit GED-205) and OX' corresponds to the ray from the origin up to the point GED-205' as shown in Equation 5:

$$\text{Efficiency (GED-205)} = OX' / OX \times 100\% = 64.5\% \quad (5)$$

Returning to units GED-101B and GED-204, we see from Figure 1 that since they are further away from the efficiency frontier than unit GED-205, their efficiency score must therefore be less than 64.5%, for they require a greater increase than unit GED-205 in their output and/or decrease in their inputs before reaching the efficiency frontier and becoming efficient. In general then, inefficient units can be ranked according to their efficiency scores. However, as the scores are relative, it is worth remembering that some of the scores are likely to change in case a new unit that turns out to be efficient is introduced into the set of course units.

Given that unit GED-205 is inefficient, we would like to know what its input values and output value should be in order for it to become efficient, that is, an HCE unit. The reduced input value(s) and augmented output value(s) that render an inefficient unit into an efficient one are known as *target values*. To determine these target values, what is needed then are the actual coordinates of point X' , the point where HCE unit GED-205' is located.

One pragmatic way of determining these target values is as follows. First, notice that any point lying on the line OX' will have a constant ratio of $PTGRADE3$ over $PCMATHS$. This constant ratio is easily obtained from the inputs for unit GED-205 to be $58.6\%/53.3\% = 1.0994$ and is defined as Equation 6. Equation 6 basically states that ratio of the target values of the inputs for point X' is a constant. In fact, this constant value 1.0994 turns out to be the slope of the line OX' and since it is nearly unity, it means that in order to increase the output, it is practically speaking as easy (or as difficult, depending on how one puts it) to decrease the input $PTGRADE3$ as it is to decrease the other input $PCMATHS$. If the slope were clearly larger than unity though, this would imply that to achieve a given increase in the output requires a greater proportionate reduction in the input $PTGRADE3$ than in the input $PCMATHS$.

$$\text{Ratio (PTGRADE3 over PCMATHS)} = 1.0994 = T_{PTGRADE3} / T_{PCMATHS} \quad (6)$$

As for the ratios of the x-coordinate (x'_{205}) and y-coordinate (y'_{205}) of the target point X' , instead of trying to read these values directly from Figure 1, one approach would be to find the intersection of line OX' and the line segment between GED-310 and GED-311. There is, however, another simple and accurate way for obtaining these ratios. We can make use of the fact that with DEA, an inefficient unit that is surrounded (or *enveloped* as the term is known) by efficient units as is the case with unit GED-205, can be rendered efficient through its efficiency score. Efficiency for the unit can be then achieved through a proportional reduction in its inputs while keeping its output(s) values constant. So we will use Equations 3 and 4 to find x'_{205} and y'_{205} respectively, but instead of using directly unit GED-205's value for $PCMATHS$ and $PTGRADE3$, we will use the corresponding values that are reduced by a factor equal to the unit's efficiency score of 64.5%. Thus for $x'_{205} = ((53.3\% \times 64.5\%) / 4.97) \times 100 = 6.92$ and for $y'_{205} = ((58.6\% \times 64.5\%) / 4.97) \times 100 = 7.61$. So the coordinates of point X' (the

HCE unit of GED-205) are (6.92, 7.61) where the X' -coordinate is the ratio of *PCMATHS* over *WPCOURSE** at point X' and the y' -coordinate expresses the ratio of *PTGRADE3* over *WPCOURSE** at point X' . Expressed mathematically, this yields Equations 7 and 8 as follows where for unit GED-205, $T_{PCMATHS}$ denotes the target value for *PCMATHS*, $T_{WPCOURSE*}$ the target value for *WPCOURSE** and $T_{PTGRADE3}$ the target value for *PTGRADE3*.

$$\text{Ratio } (T_{PCMATHS} \text{ over } T_{WPCOURSE*}) = 6.92 = (T_{PCMATHS} / T_{WPCOURSE*}) \times 100 \tag{7}$$

$$\text{Ratio } (T_{PTGRADE3} \text{ over } T_{WPCOURSE*}) = 7.61 = (T_{PTGRADE3} / T_{WPCOURSE*}) \times 100 \tag{8}$$

