JOURNAL OF SMART ECONOMIC GROWTH

JSEG

www.jseg.ro ISSN: 2537-141X Volume 9, Number 1, Year 2024

TEACHING INFORMATION AND COMMUNICATION TECHNOLOGY IN EUROPEAN UNIVERSITIES: A NON-PARAMETRIC EFFICIENCY PERSPECTIVE

Simona- Elena Grădinaru, Anamaria Aldea

Department of Economic Informatics and Cybernetics

The Bucharest University of Economic Studies

simona.gradinaru@csie.ase.ro, anamaria.aldea@csie.ase.ro

Abstract: *Education in Information Technology is becoming more important than ever to the contemporary digital society. It is one of the fields with the fastest growing career paths globally. In Europe alone, year 2016 marked a rapid growth in the number of employed persons with an education in Information and Communication Technologies and the growth persists ever since. This paper aims to examine the efficiency of teaching undergraduates in ICT programmes in European Universities. The analysis employs the non-parametric framework of VRS-DEA estimator on an empirical dataset of 96 universities, extracted from The European Tertiary Education Register, taking into the account only the human resources involved in the teaching activities: personnel, enrolled students and graduates, with a distinction between undergraduate and doctoral, given the ISCED level. After a thorough preprocessing phase that included two different clustering algorithms to identify similar and comparable institutions in terms of the ICT specialisations offered, a set of statistical tests were applied in order to identify empirically the hypotheses of the production set, that allow for choosing the appropriate efficiency estimator, which was further employed. Results draw attention towards a potential pattern for the better performing universities with ICT programs to exhibit economies to scale, while at the same time loading the teaching staff slightly more than the low performing institutions.*

JEL classification: C14, C38, I23, N34, Q5.

Keywords: Nonparametric Efficiency Estimators, DEA, Higher Education, European Universities, ICT, cluster analysis, k-means, DBSCAN.

1. INTRODUCTION

The rise of Information and Communication Technologies (ICT) in shaping the modern digital society is indisputable. In a world where everything is gradually becoming a branch of Computer Sciences, preparing the future workforce in this area of expertise becomes more important than ever. Such career paths experienced growth at accelerated paces in the last decade, while new jobs are still being developed. A group of experts gathered in a workshop by Institute for the Future for Dell Technologies (2017) estimated that more than 80% of jobs that today's students will be doing in 2030 have not been invented yet.

In the European workforce alone, year 2016 marked the most rapid growth in the number of employed persons with a higher education in ICT, increasing by almost 10% compared to the prior year and it continues to expand, according to Eurostat (2024) for European Union (28 countries). This gives rise to an interesting research perspective on the ICT programs and their efficiency in offering specialised graduates for the workforce. One of the missions of Higher Education Institutions that becomes very relevant in this context is to anticipate the future skills and to train and prepare the population with such competencies.

The educational system relies on a process that utilizes a set of resources in order to obtain a set of outcomes. In economic theory, both resources or inputs, as well as outcomes or outputs, reveal the characteristics of a production process that can be modeled mathematically. A production function requires defining these concepts, given the purpose of the study and the data availability.

This paper aims to assess the efficiency of teaching ICT programs in European universities, based on the data available at European Tertiary Education Register (ETER) with regards to human resources involved into the educational processes. A non-parametric approach to Data Envelopment Analysis (DEA) has been choosen for this purpose, an approach frequently used for evaluating efficiency of universities, as in many other studies: Daraio, Bonaccorsi and Simar (2015); Gradinaru et al (2019), Mastromarco, Toma and Daraio (2022) and many more.

The rest of the paper unfolds as follows. The brief introduction is followed by a section of Related works on efficiency in education and more specific, applied on specific curricula, study programs and specialisations. Section 3 describes the non-parametric framework employed in assessing the efficiency of universities offering ICT programs. Section 4 describes the data, motivates the choice of variables and presents the preliminary analysis. Section 5 provides a detailed overview on the empirical analysis and ends with the main findings and Conclusions.

2. RELATED WORKS

Liu et al (2013) examine the scientific research on efficiency analysis and reveal that education is placed among top 5 fields analysed with non-parametric efficiency techniques. Salerno (2003) approaches the challenges that arise in studying efficiency in higher education institutions, one of which refer to sample homonenity in order to ensure comparability of universities.

Bonaccorsi, Daraio and Simar (2006) employ partial frontiers for asssessing the efficiency of Italian universities. Some of their findings suggest a potential trade-off between teaching and research activities, implying that such tasks are rather complementary than mutually exclusive. They also shape the concept of a third mission of universities, being obliged to contribute to the economic growth of the region from which they belong. The potential compromise between teaching and research is confimed not only on the european universities in the study of Gradinaru et al (2019), but also in the context of Romania, as in the empirical study of Stoica and Aldea (2016).

Witte and Lopez-Torres (2015) offer an extensive literature review on efficiency studies at university level and group the inputs and outputs employed in efficiency models into several categories that describe and influence the educational process. The authors suggest that inputs could include variables about the students (demographics, behavioral, parents' educational attainment), about the institution (dropout rate, expenditures on teaching and research, tuition fees, academic personnel, enrolled students, graduates), but also variables describing the community in which the institution operates (percent of households with pupils and students, or Herfindahl index, as a measure for local workforce competivity). However, it's worth mentioning that the choice of variables is heavily influenced not only by the perspective chosen for the analysis, but also by the availability of the data required for an empirical study.

