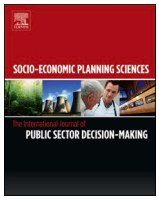
Socio-Economic Planning Sciences 92 (2024) 101845

Contents lists available at ScienceDirect

Socio-Economic Planning Sciences

journal homepage: www.elsevier.com/locate/seps



Benchmarking performance through efficiency analysis trees: Improvement strategies for colombian higher education institutions

Jose Luis Zofio [a,b], Juan Aparicio [c,d], Javier Barbero [a], Jon Mikel Zabala-Iturriagagoitia [e,f,g,*]

[a] Department of Economics, Universidad Autónoma de Madrid, Madrid, Spain
[b] Erasmus Research Institute of Management (ERIM), Erasmus University, Rotterdam, the Netherlands
[c] Centro de Investigación Operativa (CIO), Universidad Miguel Hernández, Elche, Spain
[d] Valencian Graduate School and Research Network of Artificial Intelligence, Valencia, Spain
[e] Deusto Business School, University of Deusto, San Sebastián, Spain
[f] South Eastern University Norway, Kongsberg, Norway
[g] CIRCLE, Lund University, Lund, Sweden

ARTICLE INFO

Keywords:
Machine learning
Efficiency analysis trees
Data envelopment analysis
Benchmarking
Higher education institutions
Colombia

ABSTRACT

We introduce benchmarking analysis based on state-of-the-art machine learning techniques applied to the measurement of efficiency to assess the performance of Higher Education Institutions (HEIs). We rely on Efficiency Analysis Trees (EAT) and its Convexified frontier counterpart (CEAT) to assess the efficiency of 144 private HEIs in Colombia and compare the results with those achieved with classical Data Envelopment Analysis (DEA). Both EAT and CEAT show a higher discriminatory power than DEA when determining efficiency scores. Our results identify the different splits of the production frontier, corresponding to each node of the efficiency tree, which groups HEIs according to specific management models. By identifying relevant peers for inefficient observations at the node level, we show which strategic guidelines can be adopted to improve the performance of each HEI. This process encourages mutual learning and suggests potential changes within each node leading to efficiency improvements.

1. Introduction

This study introduces benchmarking methods to analyze the performance of Higher Education Institutions (HEIs) using state-of-the-art efficiency methods based on machine learning techniques, i.e., Efficiency Analysis Trees. From an empirical perspective, the new methods focus on the analysis of the different managerial models characterizing the nodes generated by the efficiency tree, clustering Colombian HEIs according to similar input-output mixes and size. Subsequently, the analysis identifies the set of efficient observations within each node that may serve as benchmark peers for inefficient observations. The principle of least action (i.e., minimum distance) is adopted to determine the closest best performing peer to each observation within its node. Finally, the method involves the calculation of the output adjustments necessary to match the production levels of the peers, thereby identifying best practices within each node and the whole tree. Our goal is to assist decision makers responsible for the management of HEIs in designing strategies aimed at solving productive inefficiencies, which results in the improvement of the overall performance of higher education systems.

The literature has shown that one-size-fits-all models are not suitable for evaluating the performance of HEIs [1,2]. This is mainly due to the fact that HEIs tend to develop greater and better organizational capacities in those activities that offer them competitive advantages over their counterparts [3–5]. Consequently, it is expected that HEIs with different educational models and levels of performance may coexist, and that the performance of a HEI may also vary depending on the functions that are considered in such an evaluation [6]. This has led to performance measurements being increasingly used to characterize HEIs and to classify them [7,8], defining strategic groups that facilitate decision-making due to the homogeneity of the clustered HEIs [9]. Accordingly, there is an increasing academic and practical interest in using methodologies that make it possible to identify the best practices and the HEIs that may represent benchmarks to learn from Ref. [10–12], with the purpose of establishing improvement plans with specific and achievable objectives [13,14].

Efficiency evaluations of HEIs have yielded empirical evidence with

* Corresponding author. Deusto Business School, University of Deusto, San Sebastián, Spain.
E-mail address: jmzabala@deusto.es (J.M. Zabala-Iturriagagoitia).

https://doi.org/10.1016/j.seps.2024.101845
Received 30 March 2023; Received in revised form 17 January 2024; Accepted 14 February 2024
Available online 22 February 2024

important implications for improving the definition of higher education policies and strategies [7,15]. This is a result of the assistance these methodologies provide when defining better-targeted policy objectives, which are in turn adjusted to particular groups of HEIs with similar characteristics but dissimilar levels of performance. In this regard, the methodology that is proposed in this study is expected to contribute to the literature with new evidence on how to carry out such systematic evaluations of the performance of HEIs [16].

The most widely used methodology for evaluating the efficiency of HEIs is the envelopment approach generally identified as Data Envelopment Analysis (DEA), introduced by Charnes et. al [17] to, precisely, measure the performance of educational programs. Although these methods have proved themselves reliable, as shown in the many studies cited above, they suffer from overfitting, which may lead to situations where it is not possible to discriminate among observations when evaluating performance. This drawback is worsened when the number of observations is relatively small with respect to the number of inputs and outputs. Indeed, based on our results using both DEA and its non-convex counterpart known as Free Disposal Hull (FDH) in the context of Colombian private HEIs, we show that the number of efficient observations is very high. In particular, in the case of FDH this situation is extreme since all HEIs are deemed efficient with a score of one. This drawback, known as the "curse of dimensionality", can be seen as the Achilles's heel of envelopment methods.

This lack of discernment emerges from the weight flexibility that FDH and DEA offer when evaluating the efficiency of a given Decision Making Unit (DMU), as it searches for the most favorable weights. To solve this weakness, several methods have been proposed in the literature to improve the discriminatory power of DEA. The simplest methods, also adopted in this study (see Section 4), rank efficient DMUs according to their relative importance to inefficient units; for instance, how often they serve as reference peers for inefficient units (see Ref. [18]). A second method resorts to the super-efficiency approach initiated by Andersen and Petersen [19], which removes the DMU under evaluation from the reference technology, thereby obtaining an efficiency score smaller than one, which allows breaking the tie among the scores. A third strand of literature develops cross-efficiency methods (see Refs. [20,21]), which uses the DEA weights of all other DMUs in the analysis to assess bilateral efficiency, and then calculates the average, offering a distinct ranking. The fourth possibility brings inefficient frontiers into the analysis by solving 'inverted' DEA models, which measure inefficiency with respect to reference hyperplanes defined by the worst-performing DMUs [22]. Within the so-called TOPSIS methods (i.e., Technique for Order of Preference by Similarity to Ideal Solutions), it is possible to rank observations identifying the best (ideal) and worst (anti-ideal) performance. The ranking in this fifth method takes into consideration both how close and how far the DMUs are from these two benchmarks, respectively (see Ref. [23,24]). Other solutions to the weak discriminatory power of DEA are also discussed by Aldamak and Zolfaghari [25] and Balk et al [26].