The three Equations 6–8 that represent three factors are based on efficient point X' and so allow us to determine the target values for unit GED-205. In other words, these three equations are set up for unit GED-205 to become efficient and express the requirements for the target values that will specifically render unit GED-205 efficient. If, for instance, we decide that students who take unit GED-205 should be able to raise their pre-requisites *PTGRADE3* up to 60%, then by substituting 60% for $T_{PTGRADE3}$ into Equation 6 and solving, we find that $T_{PCMATHS}$ must be 54.6%. Finally, substituting $T_{PCMATHS} = 54.6\%$ into Equation 7 will yield $T_{WPCOURSE*}$ to be 7.89. This means that the set of target values $T_{PCMATHS} = 54.6\%$, $T_{PTGRADE3} = 60\%$ and $T_{WPCOURSE*} = 7.89$ is an example of target values that render unit GED-205 efficient.

Earlier, it was mentioned that an inefficient unit can be made efficient by reducing its input values proportionately as indicated by its efficiency score while keeping the output(s) constant. Alternatively, an inefficient unit can be made efficient through an increase in its output(s) values as indicated by the inverse of its efficiency score while keeping the input(s) values constant. Since the inputs used here are not easily reduced in practice as they reflect how well a student has completed the pre-requisite course units for the unit under measure, it is more useful to find out by how much the output factor needs to be increased. So unit GED-205 can become efficient by increasing its output by a factor equal to the inverse of its efficiency, that is, $1/0.645 = 1.55$ while keeping both its inputs constant. Since the product of 4.97 and 1.55 yields 7.7, then unit GED-205 would be also considered efficient if its input values were kept constant at 53.3% (for *PCMATHS*) and 58.6% (*PTGRADE3*) while its output value for *WPCOURSE** were raised all the way up to 7.7.

Thus, we have uncovered mathematically another set of target values $T_{PCMATHS} = 53.3\%$, $T_{PTGRADE3} = 58.6\%$ and $T_{WPCOURSE*} = 7.7$ to render unit GED-205 efficient. This is in contrast with the previous set of target values ($T_{PCMATHS} = 54.6\%$, $T_{PTGRADE3} = 60\%$ and $T_{WPCOURSE*} = 7.89$), which consumes more inputs by having higher pre-requisite values for *PCMATHS* (54.6% as opposed to 53.3%) and *PTGRADE3* (60% as opposed to 58.6%) and therefore correspondingly requires a higher value for the output *WPCOURSE**. Finally, we can use Equations 6–8 as a check for the newly obtained target values. Substituting 7.7 for the value of *WPCOURSE** into Equation 8 gives $T_{PTGRADE3} = 58.6\%$ which, when substituted into Equation 6 yields $T_{PCMATHS} = 53.3\%$, thus validating the obtained target values. These two sets of target values are shown in Table 7. Both sets of target values naturally result in the same HCE unit, that is unit GED-205', the one denoted as point X' in Figure 1.

Table 7. Two Different Sets of Target Values that Both Render Inefficient Unit GED-205 as Relatively Efficient

| HCE unit | Target for <i>PCMATHS</i> (%) | Target for <i>PTGRADE3</i> (%) | Target for <i>WPCOURSE*</i> | Coordinates of HCE unit for GED-205 | Relative efficiency (teaching efficacy, %) |
|----------|-------------------------------|--------------------------------|-----------------------------|-------------------------------------|--|
| GED-205' | 54.6 | 60 | 7.89 | (7.61, 6.92) | 100 |
| GED-205' | 53.3 | 58.6 | 7.7 | (7.61, 6.92) | 100 |

In addition, the efficiency frontier can be used to identify inefficient odd units. Consider unit GED-TEST, which is a totally imaginary inefficient test unit that has been added to Figure 1 for explanatory purposes only. Its HCE unit is GED-TEST', which lies outside the efficient segment determined by the reference set {GED-310, GED-311}. Unit GED-TEST is not only inefficient, but is also known as a *non-enveloped unit*. Unit GED-TEST is not properly enveloped because GED-TEST' (its HCE unit), lies on an extension of the frontier and does not fall within the proper efficiency frontier segment defined by GED-310 and GED-311, and thus GED-TEST' is enveloped only from its right side by unit GED-311, there being no efficient unit on the left side of GED-TEST'. As unit GED-TEST has only one efficient unit in its reference set {GED-311}, there is thus a strong possibility that unit GED-TEST is indeed different from the other units, and the reasons for this should be analysed. Of course, the oddity in GED-TEST might also simply be due to incorrect data. It might be useful to note that GED-TEST' is not considered by DEA to be efficient because it lies on an extension of the frontier. Unit GED-TEST' is likely to get a high efficiency score (perhaps in the neighbourhood of 0.99), but its efficiency score will remain less than unity.