In the case of Gradinaru et al (2019), the input-output space is given by PhD students and academic personnel (inputs), alongside the total graduates at ISCED 5-7 and some research outputs, such as number of publications Web of Science, average citations and number of participations to European Union Framework Programme (EU-FP) projects, EU's funding initiative for collaboration in research. One of the main findings reveal that uropean universities' efficiency in research is achieved by having an academic focus towards actively participating to the scientific community, as well as accessing or even coordinating EU-FP projects. On the opposite side, achieving efficiency in teaching requires a higher emphasis on offering as many graduates as possible. The authors suggest that most of the institutions analysed seem to focus on one mission only, in detriment of the other, with only a few universities achieving the optimal efficiency in both perspectives.

A study of Daraio, Bonaccorsi and Simar (2015) ranks the performance of european universities, based on an input space composed of academic staff and administrative personnel, together with the institutional expenditure associated with these resources and an output space defined by graduates (at ISCED levels 5 and 6) and regarding research activities (published papers, international collaboration, normalized impact factor). The authors unfold some countries that exceed the European performance standard: United Kingdom, Sweden and Switzerland, followed by Belgium, Austria, Ireland and Netherlands. The paper also discusses how universities should create a strong association between the student's profile and the educational offer, a concept at which Netherlands and Switzerland seems to excel, distinguishing from the other countries with outstanding performance in both teaching and research.

Most studies of efficiency in the educational system refer to higher education institutions (HEI) as a whole. The prevalence of scientific research with regards to STEM (Science, Technology, Engineering and Mathematics) programs, or even more specific, to separate specialisations offered by universities is rather scarce, despite the massive economic potential of STEM and ICT studies, in particular. However, this is probably also a consequence of data availability at such granular levels, most reports being focused on the universities overall.

Nevertheless, in a recent study, Mastromarco, Toma and Daraio (2022) compared the real sciences with the humanities within the Italian universities, separately for Bachelor's and Master degrees. For homogeneity purposes, they split the data by domains, keeping further only natural sciences, health sciences, social sciences versus humanities. The analysis uses a non-parametric approach for estimating the efficiency of each specialisation and education level, building a total of 8 separate models. The choice in terms of inputs refers to the ratio between academic personnel and enrolled students, and the outputs quantify the percentage of graduates within the legal duration of the courses, as well as the percentage of graduates that are satisfied with the program chosen. Results suggest that higher efficiency levels appear among strong scientific fields, which require numerous disciplines with intensive technological content, connected to the according industry, whereas the efficiency of such specialisations seems to reflect the economic reality of the local environment in which the respective institutions operate. Other studies focus on engineering programs, as De la Hoz (2021), which evaluates and classifies the efficiency of engineering specialisations in Colombia, in three phases. First, a cluster analysis is performed with the purpose of discovering similar universities, which are further employed in a Data Envelopment Analysis model for efficiency evaluation. Lastly, a Random Forest algorithm is performed on the results, in order to predict the efficiency measure previously built and to find the most predictive variables for efficiency estimates of engineering programs and specialisations.

3. METHODOLOGY

In efficiency analysis, decision-making units (DMU) are described by a set of p inputs utilized for producing q outputs, both strictly positive. In the non-parametric approach to estimate efficiency, a model is used to measure efficiency relative to a non-parametric and maximum likelihood estimate of a true frontier. To fully define this DEA model, a production set must incorporate all achievable pairs (x, y), assuming that input x can produce output y:

$$\Psi = \{(x, y) \in \mathbb{R}_+^p \times \mathbb{R}_+^q,\ x \text{ can generate } y\} \tag{1}$$

The upper boundary of the production set Ψ is of interest, as it identifies the points along this frontier as being technically efficieny, whereas the points operating beneath this boundary are found to be technically inefficient.

$$\partial\Psi = \{(x, y) \in \Psi | (\lambda x, y) \notin \Psi, \forall \lambda \in (0,1), (x, \lambda y) \notin \Psi, \forall \lambda \in (0,1)\} \tag{2}$$

A non-parametric estimator is usually used for measuring the efficiency of a decision-making unit (x, y): Data Envelopment Analysis (DEA). This estimator represents the minimal convex hyperplane that encompasses the production of all attainable points, hence the enveloping nature. One could measure the efficiency relative to the empirically determined convex frontier, $\widehat{\Psi}_{DEA}$, in various manners.

In an output-oriented case, the Farrell (1957) measure of technical efficiency for any unit is defined by the maximum attainable proportionate change in output levels, holding input levels constant, as following:

$$\lambda(x, y) = \sup\{\lambda, (x, \lambda y) \in \Psi\} \quad (3)$$

This score is subunitary, whilst highest values $\lambda(x, y) = 1$ are associated with units that achieve technical efficiency in maximizing the outputs; conversely, the DMUs with values $\lambda(x, y) < 1$ are identified as technically inefficient in the output orientation.