A shortcoming shared by the above proposals is that the optimal weights, obtained by solving the DEA models, may not be unique. Indeed, applying the simplex method identifies a feasible solution, but there might be multiple solutions besides the first one obtained, which creates uncertainty in the evaluation process, and may lead to conflicting prescriptions from a managerial perspective (e.g., in the form of multiple rates of transformation between outputs or substitution between inputs). To address the challenges associated to effective ranking and multiplicity of benchmarks, several authors are increasingly resorting to machine learning methods to approximate the production technology as an alternative to envelopment techniques. One promising alternative are the so called Efficiency Analysis Trees (EAT) introduced by Esteve et al. [27] and Aparicio et al. [28]. EAT draws from the classification and regression trees methods proposed by Breiman et al. [29]. EAT estimates performance frontiers satisfying fundamental postulates of production theory such as free disposability, data envelopment and (if required) convexity, giving rise to Convexified Efficiency Analysis Trees (CEAT). Moreover, using cross-validation methods, the above authors have shown that these techniques are more accurate in predicting the actual production frontier. Therefore, machine learning methods like EAT and CEAT, provide a real alternative to perform benchmarking analysis for inefficient observations. However, the practice of benchmarking using these methods have not been developed in the literature. This study aims to solve this gap. What ultimately matters is the possibility of comparing an underperforming observation with some peer(s) that can realistically offer guidance to improve its performance.

In this study we apply the EAT and CEAT methodological approaches to assess the performance of a set of 144 private HEIs in Colombia and show their advantage over envelopment methods, while providing an effective benchmarking environment, which is introduced in this study. The expansion that the private HEI sector is having in many countries represents an opportunity to diversify and meet market demands [30] and to provide alternatives to broaden the coverage of higher education [31]. However, currently, private HEIs are facing great difficulties related to the decrease in demand, global competition, and cross-border education, among others, thereby, showing evident signs of decline [32], which have led governments to question their sustainability over time. In this regard, the decision to focus on Colombia as a case study is twofold. First, Colombia has one of the largest degrees of privatization in higher education worldwide [33]. Indeed, Colombia's private HEIs represent 70.1% of the higher education system, representing one of the systems with the highest degree of privatization [34]. Second, and in spite of the expansion of for-profit and private HEIs worldwide [35,36], so far there is limited evidence about the performance of these HEIs in terms of efficiency.[1]

Section 2 reviews the literature by comparing standard envelopments techniques, such as non-convex FDH and convex DEA, with newly proposed counterparts like EAT and its convexified version (CEAT). Section 3 presents the results achieved with the application of the previous methodologies in the context of Colombian private HEIs. It first shows the results of the EAT technique when growing the tree approximating the benchmark technology and then compares these with those corresponding to their FDH and DEA counterparts. Section 4 illustrates how to practice benchmarking using EAT and CEAT. In it, we identify relevant benchmarks and the output adjustments that lead to efficiency enhancement strategies in each of the identified nodes of HEIs. Finally, Section 5 concludes the paper by emphasizing its contribution to benchmarking analysis as well as its implications for the management and governance of higher education.

2. Standard envelopments techniques versus efficiency analysis trees

In this section, we recall well-known concepts about standard envelopment techniques to approximate the technology like the non-convex Free Disposal Hull (FDH) and convex Data Envelopment Analysis (DEA), and compare them to newly proposed counterparts like non-convex Efficiency Analysis Trees (EAT) and Convexified Efficiency Analysis Trees (CEAT).

2.1. Free Disposal Hull and Data Envelopment Analysis

Envelopment methods, particularly DEA, is among the most popular techniques to analyze performance in education (for a recent review of the literature see Ref. [15]). It is a very flexible technique that does not require to explicitly specify the relationship between the (educational) outputs and inputs. Additionally, this methodology does not assume any

[1] Some exemptions worth stressing are Thursby and Kemp [62], Castano and Cabanda [63], Gwendolyn and Cabanda [64], Said [65] or Sav [66].

probability distribution on the data generation process and allows to deal with the multi-output multi-input production technology in a natural way. Moreover, as another advantage in comparison with parametric alternatives (e.g., Stochastic Frontier Analysis), DEA also yields benchmarking information. In particular, it provides individual knowledge about real peers that serve as targets to learn from in order to improve performance.

Let us consider that we have observed n DMUs that make up the sample $\aleph = \{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}_{i=1}^{n}$. Let we assume that DMU$_i$ consumes $\boldsymbol{x}_i = (x_{1i},..., x_{mi}) \in R_+^m, i = 1,...,n$, amounts of inputs to produce $\boldsymbol{y}_i = (y_{1i}, ..., y_{si}) \in R_+^s$, $i = 1, ..., n$, amounts of outputs. Hereinafter, we will use bold for denoting vectors, and non-bold for scalars. The relative efficiency of a given DMU$_o$ in the sample is assessed with reference to the production technology, which is defined as: $\psi = \{(\boldsymbol{x},\boldsymbol{y}) \in R_+^{m+s} : \boldsymbol{x}\ can\ produce\ \boldsymbol{y}\}$. As noted above DEA is one of the standard non-parametric methods to measure the efficiency of the DMUs (i.e., Colombian HEIs in the context of this paper). The technology generated with this method meets certain assumptions, such as: a) free disposability of inputs and outputs; meaning that if $(\boldsymbol{x},\boldsymbol{y}) \in \psi$ is a technically feasible combination, then $(\boldsymbol{x}', \boldsymbol{y}') \in \psi$, with $\boldsymbol{x}' \geq \boldsymbol{x}$ and $\boldsymbol{y}' \geq \boldsymbol{y}$, is also feasible technologically; b) data 'envelopment', that is, $(\boldsymbol{x}_i,\boldsymbol{y}_i) \in \psi$, $i = 1,...,n$; c) minimal extrapolation, which states that among all the possible subsets of R_+^{m+s} that satisfy the previous postulates, the subset provided by DEA is the smallest one (see Ref. [37]) (this last condition can also be seen as the application of the principle of parsimony or Ockham's razor); and, finally, d) convexity, which means that if $(\boldsymbol{x},\boldsymbol{y}),(\boldsymbol{x}',\boldsymbol{y}') \in \psi$, then $\lambda(\boldsymbol{x},\boldsymbol{y}) + (1-\lambda)(\boldsymbol{x}',\boldsymbol{y}') \in \psi$, for all $\lambda \in [0,1]$.