Ignoring the imaginary unit GED-TEST, we can say that among the actual units in the study, there is no evidence of an inefficient odd unit, as each inefficient unit has the same reference set that consists of two efficient units as required. Figure 1 is therefore valuable also in the sense that it points out inefficient units that are potentially not homogenous with the rest of the units. Detecting inefficient odd units without a visualizing graph is tricky, because a zero weight associated with an input/output is not always a sign of a non-enveloped unit, as pointed out by Portela and Thanassoulis (2006). The visualized efficiency frontier thus serves a dual purpose: not only does it provide a quick way of detecting a unit's efficiency, it also helps in detecting inefficient units that might be odd. Furthermore, the efficiency frontier can also pinpoint efficient odd units: if GED-TEST were located instead at the location of GED-TEST', it would be an example of an efficient unit that is very different from the rest of the units. Since unit GED-TEST would not act as a reference set to any inefficient unit, it would however, have no effect on the efficiency scores of other inefficient units.

6.1. Applying the Results to a Course Unit

It is safe to say that courses that obtain a relative efficiency score of 100% (such as GED-310 and GED-311) or nearly so, show no evidence of a poor return on allocation of teaching resources. It is also unlikely that such units are proving too difficult for the students. Assuming there is a high *PCOURSE'* value associated with such a course, then on the average, students must be performing better than expected based on their GPA, which is evidence for efficient use of teaching resources. Such high scores might also suggest that the course is not challenging enough for the students, but this is likely to be the case only if students with a low GPA are clearly performing as well as those with a higher GPA, which was not the case in this study.

From the two units that had a clearly weaker efficiency score, we selected course unit GED-101AB, which had the largest amount of participants for an in-depth analysis. The unit with the weakest efficiency score, GED-101B, differs from unit GED-101AB only in its student makeup: GED-101B is made up of Real Estate majors, but the contents for both units is the same. Together with the help of a study counsellor, and discussions with students, we concluded that students were indeed experiencing difficulties with the unit. Therefore, a thorough revamping of the module was made after it was discovered that lack of motivation for the unit's subject was one of the reasons for poor student performance.

As to effective teaching methods, the comprehensive study by Bjorklund and Fortenberry (2005) geared at engineering students identified some helpful guidelines. Among them were: (1) to encourage student-instructor interaction, (2) to develop a sense of reciprocity and mutual cooperation among the students (as opposed to rivalry), (3) to make it clear to students that there are high expectations for them, (4) to give students feedback promptly, and (5) to promote so-called "active learning techniques". The idea is thus to create an atmosphere where students are guided through difficult concepts and have someone to turn to (either the instructor or their peers) when they need

help in understanding key concepts in engineering and problem solving. In light of this, an alternative way to complete the course unit was also arranged. Instead of having the students sit for a final exam which solely determined their grade for a particular unit, new, extended tutorials were introduced. These tutorials covered all the core material and gave students the experience of putting theory into practice while improving student/teacher interaction.

This learn-by-doing method proved successful as students came to the units that are taken after GED-101AB better prepared. Replacing an exam with written assignments and class discussion is supported in a study of university students majoring in educational psychology by Tynjälä (1997), where two different groups of students completed the same course unit in one of two ways. The first group was a control group that followed traditional teaching methods that involved taking an exam while the second group was a “constructive learning” group that completed the unit through written assignments, an extensive essay and numerous in-class discussions. In the study, the constructive group of students exhibited more critical thinking as opposed to rote learning and memorization of facts.

7. Conclusions

DEA is useful because it is not based on regression or some imaginary standard of performance that may never be achieved in practice. Instead, using the input and output values of all units under measurement, an efficiency frontier (that is the equivalent of a production function) is drawn. Based on this frontier, the teaching efficacy of each unit can then be measured. Moreover, DEA can easily incorporate hundreds of course units as long as they are homogenous, that is, taught at the same institution and belong to the same curriculum. When only three factors are used as in this study, it becomes possible to illustrate and check the efficiency scores in a visual graph, which should help management to better understand the technique. The graph can also pinpoint anomalies in the results, such as an efficient unit that effectively neglected the importance of one of its inputs or outputs.