Reformulating the concepts in a probabilistic approach, as proposed by Simar and Wilson (2000), requires defining a probability for a unit (x, y) to be dominated, further decomposed in the output-oriented approach decomposed using the conditional survivor function of Y, $S_{Y|X}$ and the distribution function of X, F_X:

$$H_{XY} = Prob(Y \geq y|X \leq x) \cdot Prob(X \leq x) = S_{Y|X}(y|x) \cdot F_X(x) \quad (4)$$

In this setting, the efficiency measure in the output orientation becomes:

$$\lambda(x, y) = \sup\{\lambda \mid S_{Y|X}(\lambda y|x) > 0\} \quad (5)$$

In empirical studies, it is essential to test the convexity of the production set, since DEA is statistically consistent only under the convexity assumption. Whenever this is not respected, another estimator that relaxes the convexity assumption should be employed, as it is the case for Full Disposal Hull (FDH). Moreover, if the convexity is present, one should carefully examine whether the units operate under constant or variable returns to scale (CRS versus VRS), as estimators DEA-CRS and DEA-VRS have different convergence rates.

In order to facilitate the choice of the adequate estimator, Kneip et al (2016) and Simar and Wilson (2020) developed statistical tests for differences in average efficiency across two groups of units, for each hypothesis, testing separately convexity versus free disposability, as well as CRS versus VRS. This recently introduced approach requires splitting the sample into two subsamples in a process repeated many times, each of those subsets being further employed into the separate estimators: DEA-VRS and FDH, for testing the convexity, DEA-CRS and DEA-VRS, respectively, for testing the returns to scale. When significant differences are identified between the averages of these estimators over the two subsamples, the null hypothesis of the test is accepted: convexity or constant returns to scale.

When performing both tests, this later translates into imposing these assumptions on the production set, thus employing for the efficiency assessment either DEA for convex production datasets or FDH for free disposability, as well as the appropriate DEA version depending on the confirmed returns to scale. The fundamental benefit of this technique consists in allowing the researcher to choose empirically the adequate estimator, given the nature of the production set one aims to study.

4. DATA OVERVIEW

The database of European Tertiary Education Register (ETER) was used as source for gathering information on the European universities that offer ICT study programs. The raw data contains aggregate figures for around 2,964 Higher Education Institutions (HEIs), but a large amount of them have partial or incomplete information available for ICT specialisations. After imposing the nonnegativity constraint to all variables subject for the efficiency analysis, a valid sample was identified containing only 228 HEIs that offer all ISCED levels: Bachelor's and Master degrees, as well as doctoral studies. The indicators subject of investigation are their according resources involved in the teaching undergraduate activities: academic personnel, PhD students and Bachelor degrees offered, all applied only for Information and Communication Technologies. Table 1 highlights the variables extracted for describing the teaching activities.

Table 1. Variable Description.

Variable	Definition
Academic staff (HC)	Personnel whose primary assignment is instruction or research
Students enrolled at ISCED 8	PhD Candidates, potentially involved in teaching activities
Graduates at ISCED 6	Graduates with a Bachelor's degree

Subsequently, a thorough inspection of the data was conducted. A detailed analysis is needed, as it appears that data is far from homogenous, an issue illustrated by the broad data ranges that also reveal some potential outliers. At a glace, it may appear that the dataset contains a broad range of ICT programs and specialisations, from very small to very large and perhaps even prestigious. Such heterogeneity may affect and actually bias the efficiency frontier, due to which a meticulous analysis is required. Table 2 shows the summary statistics for the initial dataset of 228 universities with ICT programs, including the teaching workload with undergraduates, defined as the ratio between students enrolled at ISCED 6 and academic staff. One could easily spot the lack of comparability between these programs, as they vary broadly. Continuing the deep dive, data seems rather heterogenous from multiple perspectives. First, the sample contains a few observations with very small inputs and output, which aren't necessarily comparable with the other institutions.

Table 1. Descriptive statistics of initial heterogenous data.

Variable	Usage	Minimum	Median	Mean	Maximum
Academic staff (HC)	input	5	69.5	95.1	452
Students enrolled at ISCED 8	input	2	43	60.1	319
Graduates at ISCED 6	output	2	85.5	111.1	440
Teaching workload	descriptive	.03	9.2	13.4	111.8

Secondly, a similar pattern emerges when it comes to teaching workload with undergraduates, which may involve significantly different efforts in the teaching activities and perhaps institutions that are impossible to compare from the perspective of their production process. This assumption led to filter the data to cut a small proportion of each distribution's left tail, to ensure a reasonable size of ICT programs, as well as dropping the institutions with teaching load with undegraduates significantly higher than the average load of the data set. Imposing all these constraints led to keeping further 129 institutions of interest.

Table 2. Descriptive statistics of the filtered dataset.

Variable	Usage	Mininum	Median	Mean	Maximum
Academic staff (HC)	input	25	97	125.3	452
Students enrolled at ISCED 8	input	13	52	73.3	319
Graduates at ISCED 6	output	17	83	101.7	367
Teaching workload	descriptive	.66	7	6.8	12.8

The resulting sample was subject to a principal components analysis, revealing that more a high proportion of the original variance can be explained by only two dimensions, and thus aggregating together the inputs. This was further introduced into a k-means clustering, for the purpose of identifying groups of homogenous observations, if the dataset still contains extreme values. The algorithm created two more homogenous groups, the best solution in terms of its silhouette, with 21 and 108 observations each; the second cluster includes small and medium size ICT programs, whereas the first is composed of large universities, that through comparison could be easily treated as outliers.