Among all the technical efficiency measures in the literature, there is a family that stands out for its interpretability and fulfilment of interesting properties. We are referring to the radial measures. In particular, the output-oriented radial measure evaluates each DMU$_o$ by equi-proportionally augmenting the outputs as much as possible while inputs remain constant. Many studies justify the adoption of an output orientation when evaluating the performance of educational organizations [38–43]. Taking advantage of the fact that the DEA technology corresponds to a polyhedral set, this measure can be determined through the following linear optimization model.

$$\begin{aligned} \varphi^{DEA}(\boldsymbol{x}_o,\boldsymbol{y}_o) = \max_{s.t.} \quad & \varphi \\ & \sum_{i=1}^{n} \lambda_i x_{ji} \leq \varphi x_{jo}, \quad j = 1,...,m, \\ & \sum_{i=1}^{n} \lambda_i y_{ri} \geq y_{ro}, \quad r = 1,...,s, \\ & \sum_{i=1}^{n} \lambda_i = 1, \\ & \lambda_i \geq 0, \quad i = 1,...,n. \end{aligned} \tag{1}$$

Using this measure, DMU$_o$ is technically efficient—i.e., belongs to the performance (technological) frontier—if $\varphi^{DEA}(\boldsymbol{x}_o,\boldsymbol{y}_o) = 1$. Otherwise, i.e., $\varphi^{DEA}(\boldsymbol{x}_o,\boldsymbol{y}_o) > 1$, DMU$_o$ is classified as technically inefficient. Additionally, model (1) provides benchmarking information. The output targets are calculated as $\varphi^{DEA}(\boldsymbol{x}_o,\boldsymbol{y}_o)\cdot y_{ro}$, for $r = 1,...,s$, and the reference benchmarks for DMU$_o$ are identified through the optimal values of the decision variables lambda. In particular, $\lambda_i^* \neq 0$ identifies DMU$_i$ as benchmark for the assessed unit DMU$_o$.

The flexibility of DEA can be further increased by removing the postulate of convexity (see Ref. [44] or [45]). In that case, the technique is known as Free Disposal Hull [46] and yields a stepwise efficient frontier in contrast to DEA, which produces a piece-wise linear border of the technology. The output-oriented radial measure can be also calculated under FDH by using model (1) but substituting the last constraint with $\lambda_i \in \{0,1\}$, $i = 1, ..., n$, whose efficiency for DMU$_o$ is denoted by $\varphi^{FDH}(\boldsymbol{x}_o,\boldsymbol{y}_o)$.

Regarding envelopment methods, and from a statistical point of view, overfitting is a problem that happens when you have a perfect fit of your model on the data sample. When this occurs, the model unfortunately cannot perform accurately against unseen data, which is usually related to a large generalization error. In words of Hastie et al. [47], p. 221), "… *a model with zero training error is overfit to the training data and will typically generalize poorly*." Standard machine learning techniques aim to identify the actual function that is behind the data generating process (see, e.g. Ref. [48]). If the precise equilibrium is struck between the ability of the model to learn any dataset without error and the accuracy achieved on a particular dataset (the observations), then an appropriate estimation of the underlying function being approximated will be attained. This ability to learn any possible dataset is linked to the notion of generalization error (also called out-of-sample error in the literature). The theoretical generalization error of a model cannot be calculated in general, but it may be approximated by resorting to test samples or cross-validation. In this context, envelopment techniques like FDH and DEA, which put the efficient frontier as close as possible to the data sample due to the minimal extrapolation principle (particularly FDH that envelops the data more tightly), can correctly measure efficiency for a particular set of observations (DMUs) following a sample-specific-based evaluation, but, at the same time, suffer from overfitting. This last feature limits its inferential capability, at least for small data samples, a point that is important when one of the objectives of the study is saying something about the underlying function behind the data generating process that produced the observations. One direct impact of this overfitting problem on the results determined through FDH and DEA is that an important part of the DMUs under evaluation turn out to be technically efficient (see the discussion in the introduction and the results presented in the empirical section where all DMUs are deemed efficient under the FDH approach). A problem that can be aggravated if the relationship between the number of variables (inputs and outputs) and the sample size is inadequate (see, e.g. Ref. [49]). Some recent approaches have attempted to solve the overfitting problem in efficiency evaluation by tailoring machine learning techniques. One outstanding proposal is that known as Efficiency Analysis Trees which we describe in the next subsection.

2.2. Non-convex and convex efficiency analysis trees

A recent technique related to stepwise functions for estimating efficient frontiers by machine learning techniques is Efficiency Analysis Trees (EAT) (see Ref. [27,28]). EAT is inspired in the Classification And Regression Trees proposed by Breiman et al. [29]. EAT builds technologies that satisfy desirable theoretical postulates as FDH except minimal extrapolation; that is, free disposability of inputs and outputs, 'data envelopment', i.e., contains all the observations, and, if demanded, convexity, giving raise to CEAT. Additionally, by abandoning minimum extrapolation, EAT provides a non-overfitted estimation of the underlying production possibility set. However, both EAT and FDH build stepwise surfaces as estimates of production functions. In Fig. 1, we show a graphical example of these two techniques in action.

Next, we detail the main stages of the algorithm that allows determining the EAT estimate of the production possibility set. Given the data sample (considered learning sample) $\aleph = \{(\boldsymbol{x}_i,\boldsymbol{y}_i)\}_{i=1,...,n}$, the first node of the tree, t_1, contains all the observations and is divided into two child nodes, which will be also split in subsequent stages of a recursive algorithm. In this regard, the general splitting step is as follows. Let t be the node to be split. Node t contains a data subset of $\aleph$. Then, the algorithm chooses input j.$j = 1...,m$, and a threshold value for this input, $s_j \in S_j$. The criterion employed for this choice is grounded on minimizing the sum of the Mean Square Error (MSE) linked to the data belonging to the left child node (the observations that satisfy $x_j < s_j$) and the MSE determined for the data that belongs to the right child node (the observations that meet $x_j \geq s_j$). Mathematically speaking, the split consists in selecting the best combination (x_j^*, s_j^*) that minimizes the following