We feel that without this DEA analysis, we would have not thought that unit GED-101AB needed our attention, for the student feedback did not reveal any significant problems in the way the lectures or tutorials were being carried out. Neither had a core contents analysis for the unit (using Bloom’s Taxonomy) indicated any problems. The DEA study gave us a more objective perspective in understanding the sources of the difficulties experienced by students and showed the variations in students’ performances between different units.

At first, the aim of the study was to use DEA to measure teaching efficiency, but that proved too daunting a task as measures such as teacher-student contact time or student satisfaction with the unit would have needed to be included. We also wanted to avoid the use of SET which, according to Ramsden (1991) can be subject to “formidable problems”.

Instead, we opted to conduct a study that showed the relative performance of each unit, that is, the efficient use of teaching resources allocated to the course unit. In essence, the performance scores of each unit act as a guideline so that the higher the efficiency score, the less likely the unit is in need of revision of its contents. As no use of financial resources was made, the teachers’ efforts are reflected in the final scores that are based on the relative student performance in the unit while taking into account the students’ previous performance in the pre-requisites and their general academic performance. As the grade obtained in a unit is not used as such but rather compared to the student’s *WGPA*, errors due to subjective differences in evaluating students are reduced. This addresses the concerns about using examination results as PIs when they may be non-comparable due to different grading criteria (Kong & Fu, 2012). Moreover, two independent studies have confirmed that the problem of non-comparable grades is not apparent in sets of course units taken from the same syllabus or curriculum in a university (Young, 1993).

It would be interesting to repeat this study after a few years or so with a new batch of students. If our revision for the contents of course unit GED-101AB was successful, one would expect that the

relative efficiency for that unit would then be higher. This is built on the assumption that in the new students batch, students with relatively high values for the inputs *PTGRADE3* and *PCMATHS* would correspondingly obtain better marks in the revised unit (i.e. reflected in *WPCOURSE*), even though their abilities may not be significantly higher than the currently examined batch.

A final observation regarding these results is the role played by student motivation and goal orientation. It should be remembered that a gifted student might be obtaining poor scores in a course simply due to a lack of motivation in that particular course subject. As to whether or not that lack of motivation is due to something that needs to be improved in the course (i.e. the course material or the lectures) is harder to detect from the obtained results, since as pointed out by Pulkka and Niemivirta (2013) students can perceive the course material differently depending on how goal oriented they are. In some cases a student evaluation of the unit, such as a FASI-type (Formative Assessment of Instruction) instrument, as detailed in Adams and Wieman (2011) may shed light on the reasons behind a poor score.

If a university aims to maintain a standard of excellence in its teaching, then there must be continuous curriculum monitoring that allows for wise distribution of education resources as well as an on-going development of the course units, both in their contents and in their teaching methods. The study presented herein can measure the relative efficiency of course units with relative ease, requiring mainly access to the student information database and suitable software to pre-process the data. Since time and personnel resources for developing teaching are typically limited, the results obtained can be used to focus the development efforts onto those units that most urgently need attention.

Acknowledgements

We gratefully acknowledge the kind assistance provided by Study Counsellor Ms Päivi Kauppinen at the School for Engineering, Aalto University for her many helpful suggestions and other support. The helpful comments by the anonymous referee are also deeply appreciated.

Author details

Sami El-Mahgary¹
E-mail: sami.mahgary@aalto.fi
Petri Rönholm²
E-mail: petri.ronholm@aalto.fi
Hannu Hyyppä²
E-mail: hannu.hyyppa@metropolia.fi
Henrik Haggren²
E-mail: henrik.haggren@aalto.fi
Jenni Koponen³
E-mail: jenni.i.koponen@aalto.fi

¹ Department of Computer Science and Engineering, Aalto University, Room A145, P.O. Box 15400, Aalto 00076, Finland.

² Department of Surveying, School of Engineering, Aalto University, P.O. Box 11200, Aalto 00076, Finland.

³ Strategic Support for Research and Education, Educational Development Team, Aalto University, P.O. Box 19205, Aalto 00076, Finland.

Article Information

Cite this article as: Evaluating the performance of university course units using data envelopment analysis, S. El-Mahgary P. Rönholm, H. Hyyppä, H. Haggren & J. Koponen, *Cogent Economics & Finance* (2014), 2: 918856.

Cover image

Source: Author.