To ensure there weren't any potential remaining extreme values, the cluster containing 108 universities was fed into a DBSCAN clustering algorithm, that can also perform outlier detection. Surprisingly, after finding the appropriate parameters that optimize cluster silhouette, DBSCAN identified a group of 5 potential outliers and a homogenous cluster of 103 institutions. Lastly, after a thorough examination of distributions, an additional 7 universities were identified as outliers. Ultimately, the homogenous group of 96 ICT programs was selected for further analysis.

Table 4. Descriptive statistics of the homogenous dataset.

Variable	Usage	Mininum	Median	Mean	Maximum
Academic staff (HC)	input	25	82.5	91.2	208
Students enrolled at ISCED 8	input	14	48	52.5	117
Graduates at ISCED 6	output	17	72.5	102	180
Teaching workload	descriptive	1.7	7.2	7	12.8

A glimpse into Table 4 denotes a more homogenous data set than both the initial and the filtered set of HEIs. Variable ranges were significantly reduced, being narrowed from 452 to 208 for academic personnel, from 367 to 180 for graduates with a Bachelor's degree and from 319 to 117 for PhD students.

A preliminary analysis of the clean data set is illustrated in Figure 1. Teaching personnel, both academic staff and PhD students, was plotted against the students enrolled and graduates at ISCED 6, with the purpose of identifying any institutions that may have a bias on the efficiency frontier. Across all dimensions of interest, most universities are located within the bottom left quadrant of the plot, revealing a potential degree of overlapping between the variables.

JOURNAL OF SMART ECONOMIC GROWTH

www.jseg.ro ISSN: 2537-141X Volume 9, Number 1, Year 2024

The first chart plots the academic staff and the students enrolled for Bachelor degrees, in which a few German institutions are distinguished by large academic personnel, as seen in the right quadrant: Tübingen University (DE0008), University of Stüttgart (DE0007), University of Bremen (DE0040), University of Kaiserslautern (DE0085). A similar pattern appears the third chart, when comparing academic staff and graduates.

The second plot brings into perspective the students enrolled and the graduates, according to which three HEIs are prominent by highest number of degrees offered, as it is the case for University of Lincoln (UK0065), Brunel University London (UK0017) and Ludwig Maximilian University of Münich (DE0021). Interestingly, The University of Lincoln seems to distinguish from the remaining data through multiple perspectives.

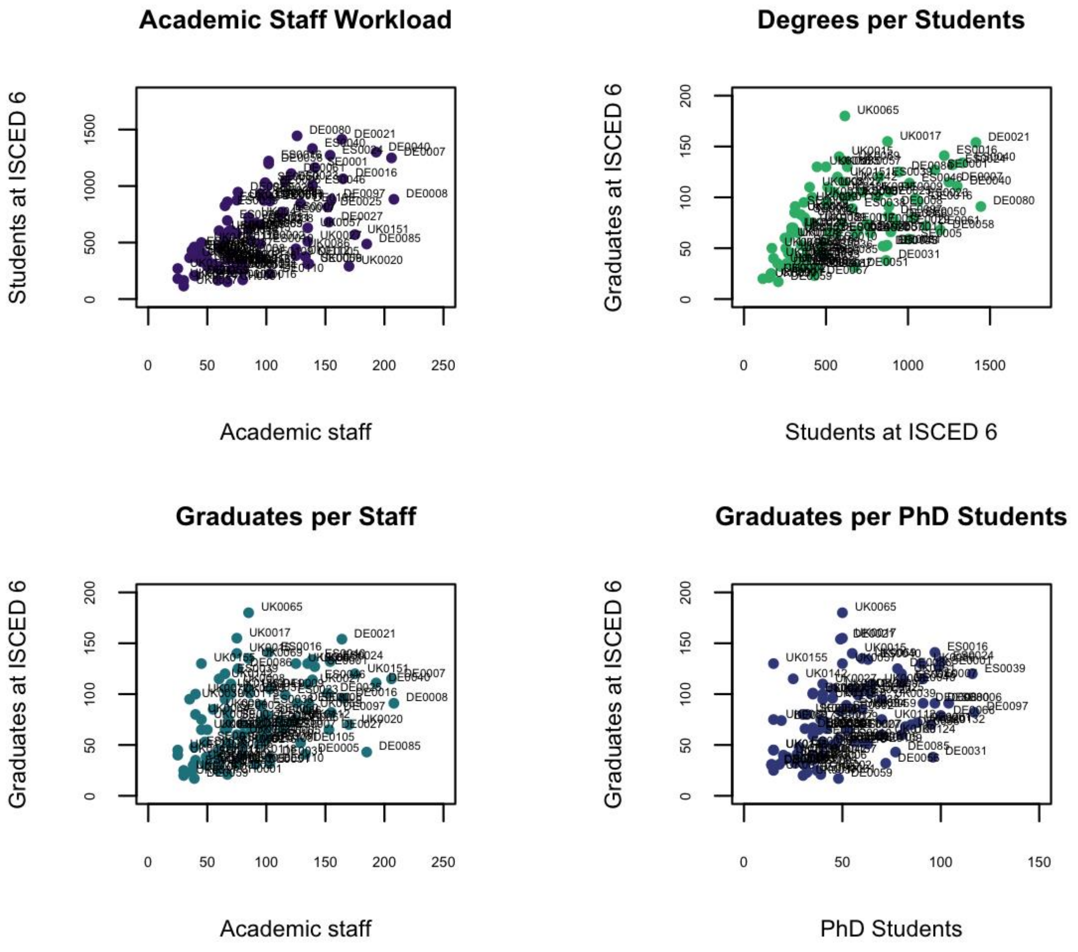

Figure 1. Preliminary analysis.