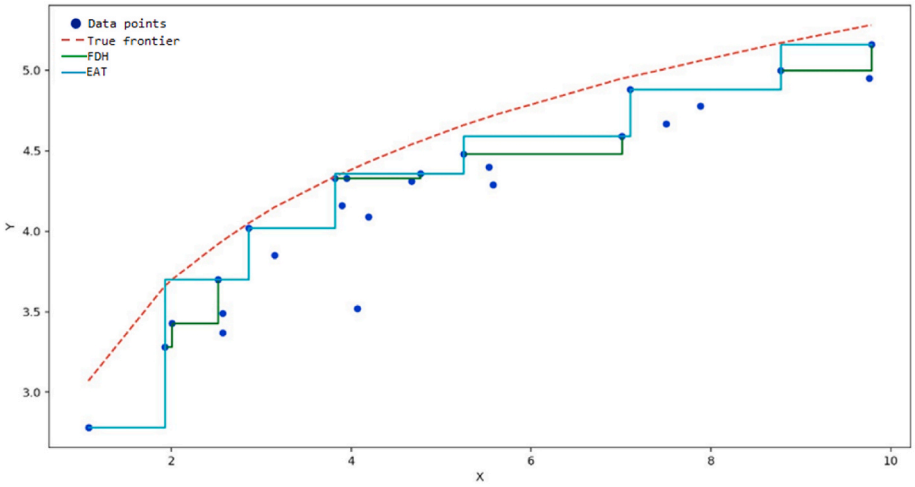


Fig. 1. Graphical illustration of EAT and FDH for a single input and a single output.
Source: own elaboration.

expression:

$$\min\left(R(t_L)+R(t_R)\right)=\frac{1}{n}\sum_{(\boldsymbol{x}_i,\boldsymbol{y}_i)\in t_R}\sum_{r=1}^{s}\left(y_{ri}-y_r(t_L)\right)^2+\frac{1}{n}\sum_{(\boldsymbol{x}_i,\boldsymbol{y}_i)\in t_R}\sum_{r=1}^{s}\left(y_{ri}-y_r(t_R)\right)^2 \tag{2}$$

where $y_r(t)$ denotes the estimation of the r-th output at node t. As for the number of nodes generated, every recursive algorithm needs a stopping rule. In the case of EAT, the stopping rule is $n(t)\leq n_{\min}$, where $n(t)$ denotes the number of observations at node t. Regarding $n_{\min}$, it is a parameter to be tuned and usually takes values among $\{5,10,15,20\}$. Put in words, a node is terminal and is no further split if the number of observations that belong to that node is less or equal than $n_{\min}$.

An important issue in EAT methods from a frontier perspective is the definition of $y_r(t)$ to ensure free disposability of inputs and outputs as well as the enveloping property. The satisfaction of the enveloping property, that is, the estimated technology contains all the observations (at least those belonging to the learning sample), may be easily fulfilled. To do that, the basic idea is to play with the notion of maximum rather than the mean of the output values (as standard classification and regression trees methods do). However, the satisfaction of free disposability while the tree structure is being created is challenging. In the EAT algorithm, the property is satisfied in streaming. After implementing each split, a region in the input space is identified: the "support" of node t, which is defined as $R_t=\{\boldsymbol{x}\in R_+^m: a_{jt}\leq x_j<b_{jt}, j=1,\ldots,m\}$. The values a_{jt} and b_{jt} come from the (optimal) thresholds chosen during the splitting method. Next, we state the notion of (input) Pareto-dominance linked to nodes. Let $k=1,\ldots,K$ be the splits completed at a certain point of the algorithm and $T_k(\aleph)$ the tree created after the k-th split, with $\widetilde{T}_k(\aleph)$ representing the set of terminal nodes in such tree structure $T_k(\aleph)$. Furthermore, let $t^*\in\widetilde{T}_k(\aleph)$ be a specific node to be split at this moment. In this regard, let $T(k|t^*\rightarrow t_L,t_R)$ be the tree associated with this step where node t^* is split into nodes t_L and t_R. Then, given a generic node t, the set of (input) Pareto-dominant nodes is defined as $I_{T_k(\aleph)}(t)=\{t'\in\widetilde{T}_k(\aleph)\backslash t:\exists\boldsymbol{x}\in R_t,\ \exists\boldsymbol{x}'\in R_{t'}\ \text{such that}\ \boldsymbol{x}'\leq\boldsymbol{x}\}$. This notion is the clue for satisfying the property of free disposability. This definition is linked to the output estimation process at each node to be split. Specifically, $y_r(t_R)=y_r(t^*)$, $r=1,\ldots,s$. That is, the output estimation of the right child node coincides with the output estimation of its parent node. Regarding the output estimation of the left child node, it is stated as follows:

$$y_r(t_L)=\max\left\{\max\left\{y_{rj}:\left(\boldsymbol{x}_j,\boldsymbol{y}_j\right)\in t_L\right\},y_r\left(I_{T(k|t^*\rightarrow t_L,t_R)}(t_L)\right)\right\},r=1,\ldots,s, \tag{3}$$

where $y_r(I_{T(k|t^*\rightarrow t_L,t_R)}(t_L))=\max\{y_r(t'):t'\in I_{T(k|t^*\rightarrow t_L,t_R)}(t_L)\}$.

Put in words, the output estimation of the left child node, $y_r(t_L)$, depends on the maximum value observed for the r-th output over the sample data that belongs to the left child node and the maximum value determined over the output estimations corresponding to the nodes that Pareto-dominate node t_L.

After applying the EAT algorithm, we get a tree structure that may be too 'deep', using the terminology utilized in regression trees. Unfortunately, deep trees usually suffer from the same problem as FDH, that is, overfitting. A priori, there is not a simple solution to overcome this drawback in the case of FDH. However, in the case of EAT, due to its tree structure, it is possible to implement a solution inspired in the pruning process introduced by Breiman et al. [29]. Accordingly, let $T(\aleph)$ be the final tree structure after the application of the EAT algorithm and the pruning process. Also, let $\boldsymbol{d}^{T(\aleph)}(\boldsymbol{x})$ be the multidimensional output estimator defined from the tree $T(\aleph)$, that is, $d_r^{T(\aleph)}(\boldsymbol{x})=\sum_{t\in T(\aleph)}y_r(t)I(\boldsymbol{x}\in t)$, for all $r=1,\ldots,s$, with $I(\cdot)$ being the indication function. From $\boldsymbol{d}^{T(\aleph)}(\boldsymbol{x})$, the estimator of the production possibility set associated with EAT is defined as:

$$\widehat{\Psi}_{T(\aleph)}=\left\{(\boldsymbol{x},\boldsymbol{y})\in R_+^{m+s}:\boldsymbol{y}\leq\boldsymbol{d}^{T(\aleph)}(\boldsymbol{x})\right\} \tag{4}$$

As Esteve et al. [27] proved, $\widehat{\Psi}_{T(\aleph)}$ contains all the observations and meets free disposability in inputs and outputs. Additionally, $\widehat{\Psi}_{T(\aleph)}$ does not satisfy minimal extrapolation, what implies that $\widehat{\Psi}_{T(\aleph)}$ contains, as a subset, the technology generated by FDH. Moreover, the efficient frontier provided by EAT has a staggered shape. All these features made the authors of the EAT technique to claim that Efficiency Analysis Trees can be seen as a 'pruned' version of FDH, which avoids overfitting, or as a FDH-type out-of-sample estimator of the underlying technology, using a terminology that belongs to the field of machine learning.