References

Adams, W. K., & Wieman, C. E. (2011). Development and validation of instruments to measure learning of expert-like thinking. *International Journal of Science Education*, 33, 1289–1312.
<http://dx.doi.org/10.1080/09500693.2010.512369>

Alwadood, Z., Noor, N. M., & Kamarudin, M. F. (2011, September 25–28). *Performance measure of academic departments using data envelopment analysis*. Paper presented at the 2011 IEEE Symposium on Business Engineering and Industrial Applications, Langkawi, Malaysia.

Barnetson, B., & Cutright, M. (2000). Performance indicators as conceptual technologies. *Higher Education*, 40, 277–292.
<http://dx.doi.org/10.1023/A:1004066415147>

Barros, C. P., & Weber, L. W. (2009). Productivity growth and biased technological change in UK airports. *Transportation Research Part E: Logistics and Transportation Review*, 45, 642–653.

Beasley, J. E. M. (1995). Determining teaching and research efficiencies. *Journal of the Operational Research Society*, 46, 441–452.
<http://dx.doi.org/10.1057/jors.1995.63>

Becker, W. E. (2004). Quantitative research on teaching methods in tertiary education. In W. E. Becker & M. L. Andrews (Eds.), *The scholarship of teaching and learning in higher education* (pp. 265–309). Bloomington: Indiana University Press.

Beran, T., & Violato, C. (2005). Ratings of university teacher instruction: How much do student and course characteristics really matter? *Assessment & Evaluation in Higher Education*, 30, 593–601.
<http://dx.doi.org/10.1080/02602930500260688>

Bessent, A., Bessent, W., Elam, J., & Clark, T. (1988). Efficiency frontier determination by constrained facet analysis. *Operations Research*, 36, 785–796.
<http://dx.doi.org/10.1287/opre.36.5.785>

Biggeri, L., & Bini, M. (2001). Evaluation at university and state level in Italy: Need for a system of evaluation and indicators. *Tertiary Education and Management*, 7, 149–162.
<http://dx.doi.org/10.1080/13583883.2001.9967048>

Bjorklund, S., & Fortenberry, N. L. (2005). *Final report: Measuring student and faculty engagement in engineering education* (CASEE Report 5902001-2005-0705). Washington, DC: The National Academy of Engineering.

Bobe, B. (2009, July 5–7). *Evaluating the efficiencies of university faculties: Adjusted data envelopment analysis*. Paper presented at the Accounting and Finance