When analysing universities with ICT specialisations, the output orientation appears to be suitable for the higher education institutions, as they rather aim to maximize the teaching output than trying to minimize the available resources. With this assumption in mind, a teaching model was built with two inputs that describe the personnel involved in the teaching activities for ICT Bachelor's degree courses, mainly the academic staff and the PhD candidates in this field, and a single output given by the number of graduates, thus the number of ICT Bachelor's degree offered in 2016 by each university.

With the purpose of assessing the efficiency frontier with or without the convexity of the attainable set, a bidimensional dataset is preferred for interpretability, containing a unidimensional input and a unidimensional output. Moreover, perhaps not all PhD candidates in ICT fields are involved in the teaching activities, whereas some academic personnel may not be involved in undergraduate courses. This is a real challenge in practice, when such accurate data is rarely available. Instead, the researcher is rather obliged to either estimate empirically the way of aggregating such variables, make additional assumptions based on the available information, like simply cummulating the potential human resources involved into teaching undergraduates, or even ignore this aspect.

With respect to this aspect, we employed the dimensionality reduction method proposed by Daraio and Simar (2007), based on Mouchart and Simar (2002) for aggregating the input variables. This technique uses the standardized inputs combined in a linear combination that returns the aggregated input, using as weights in this trasformation the resulting values of the first eigen vector given by single value decomposition of matrix of standardized inputs X.

In this way, the weights for creating an aggregated input are obtained empirically. The input aggregated with this dimensionality reduction has a high degree of overlapping with the initial inputs, returning a correlations of around +.9 with both the academic personnel and PhD students, supporting that the loss of information in dimension reduction is minimal.

The next sections describes the steps performed for the analysis and the decisions made along the way based on empirical findings. A detailed set of interpretations is given after estimating the efficiency of universities from the perspective of teaching ICT undergraduates.

RESULTS

As detailed in the methodological section, it is essential to test the convexity hypothesis and the returns to scale of the production set. For this purpose, we employed the FEAR package in R, which returns a statistical test and its according p-value, as developed by Kneip et al (2016) and Simar and Wilson (2020). Table 5 shows the statistics generated for the homogenous group of 96 HEIs, when performing a test on 10 splits and 1000 bootstrap replications.

Table 5. Results of hypothesis testing using 10 splits and 1000 bootstrap replication.

Hypothesis	Null Hypothesis	$\{\hat{T}_{96,b}\}_{b=1}^{B}$	$\hat{p}_T$
Convexity	Convexity	-3.35	.98
Returns to scale	CRS	4.52	.03

For a test of size $\alpha = .95$, the null hypotheses of either CRS or convexity are rejected when $\hat{p}_T$ returns values below $1 - \alpha = .05$. In the statistical framework, $\hat{p}_T$ quantifies the probability of observing the result when the null hypothesis is true, often interpreted as chances of being wrong when accepting the null hypothesis. According to the results provided in Table 5, the analysed production set is indeed *convex* and subject to *variable returns to scale*. These findings lead towards the use of VRS-DEA estimator in the efficiency analysis, despite its less constrained version, FDH, which is less appropriated when convexity has been shown.

In view of assessing the teaching efficiency of ICTs programs in European universities, the output orientation is intuitive in the context of maximizing the number of degrees offered. After testing and identifying the adequate hypotheses, we build VRS-DEA efficiency frontier for assessing the teaching in ICT programs, as illustrated in Figure 2. In this configuration, the average efficiency estimate reaches .46, a criterion based on which 53 are marked as inefficient, due to a lower efficiency estimate and the remaining 43 institutions are marked as efficient, three of them achieving maximum efficiency. These refer to British universities only, as the chart below suggests: UK0121, UK0155 and UK0065.