In the case of EAT, if we want to determine the output-oriented radial measure—counterpart to the FDH model (1) with $\lambda_i\in\{0,1\}$, $i=1,\ldots,n$—to gauge technical efficiency of input-output bundle $(\boldsymbol{x}_o,\boldsymbol{y}_o)$, we must obtain an optimal solution of the following linear program:

$$\begin{aligned}\varphi^{EAT}(\boldsymbol{x}_o,\boldsymbol{y}_o) = \quad & \max \quad && \varphi \\ & s.t. \\ & && \sum_{t\in\widetilde{T}(\aleph)} \lambda_t a_{jt} \leq \varphi x_{jo}, && j=1,...,m, \\ & && \sum_{t\in\widetilde{T}(\aleph)} \lambda_t d_r^{T(\aleph)}(\boldsymbol{a}_t) \geq y_{ro}, && r=1,...,s, \\ & && \sum_{t\in\widetilde{T}(\aleph)} \lambda_t = 1, \\ & && \lambda_t \in \{0,1\}, && t\in\widetilde{T}(\aleph).\end{aligned} \tag{5}$$

The difference between model (5) and the typical optimization model linked to FDH (in its enveloping form) is the left-hand side of the first two groups of constraints. Instead of using the information of inputs and outputs of the data sample, the EAT model resorts to the points $\boldsymbol{a}_t$ associated with the support of each node $t\in\widetilde{T}(\aleph)$, that is, each terminal node of the final tree. To see the optimizations programs that must be solved in the case of other (non-radial) technical efficiency measures we refer the reader to Aparicio et al. [28].

Another advantage of EAT against FDH is that, as we show in the empirical application to Colombian private HEIs, the former is able to provide a graphical representation of the efficient frontier of the technology even for high dimensions (a large number of inputs and outputs). The reason is that EAT is linked to the building of a tree structure. The terminal nodes of that structure provide information on the shape of the efficient frontier regardless of the number of variables. Several examples can be found in Aparicio et al. [28] for different real-world applications.

Additionally, to complete the modelling of the technology using EAT methods, we may use the stepwise frontier estimated by EAT as the base to determine a piece-wise linear estimation of the efficient frontier under the assumption of convexity of the production possibility set. To do that, it is enough to substitute the last constraint in (5) with the restriction $\lambda_t \geq 0$, $t\in\widetilde{T}(\aleph)$ (see Ref. [28]). The same idea is behind the relationship between FDH and DEA. When the frontier linked to EAT is 'convexified', the technique is known as CEAT (Convexified Efficiency Analysis Trees), which is represented by

$$conv\left(\widehat{\Psi}_{T(\aleph)}\right) = \left\{(\boldsymbol{x},\boldsymbol{y})\in R_+^{m+s} : \boldsymbol{y}\leq \boldsymbol{d}^{T(\aleph)}(\boldsymbol{x}), \sum_{t\in\widetilde{T}^*}\lambda_t = 1, \boldsymbol{\lambda}\geq 0_{T(\aleph)}\right\}. \tag{6}$$

And the output-oriented radial score obtained through it, denoted as $\varphi^{CEAT}(\boldsymbol{x}_o,\boldsymbol{y}_o)$, can be compared with the score determined by DEA —i.e., model (1), since both methods yield technologies that satisfy the same set of postulates, except, once again, minimal extrapolation, which is met only in the case of DEA.

3. The performance of Colombian higher education institutions through efficiency analysis trees

3.1. The Colombian HEI system, statistical sources and chosen input and output variables

In terms of the scale proposed by Trow [50], it could be said that Colombia is in the process of massification of its higher education. From having an elite system in year 2000 (18.7% coverage rate), the country made a transition to having a system close to being massive in 2015 (49% coverage rate) [51]. The privatization of higher education has been a key determinant in this process of massification [52]. As argued above, Colombia is the third country with the largest share of private spending on HEIs with respect to GDP, being only surpassed by the United States and Chile [53]. At the end of 2018, the Colombian higher education system was made up of 301 HEIs. 71.1% of these HEIs belong to the private sector. According to Uribe [54] and Pineda and Celis [34] the growth of private HEIs in Colombia is a result of the lack of public resources for higher education and the formulation of policies that encourage the growth and expansion of private institutions.

To characterize private HEIs in Colombia and assess their performance, we select those variables that are directly related to the managerial strategies applied by them and which represent sources of institutional differentiation [55,56], while also explaining the market share to which they are directed [57]. The database we have gathered combines two sources of data (see Table 1). On the one hand, we consider the publicly available data provided by the Ministry of Education of Colombia [58] for year 2019. This includes the following input variables: (i) number of full professors (x_1); (ii) number of full-time lecturers with PhD degree (x_2); (iii) number of full-time lecturers per active program (x_3); (iv) number of undergraduate programs (x_4); (v) number of master and PhD programs (x_5); (vi) number of students enrolled in undergraduate programs (x_6); (vii) number of students enrolled in master and PhD programs (x_7); and the following output variables: (i) number of students graduated in undergraduate programs (y_1); and (ii) number of students graduated in master and PhD programs (y_2). Year 2019 is chosen to avoid the impact that the Covid-19 pandemic may have had on the performance of Colombian HEIs. The total population of Colombian private HEIs for which the data from the Ministry of Education were available for year 2018 amounts to 171.