- Association of Australia and New Zealand (AFAANZ) Conference, Adelaide, Australia.
- Boussofiene, A., Dyson, R. G., & Thanassoulis, E. (1991). Applied data envelopment analysis. *European Journal of Operational Research*, 52, 1–15.
[http://dx.doi.org/10.1016/0377-2217\(91\)90331-O](http://dx.doi.org/10.1016/0377-2217(91)90331-O)
- Brockx, B., Spooren, P., & Mortelmans, D. (2011). Taking the grading leniency story to the edge. The influence of student, teacher and course characteristics on student evaluations of teaching in higher education. *Educational Assessment, Evaluation and Accountability*, 23, 289–306.
<http://dx.doi.org/10.1007/s11092-011-9126-2>
- Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429–444.
[http://dx.doi.org/10.1016/0377-2217\(78\)90138-8](http://dx.doi.org/10.1016/0377-2217(78)90138-8)
- Crumbley, L., Henry, B. K., & Kratchman, S. H. (2001). Student's perceptions of the evaluation of college teaching. *Quality Assurance in Education*, 9, 197–207.
<http://dx.doi.org/10.1108/EUM000000000006158>
- Duke University. (2011). *International credentials guide 2011–2012*. Durham, NC: Graduate School, Durham University.
- Ekstrand, J. (2006, November 2–3). *Measuring and explaining economic students learning efficiency*. Paper presented at the 25th Arne Ryde Symposium, Lund University, Sweden.
- El-Mahgary, S. (1995). Data envelopment analysis. *OR Insight*, 8, 15–22. <http://dx.doi.org/10.1057/ori.1995.21>
- El-Mahgary, S., & Lahdelma, R. (1995). Data envelopment analysis: Visualizing the results. *European Journal of Operational Research*, 83, 700–710.
[http://dx.doi.org/10.1016/0377-2217\(94\)00303-T](http://dx.doi.org/10.1016/0377-2217(94)00303-T)
- El-Mahgary, S., & Soisalon-Soininen, S. (2007, September 3–7). *A two-phased visual query interface for relational databases*. Paper presented at the 18th International Conference on Database and Expert System Applications (DEXA), Regensburg, Germany.
- Emrouznejad, A., Parker, B., & Tavares, G. (2008). Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. *Socio-Economic Planning Sciences*, 42, 151–157. <http://dx.doi.org/10.1016/j.seps.2007.07.002>
- Hanke, M., & Leopoldseder, T. (1998). Comparing the efficiency of Austrian universities: A data envelopment analysis application. *Tertiary Education and Management*, 4, 191–197.
- Harvey, L. (1999, May). *Evaluating the evaluators*. Opening keynote at the Fifth Biennial Conference of the International Network of Quality Assurance Agencies in Higher Education, Santiago, Chile.
- Johnes, J. (2006a). Measuring teaching efficiency in higher education: An application of data envelopment analysis to economics graduates from UK universities 1993. *European Journal of Operational Research*, 174, 443–456.
<http://dx.doi.org/10.1016/j.ejor.2005.02.044>
- Johnes, J. (2006b). Measuring efficiency: A comparison of multilevel modelling and data envelopment analysis in the context of higher education. *Bulletin of Economic Research*, 58, 75–104.
<http://dx.doi.org/10.1111/boer.2006.58.issue-2>
- Köksal, G., & Nalçacı, B. (2006). The relative efficiency of departments at a Turkish engineering college: A data envelopment analysis. *Higher Education*, 51, 173–189.
<http://dx.doi.org/10.1007/s10734-004-6380-y>
- Kong, W.-H., & Fu, T.-T. (2012). Assessing the performance of business colleges in Taiwan using data envelopment analysis and student based value-added performance indicators. *Omega*, 40, 541–549.
<http://dx.doi.org/10.1016/j.omega.2011.10.004>
- Lindsay, A. (1982). Institutional performance in higher education: The efficiency dimension. *Review of Educational Research*, 52, 175–199.
<http://dx.doi.org/10.3102/00346543052002175>
- Longden, B. (2006). Interpreting student early departure from higher education through the lens of cultural capital. *Tertiary Education and Management*, 10, 121–138.
- Mahmoudvand, R., & Hassani, H. (2009). Two new confidence intervals for the coefficient of variation in a normal distribution. *Journal of Applied Statistics*, 36, 429–442. <http://dx.doi.org/10.1080/02664760802474249>
- Maingot, M., & Zeghal, D. (2008). An analysis of voluntary disclosure of performance indicators by Canadian universities. *Tertiary Education and Management*, 14, 269–283.
<http://dx.doi.org/10.1080/13583880802481666>
- Martín, E. (2007). 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, 57–79.
- Metters, R. D., Vargas, V. A., & Whybark, D. C. (2001). An investigation of the sensitivity of DEA to data errors. *Computers & Industrial Engineering*, 41, 163–171.
[http://dx.doi.org/10.1016/S0360-8352\(01\)00050-X](http://dx.doi.org/10.1016/S0360-8352(01)00050-X)
- Mohamad Ishak, M. I., Suhaida, M. S., & Yuzainee, M. Y. (2009, April 14–17). *Performance measurement indicators for academic staff in Malaysia private higher education institutions: A case study in unites*. Paper presented at the Performance Measurement Association Conference, University of Otago, New Zealand.
- Murillo-Zamorano, L. R. (2004). Economic efficiency and frontier techniques. *Journal of Economic Surveys*, 18, 33–77.
<http://dx.doi.org/10.1111/j.1467-6419.2004.00215.x>
- Norman, M., & Stoker, B. (1991). *Data envelopment analysis—The assessment of performance*. New York, NY: Wiley.
- Portela, M. C. A., & Thanassoulis, E. (2006). Zero weights and non-zero slacks: Different solutions to the same problem. *Annals of Operations Research*, 145, 129–147.
<http://dx.doi.org/10.1007/s10479-006-0029-4>
- Pulkka, A., & Niemivirta, M. (2013). Adult students' achievement goal orientations and evaluations of the learning environment: A person-centred longitudinal analysis. *Educational Research and Evaluation*, 19, 297–322. <http://dx.doi.org/10.1080/13803611.2013.767741>
- Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The course experience questionnaire. *Studies in Higher Education*, 16, 129–150.
<http://dx.doi.org/10.1080/03075079112331382944>
- Sarkis, J., & Seol, I. (2010). Course evaluation validation using data envelopment analysis. *The Accounting Educator's Journal*, 20, 21–32.
- Sarrico, C. S., & Dyson, R. G. (2000). Using data envelopment analysis for planning in UK universities—An institutional perspective. *Journal of the Operational Research Society*, 51, 789–800.
- Taylor, J. (2001). The impact of performance indicators on the work of university academics: Evidence from Australian universities. *Higher Education Quarterly*, 55, 42–61. <http://dx.doi.org/10.1111/hequ.2001.55.issue-1>
- Tynjälä, P. (1997). Developing education students' conceptions of the learning process in different learning environments. *Learning and Instruction*, 7, 277–292.
[http://dx.doi.org/10.1016/S0959-4752\(96\)00029-1](http://dx.doi.org/10.1016/S0959-4752(96)00029-1)
- Tzeremes, N., & Halkos, G. (2010). *A DEA approach for measuring university departments' efficiency*. MPRA Paper 24029. University Library of Munich, Germany.
- Young, J. (1993). Grade adjustment methods. *Review of Educational Research*, 63, 151–165.
<http://dx.doi.org/10.3102/00346543063002151>