JSEG
JOURNAL OF SMART ECONOMIC GROWTH
www.jseg.ro ISSN: 2537-141X Volume 9, Number 1, Year 2024

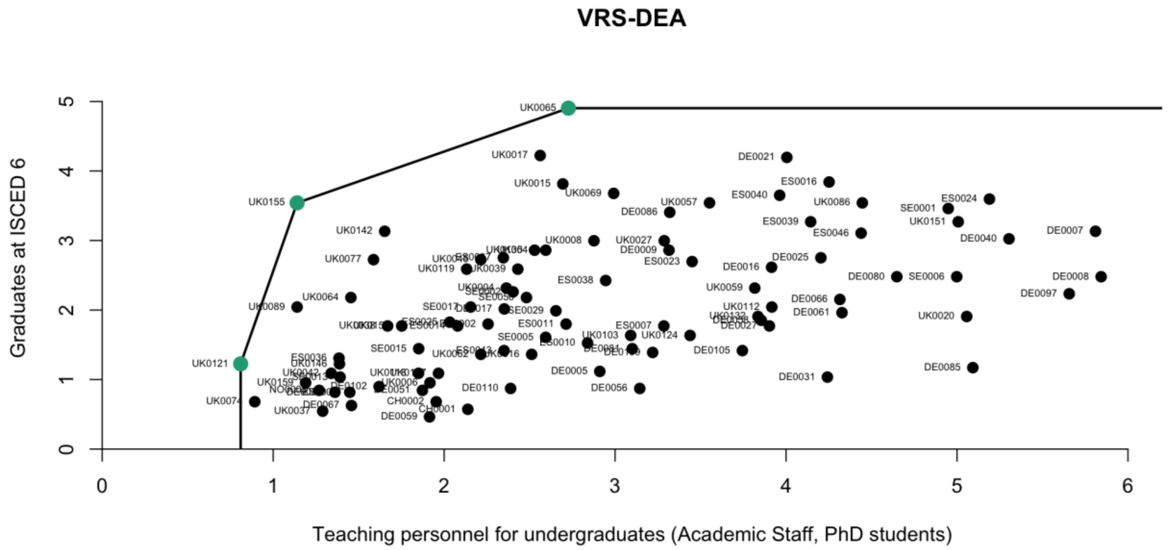

Figure 2. VRS-DEA efficiency frontier.

First of them, UK0121, while operating at the lowest scale in terms of resources, is a balanced example as it managed to offer 45 Bachelor degrees in ICTs in 2016, teaching the undergraduates with 25 employees in the ICT academic personnel and 15 PhD students. It seems that a more emphasis is placed on having enough academic staff well prepared for broad topics included in their curricula, campuses upgraded with modern infrastructure and high-tech equipment and labs (some of them being accessible 24/7), as well as having dedicated IT facilities that refer to technical operations or online learning modules and resources. However, these benefits are available to all students regardless of their study field, which also reflects into a high student satisfaction.

UK0155, located in Aberdeen, Scotland, managed to offer 130 Bachelor degrees in ICT in 2016, using as teaching personnel 15 PhD students and an academic staff of 45 employees. The School of Computing within this university offers a range of courses that reflects the modern computing landscape, many of them being recognized by the British Computer Society. The ICT profiles are said to approach teaching and learning through a combination of techniques, like interactive lectures, group projects and practical and innovative exercises in their computing labs, also offering virtual learning. The School of Computing within this university offers a broad range of specialisations for Bachelor of Science, like Application Software Development, Computing and Creative Design, Cyber Security and Network Management, Data Science, Digital Media, or even Games Design or Web and Mobile Design.

Lastly, UK0065 gave a total of 180 graduates with a Bachelor's degree in ICT in 2016, having an academic staff of 85 employees and 50 PhD students in the same domain. The Bachelor's degree in Computer Science provides students the knowledge and skills to develop a broad range of computing solution for real world and business problems. A particular emphasis is given to cutting-edge topics, like machine learning and artificial intelligence, that are included in the 3-years curricula, together with core disciplines and digital tehnologies, like cloud computing, robotics or big data. Interestingly, the university also allow students to choose a placement year in the industry in order to gain experience, during which no tuition fees are being applied. Moreover, the university has launched a collaboration with an industrial automation company in 2016, that allow students to explore real life problems and challenges and to tackle them with modern automation solutions, especially when it comes to robotic applications.

Although not directly situated on the efficiency frontier, other examples of highly efficient ICT specialisations appear in UK0017, DE0021 or ES0016. At the opposite pole, among the worst performing universities in maximizing ICT graduates are institutions from Switzerland (CH0001, CH0002) or Germany (DE0059). What these have in common is a relatively low number of graduates when compared to the teaching personnel involved in undergraduate activities.

As mentioned above, 53 universities were marked as efficient since they register efficiency estimates below the sample average, and the remaining 43 institutions are marked as efficient, the three previously described institutions achieving maximum efficiency. Figure 3 shows the differences between these two categories. A glimpse into the chart suggests that on average the highly efficient ICT profiles tend to offer more than double Bachelor degrees in comparison to the low efficient specialisation (around 110 graduates versus 50). However, this could be an effect on operating on a larger scale, as confirmed by both the number of students enrolled and the according teaching personnel.