On the other hand, we consider the rankings provided by the Colombian consulting firm Sapiens, which has published reports and classifications derived from the analysis of the dynamics of HEIs in Colombia for more than 20 years.[2] These include the following rankings which are added to the previous outputs: (iii) Ranking of Colombian HEIs according to their level of technological development and innovation (y_3); (iv) Ranking of Colombian HEIs according to their level of production of scientific articles (y_4); (v) Ranking of Colombian HEIs according to their level of generation of new knowledge (y_5); and (vi) Ranking of Colombian HEIs according to the level of social appropriation of their knowledge, IPRs (y_6).[3] In these cases, the latest available year is chosen for each of these rankings, which range between 2018 and 2021. The size of the final sample of Colombian private HEIs for which the total amount of all indicators could be gathered amounts to 144. As it can be observed in Table 1, our final model includes 7 input variables and 6 output variables. All input and output variables, x_i, and y_i, are normalized by the maximum observed value (i.e., $\frac{z_i}{\max(z_i)}\times 100$), $z_i = x_i$, y_i, so their value ranges in all cases between 0 and 100.[4]

In the following subsections, we describe the results of the EAT technique when growing the tree approximating the benchmark technology and, once the efficiency scores of the Colombian HEIs are calculated under the non-convexity and convexity assumptions, compare the attained results with those corresponding to their FDH and DEA counterparts.

3.2. Tree growth representing Colombian HEIs

Here we report the results corresponding to the benchmark technology obtained when applying the EAT method. We follow the algorithm described in section 2.2, predicting the maximum of the output vector by splitting the observed sample conditioned to the values of the observed inputs. Fig. 2 presents the complete tree that has been obtained by applying the algorithm previously described. When performing the node splitting, the algorithm follows the heuristic method such that the number of the (input) Pareto-dominants nodes is low, which is normally

[2] See: https://www.srg.com.co/conocenos/.

[3] To find the details on the indicators that are considered to elaborate these four rankings see, respectively: https://www.srg.com.co/dtisapiens/metodologia/; https://www.srg.com.co/artsapiens/metodologia/; https://www.srg.com.co/gncsapiens/metodologia/; and https://www.srg.com.co/noticias/reporte-asc-sapiens-2018/.

[4] Appendix 1 provides the descriptive statistics of the variables considered without normalization.

Table 1
Variables considered and descriptive statistics.

		Variable	Average	Median	Max	Min	Stand. Dev.
Inputs	Lecturers	x_1 – # full professors	13.007	6.991	100.0	0.0	16.728
		x_2 – # full-time lecturers with PhD degree	6.812	1.985	100.0	0.0	13.864
		x_3 – # full-time lecturers per active program	22.055	18.607	100.0	0.0	17.022
	Diversity of programs	x_4 – # undergraduate programs	17.426	14.433	100.0	0.0	13.860
		x_5 – # master and PhD programs	11.756	5.631	100.0	0.0	16.450
	Students	x_6 – # students enrolled in undergraduate programs	6.166	4.081	100.0	0.0	9.755
		x_7 – # students enrolled in master and PhD programs	8.675	3.733	100.0	0.0	14.154
Outputs	Training	y_1 – # students graduated in undergraduate programs	5.947	3.475	100.0	0.0	9.820
		y_2 – # students graduated in master and PhD programs	10.034	4.667	100.0	0.0	15.224
	Scientific – technological production	y_3 – Ranking DTI-Sapiens	8.224	4.268	100.0	0.0	13.312
		y_4 – Ranking ART-Sapiens	7.761	2.469	100.0	0.0	14.914
		y_5 – Ranking GNC-Sapiens	15.118	7.879	100.0	0.0	19.286
	Impact on the territory	y_6 – Ranking ASC-Sapiens	13.758	8.000	100.0	0.0	15.968

Source: own elaboration.

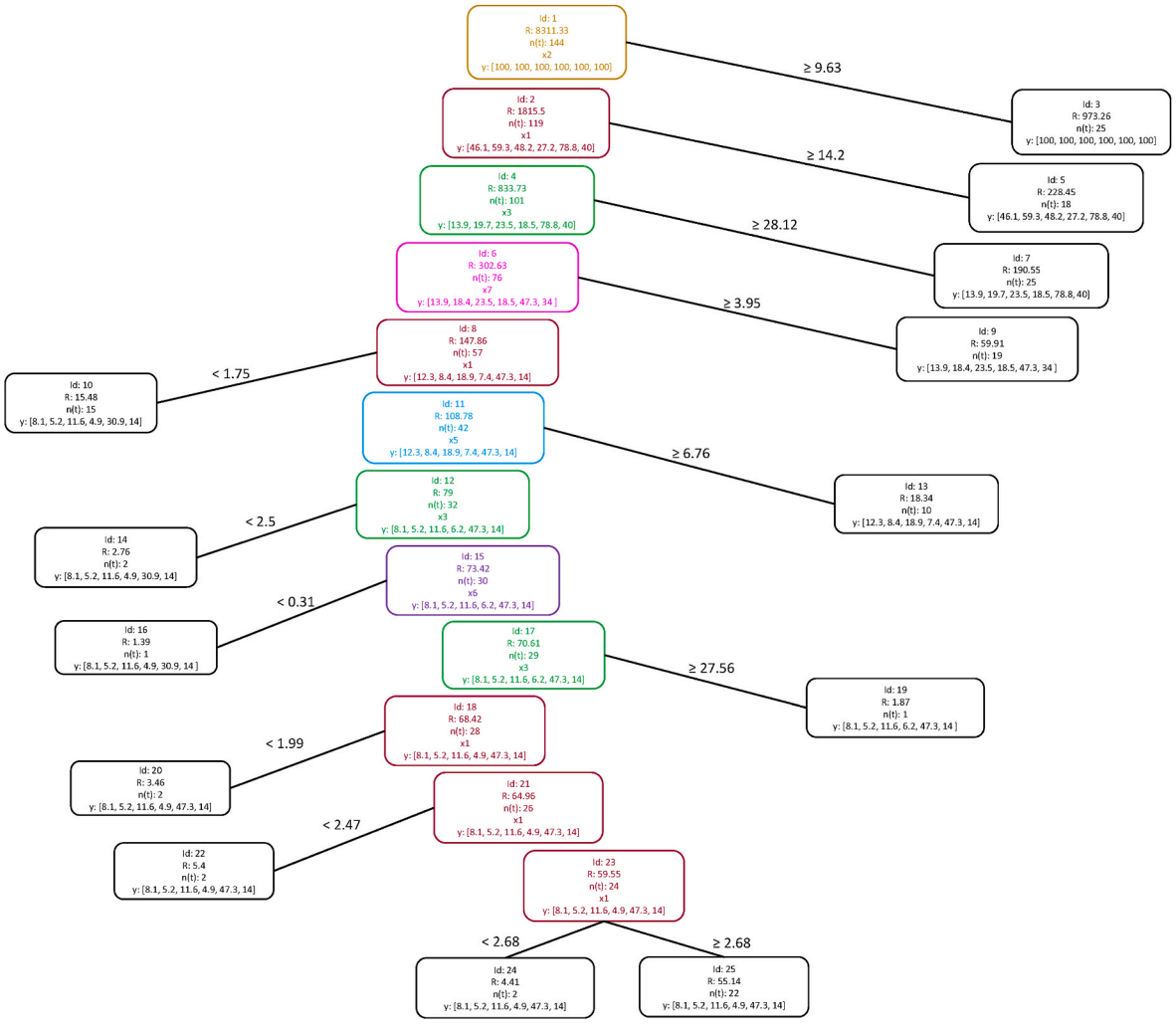


Fig. 2. Efficiency analysis tree.

associated with the left-branch (child) nodes. The tree, as well as the scores corresponding to the different EAT, CEAT and standard FDH and DEA efficiency measures have been obtained running the *eat* package implemented in R as described in Esteve et al. [59].