APPENDIX

The DEA model (multiplier form)

$$\max \theta_k = \sum_{r=1}^s w_r y_{rk} \tag{9a}$$

subject to

$$\sum_{i=1}^m \mu_i X_{ik} = 1 \tag{9b}$$

$$\sum_{r=1}^s w_r y_{rj} - \sum_{i=1}^m \mu_i X_{ij} \leq 0 \quad j=1, \dots, N \tag{9c}$$

$$w_r \geq \varepsilon, \quad \mu_i \geq \varepsilon \quad r=1, \dots, s \quad i=1, \dots, m \tag{9d}$$

Although in DEA efficiency is measured as the ratio of weighed outputs over weighed inputs, in practice, as devised by Charnes et al. (1978) it is the above linear programming that is solved. The model assumes that for a total set of N units, there are m inputs and s outputs. The above model is solved for each unit j to be evaluated. For a unit j , its i th input is denoted by x_{ij} , while its r th output is denoted by y_{rj} . The model is known as the multiplier form as the unknowns are the output weights w_r and the input weights μ_i . The model seeks to determine the values for these weights so as to optimize and maximize the efficiency θ of unit j that is being solved. The constraints from Equations 9b and 9c together imply that the obtained output and input weights must be such that they do not cause the efficiency of any other unit, including the unit j being measured to exceed unity. Additionally, Equation 9d restricts the input and output weights to be positive and greater than an infinitesimal value ε .

Table 8. The Number of Credits Awarded, the Number of Participants, and the Required Pre-requisites for Each Unit

| Unit ID | Credits awarded | Number of students enrolled in unit | Module pre-requisites {(Maths); (GED)} |
|-----------|-----------------|-------------------------------------|--|
| GED-101A | 4 | 122 | {{(MATH-1110); None} |
| GED-101B | 4 | 189 | {{(MATH-1110); None} |
| GED-101AB | 4 | 311 | {{(MATH-1110); None} |
| GED-102 | 3 | 122 | {{(MATH-1110); None} |
| GED-103 | 3 | 77 | {{(MATH-1110); None} |
| GED-204 | 3 | 32 | {{(MATH-1110), (MATH-1131); GED-102} |
| GED-205 | 3 | 30 | {{(MATH-1110), (MATH-1131); GED-103} |
| GED-206 | 4 | 37 | {{(MATH-1110), (MATH-1131), (MATH-1262); GED-101} |
| GED-310 | 6 | 19 | {{(MATH-1110), (MATH-1131), (MATH-1262); GED-205} |
| GED-310' | 6 | 33 | {{(MATH-1110), (MATH-1131), (MATH-1262); GED-205} |
| GED-311 | 6 | 13 | {{(MATH-1110), (MATH-1131), (MATH-1262), (MATH-2104); GED-204} |
| GED-318 | 2 | 23 | {{(MATH-1110), (MATH-1131), (MATH-2104); GED-203} |



© 2014 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 3.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.



***Cogent Economics & Finance* (ISSN: 2332-2039) is published by Cogent OA, part of Taylor & Francis Group.**

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com