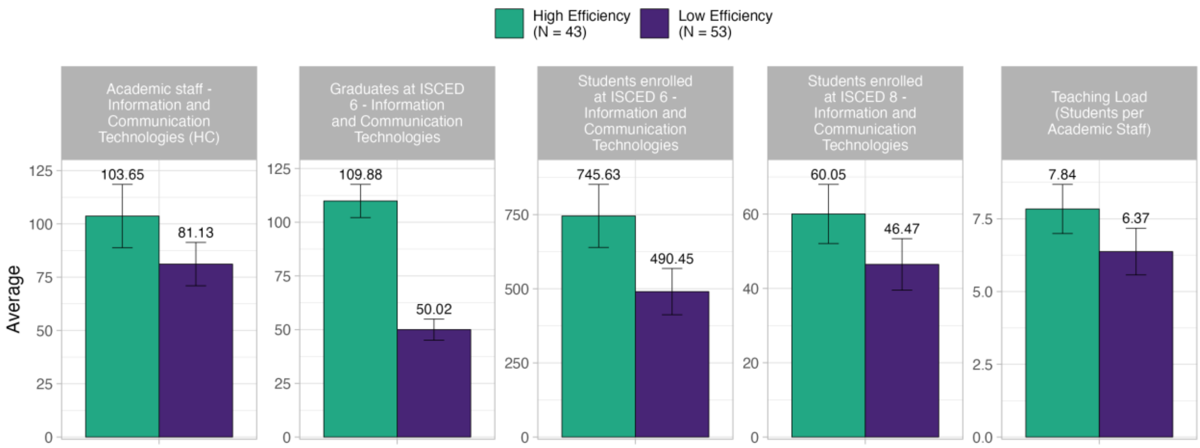

Figure 3. Average differences between the two groups of ICT programs.

The institutions with low efficiency ICT specialisations appear to offer significantly fewer degrees, less than a half than the highly efficient profiles, despite the personnel and students enrolled being with only 20-30% lower than the other HEIs. Interestingly, when it comes to teaching load, as defined by the ratio of undergraduates and academic staff, the efficient units seem to be slightly more overload than the inefficient ones: approximatively 7.8 students per teacher versus around 6.4. Although the difference appears to be statistically significant, it's worth reminding the reader that for comparability purposes, the analysis itself was focused on institutions with a teaching load lower than the European average among the universities with available data, hence the rather low values.

Ultimately, Figure 4 presents the distribution of universities that achieve high efficiency and low efficiency by European countries. As previously mentioned, the three universities with maximum efficiency below to the UK, whose institutions included into the analysis achieve a high level of efficiency in teaching ICT graduates, as more than half of British universities are classified as highly efficient. This looks significantly different than the remaining countries.

The most similar behavior appears to Spain, which registers a perfect balance between efficient and inefficient universities with ICT programs. Sweden and Germany seem to have more inefficient institutions from the perspective of maximizing the ICT Bachelor degrees, whereas Norway and Switzerland have only inefficient universities in the current sample.

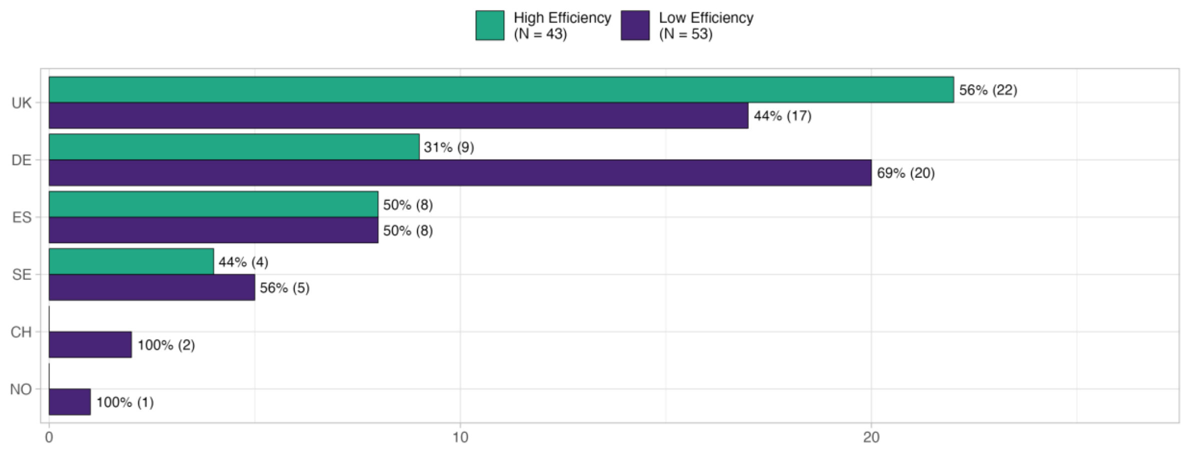

Figure 4. Distribution of high efficiency versus low efficiency ICT profiles within countries.

This is an insightful outlook, as it suggests that these four countries might have a focus on other priorities, such as offering more advanced degrees (Master, Doctoral), emphasizing more the research activities or perhaps giving a higher importance to courses quality than simply obtaining as many graduates as possible. More research is required in order to reveal the potential causes of such inefficiencies, as well as encompassing the more advanced courses into analysis, Master and Doctoral.

CONCLUSIONS

The digital revolution that society experiences nowadays influences everything, from the smartphones we use on a daily basis to the most recent advances in artificial intelligence that bring upcoming transformations yet not fully understood. Experts declare that the future will brings many jobs that have not been even invented so far, for which universities have to prepare the population with such competencies, whilst anticipating these skills of the future.

Education in Information Technolohy is becoming more important than ever to the contemporary digital society. Such specialisations provide a broad range of knowledge, skills and tools to tackle and solve real-world challenges and business problems in many fields, which have recently become in high demand given the most latest technology advancements. It is one of the fields with the fastest growing career paths globally and in Europe alone, year 2016 marked a rapid growth in the number of employed persons with an education in ICT and the growth persists ever since.