We start the description of the tree with the root node at the top. We observe that the key variable that minimizes the sum of the squared errors for the left and right branches (or children) according to expression (2), is the number of full-time lecturers with PhD degree (x_2), with the right branch containing all HEIs (25) with $x_2 \geq 9.63$, which is the optimal threshold splitting the sample. Therefore, the left branch, t_2, of the subtree T_1 composed by $\{t_1, t_1, \ldots, t_{25}\}$ includes the remaining 119 HEIs with $x_2 < 9.63$. We observe that the right node t_3 (Id = 3), inherits the maximum observed values of the output vector of the root node: $y(t_0) = y(t_3) = [100, 100, 100, 100, 100, 100]$, as presented in expression (3). On the contrary, the left branch t_2 (Id = 2) shows the output estimation $y(t_2) = [46.1, 59.3, 48.2, 27.2, 78.8, 40]$, corresponding to (lower valued) inputs—once the split is performed, which constitute the 'support' of node t_2. Consequently, in each node, we find an

identification number (Id), the mean squared error (R), the sample size, the split executed for input j and threshold, $s_j \in S_j$, and, finally, the estimated output. Note that, as anticipated and from a data visualization perspective, it is possible to illustrate the whole production technology, as opposed to FDH and DEA, which are limited to no more than three dimensions.

Following the algorithm, the tree is subsequently grown (or expanded) by way of consecutive k+1 splits under the Pareto-dominance condition, which yields successive nodes $\widetilde{T}_{k+1}$. In this empirical application as many as 24 nodes are calculated (excluding the root node). In Fig. 2 those nodes that are partitioned are represented as colored rectangles, while the 13 end nodes (leaves) are represented as black rectangles. In these 13 terminal nodes, normally located at the bottom or the sides of the tree, the final estimation of the response variable (the output) is also reported. In this way, we can visualize the stepwise frontier by visiting the different branches of the tree from top to bottom. For the case of terminal nodes, the splitting process ends because either the Pareto-dominance criteria cannot be observed, so no more partitions are possible, or the stopping rule is met (in our case, $n_{\min} = 10$ after executing a test sample-based validation method), and then it is marked as a leaf node of the tree. We remark that the tree represented in Fig. 2 is the result of applying the process on the data described in section 2.2, yielding the 'pruned' (or optimal) version of the tree, T. The process prevents the overfitting of the data and allows identifying an accurate version of the performance frontier.

3.2.1. Productive efficiency in Colombian HEIs: output-oriented EAT vs. FDH and CEAT vs. DEA

We now compare the results obtained when calculating the radial output-oriented efficiency measure using the tree (EAT and CEAT) and envelopment frontiers (FDH and DEA). Once we have generated the optimal tree, inducing the production possibility set for each supporting vector of inputs, we first calculate the EAT efficiency scores and compare them to their FDH counterparts. As previously argued, and following common practice in efficiency studies in the educational sector, we adopt an output-oriented approach, by which HEIs aim at improving their academic results given the available resources.

Therefore, we start comparing the models for the non-convex estimation of the technology. Table 2 presents the individual efficiency scores of the ten best and ten worst performing private HEIs in Colombia, as well as their descriptive statistics for the whole sample at the bottom of the table. HEIs are ordered from best performing (i.e., smallest scores equal one) to worst performing (i.e., largest scores), first according to their EAT scores, and then based on their CEAT values. The EAT output-oriented efficiency scores, $\varphi^{EAT}(\boldsymbol{x}_o,\boldsymbol{y}_o) := \max\{\varphi_o \in R : (\boldsymbol{x}_o, \varphi_o\boldsymbol{y}_o) \in \widehat{\Psi}_T\}$, with $\lambda_i \in \{0,1\}$, $i = 1,...,n$, empirically calculated through program (5) are reported in the second column, while their FDH counterparts, $\varphi^{FDH}(\boldsymbol{x}_o,\boldsymbol{y}_o) := \max\{\varphi_o \in R : (\boldsymbol{x}_o,\varphi_o\boldsymbol{y}_o) \in \widehat{\Psi}_{FDH}\}$, calculated through program (1) are presented in the third column. All ten best performing schools are EAT and FDH efficient with unitary efficiency scores, while the scores of the ten worst performing schools illustrate that FDH technology is enveloped by the EAT technology—see Proposition 2 (iii) in Aparicio et al. [28]. Consequently, the efficiency scores of the former are either equal or smaller that the latter: $\varphi^{EAT}(\boldsymbol{x}_o, \boldsymbol{y}_o) \geq \varphi^{FDH}(\boldsymbol{x}_o,\boldsymbol{y}_o)$. Although not entirely surprising, we observed that all the considered HEIs are efficient under the FDH envelopment approach.

This result illustrates the already discussed overfitting shortcomings of traditional envelopment techniques. Indeed, all Colombian private HEIs are efficient under the FDH approach because of the minimum extrapolation requirements of this approach and the many input and output dimensions involved in the evaluation of educational systems. This renders the ranking and benchmarking process driving any efficiency analysis meaningless, providing further justification for techniques like Efficiency Analysis Trees which do not endure this drawback. To further illustrate this limitation, we portray the pairwise box-plots and kernel density distributions of the calculated efficiency scores in the left-hand side of Fig. 3.

Table 2
Comparing output-oriented efficiency scores: EAT vs. FDH, and CEAT vs. DEA. Selected observations.