The purpose of this paper is to assess the efficiency of teaching undergraduates in European universities with ICT specialisations, according to 2016. The analysis employs the nonparametric frontier of VRS-DEA estimator on an empirical dataset of 96 institutions extracted from The European Tertiary Education Register (ETER). The proposed model considers the personnel involved in teaching undergraduates to be given by academic staff and PhD students, as part of their doctoral studies, while aiming for maximizing the Bachelor degrees offered.

The data was subject to a thorough examination before reaching to a sample of comparable institutions. In order to ensure comparability, a small proportion of the distributions' left tail was cut, a filter was applied based on teaching workload, then a homogenous group of universities was identified after applying k-means and DBSCAN, to filter out any potential outliers that could bias the estimator. The last step before estimation consisted in testing the hypotheses of the production set, thus confirming empirically the convexity and variable returns to scale, which further led to employing the adequate estimator, which in this case refers to VRS-DEA.

The efficiency analysis revealed three institutions with maximum efficiency, situated on the frontier and another 50 highly efficient universities with estimates above the sample mean, whereas the remaining 43 institutions were marked as efficient. The differences between the two categories seem to be given mostly by the graduates obtained, the high-efficient category registering more than double graduates than the low-efficient group. However, this looks like confirming the existence of economies of scale among the high-efficient group of institutions with ICT programs.

Potential future research involves expanding the study of efficiency in teaching activities of Information and Communications Technology towards the more advanced courses, Master and Doctoral. Furthermore, since this domain is just a niche from the broader umbrella of science, technology, engineering and mathematics (STEM), an interesting future direction could study the efficiency of teaching in all these domains, across all tertiary education levels: Bachelor's, Master and Doctoral.

CONFLICTS OF INTEREST AND PLAGIARISM: The authors declare no conflict of interest and plagiarism.

REFERENCES

1. Bonaccorsi, A., Daraio, C., Simar, L. (2006). Advanced indicators of productivity of universities. An application of robust nonparametric methods to Italian data. *Scientometrics*, 66(2), 389-410.
2. Daraio, C., Simar, L. (2007). *Advanced. Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications*. Springer Science & Business Media.
3. Daraio, C., Bonaccorsi, A., Simar, L. (2015). Efficiency and economies of scale and specialization in European universities: A directional distance approach. *Journal of Informetrics*, 9(3), 430-448.
4. De la Hoz, E. (2021). Assessing and classification of academic efficiency in engineering teaching programs. *Full Research Paper.*
5. Eurostat (2024). *Employed persons with ICT education by educational attainment level* – annual data [Data set]. Available at: https://ec.europa.eu/eurostat/databrowser/view/isoc_ski_itedu/default/table?lang=en
6. Farrell, M.J. (1957). The measurement of the productive efficiency. *Journal of the Royal Statistical Society: series A (General)*, 120(3), 253-281.
7. Gradinaru, S., Aldea, A., Besir, L., Albu, C. (2019). Are teaching and research activities mutually exclusive? A data mining study on European universities. *Proceedings of the International Conference on Applied Statistics*, 1(1), 275-287.
8. Institute for the Future for Dell Technologies (2017). Available at: https://www.delltechnologies.com/content/dam/delltechnologies/assets/perspectives/2030/pdf/SR1940_IFTFforDellTechnologies_Human-Machine_070517_readerhigh-res.pdf
9. Kneip, A., Simar, L., Wilson, P.W. (2016). Testing hypotheses in nonparametric models of production. *Journal of Business & Economic Statistics*, 34(3), 435-456.
10. Liu, J.S., Lu. L.Y., Lu, W.M., Lin, B.J. (2013). A survey of DEA applications. *Omega*, 41(5), 893-902.
11. Mastromarco, C., Toma, P., Daraio, C. (2022). Teaching Efficiency of Italian Universities: A Conditional Frontier Analysis. *Teaching, Research and Academic Careers*, 319-358.
12. Mouchart, M., Simar, L. (2002). Efficiency analysis of Air Controllers: first insights. Consulting report, 2002.

13. Salerno, C.S. (2003). *What we know about the efficiency of higher education institutions: The best evidence.* Center for Higher Education Policy Studies (CHEPS).
14. Simar, L., Wilson, P.W. (2000). Statistical inference in nonparametric frontier models: The state of the art. *Journal of Productivity Analysis*, 13, 49-78.
15. Simar, L., Wilson, P.W. (2020). Hypothesis testing in nonparametric. Models of production using multiple sample splits. *Journal of Productivity Analysis*, 53, 287-303.
16. Stoica, M., Aldea, A. (2016). Efficiency of teaching and research activities in Romanian universities: An order-alpha partial frontiers approach. *Economic Computation & Economic Cybernetics Studies & Research*, 50(4), 169-186.
17. Wilson, P.W. (2008). FEAR: A software package for frontier efficiency analysis with. R. *Socio-economic planning sciences*, 42(4), 247-254.
18. Witte, K.D., Lopez-Torres, L. (2015). Efficiency in. education: A review of literature and a way forward. *Journal of the Operational Research Society*, 68, 339-363.