	Non-Convex		Convex	
Score	$\varphi^{EAT}(\boldsymbol{x}_o,\boldsymbol{y}_o)$	$\varphi^{FDH}(\boldsymbol{x}_o,\boldsymbol{y}_o)$	$\varphi^{CEAT}(\boldsymbol{x}_o,\boldsymbol{y}_o)$	$\varphi^{DEA}(\boldsymbol{x}_o,\boldsymbol{y}_o)$
HEIs				
1	1.000	1.000	1.000	1.000
5	1.000	1.000	1.000	1.000
89	1.000	1.000	1.000	1.000
90	1.000	1.000	1.000	1.000
101	1.000	1.000	1.000	1.000
131	1.000	1.000	1.000	1.000
109	1.000	1.000	1.007	1.000
104	1.000	1.000	1.130	1.000
3	1.000	1.000	1.136	1.000
43	1.000	1.000	1.141	1.000
140	5.555	1.000	11.793	1.000
7	5.831	1.000	11.874	1.679
55	6.231	1.000	6.231	1.060
82	7.000	1.000	9.502	1.000
128	7.000	1.000	14.896	2.073
135	7.309	1.000	7.763	1.000
132	7.501	1.000	11.716	1.000
129	8.327	1.000	18.015	1.000
138	9.294	1.000	20.469	1.000
120	13.256	1.000	23.948	1.362
Average	2.394	1.000	5.349	1.082
Median	1.932	1.000	4.383	1.000
Max.	13.256	1.000	23.948	2.073
Min.	1.000	1.000	1.000	1.000
Stand. Dev.	1.814	0.000	4.055	0.166

Source: own elaboration.

We may now undertake the comparison between the convexified versions of EAT and FDH For this purpose, we calculate the output-oriented CEAT and DEA efficiency scores corresponding to $\varphi^{CEAT}(\boldsymbol{x}_o,\boldsymbol{y}_o) := \max\{\varphi_o \in R : (\boldsymbol{x}_o, \varphi_o\boldsymbol{y}_o) \in conv(\widehat{\Psi}_{T(8)})\}$ and $\varphi^{DEA}(\boldsymbol{x}_o,\boldsymbol{y}_o) := \max\{\varphi_o \in R : (\boldsymbol{x}_o,\varphi_o\boldsymbol{y}_o) \in \widehat{\Psi}_{DEA}\}$, respectively. The fourth and fifth rows of Table 2 present the efficiency scores for both series. We observe once again that a large proportion of HEIs are efficient in the DEA approach: 97, representing 67.36% of the sample. We also find within convex models that the DEA efficiency scores are either equal or smaller than their CEAT counterparts: $\varphi^{CEAT}(\boldsymbol{x}_o,\boldsymbol{y}_o) \geq \varphi^{DEA}(\boldsymbol{x}_o,\boldsymbol{y}_o)$. The right-hand side of Fig. 3 compares the box-plots and kernel density functions of both distributions. We observe that both series are hardly comparable due to the large number of efficient observations in the traditional approach. On this occasion it is possible to calculate Spearman's correlation, although the value is rather low at $\rho(\varphi^{CEAT},\varphi^{DEA}) = 0.391$ ($p = 0.0000$), showing that the rankings underlying both distributions are substantially different. This result is also observed even if only the values of inefficient HEIs under DEA were compared to their CEAT counterparts. The conclusion is that DEA offers an optimistic assessment of the efficiency levels of the Colombian private HEIs, which portraits an incorrect picture of the systems' performance, despite the obvious attractiveness to university managers and government officials.

Given that the only efficiency measurement that can be meaningfully done is that offered by the tree methodology we now compare the scores *between* non-convex EAT and convex CEAT. It is observed that, individually, $\varphi^{CEAT}(\boldsymbol{x}_o,\boldsymbol{y}_o) \geq \varphi^{EAT}(\boldsymbol{x}_o, \boldsymbol{y}_o)$, while the mean average inefficiency in the convex approach more than doubles that of its non-convex counterpart: $\overline{\varphi}^{CEAT} = 5.35 > \overline{\varphi}^{EAT} = 2.35$. Also, looking at the box-plots in Fig. 3, the dispersion of the distributions within the inter-quartile ranges represented by the whiskers appears to be relatively high. As for the number of efficient observations, we see that under the CEAT approach it is substantially smaller than that corresponding to

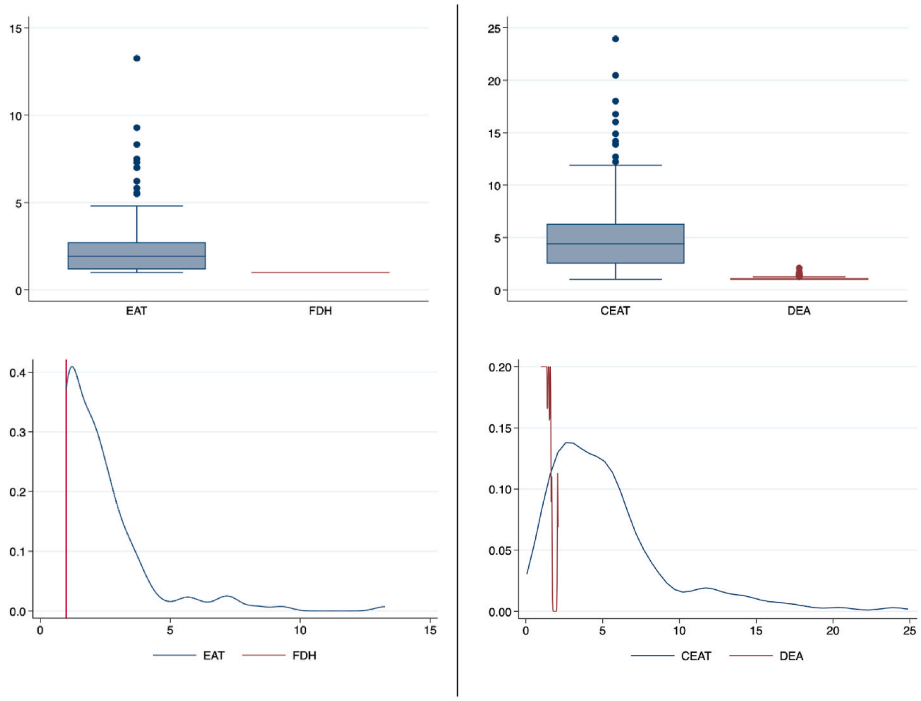


Fig. 3. Box-plots and kernel distributions: EAT vs. FDH and CEAT vs. DEA. Source: own elaboration.

EAT, i.e., convex models allow for a better discrimination among observations. Indeed, while just 6 HEIs are efficient under CEAT (4.16% of the sample), 28 (19.4%) are efficient under EAT. Although the ranking compatibility of both distributions is relatively high at $\rho(\varphi^{EAT}, \varphi^{CEAT}) = 0.6223$, these results show the advantage of using the new methods, particularly Convexified Efficiency Analysis Trees, CEAT, to identify production peers and ranking observations—in passing, we remark that the number of efficient observations under non-convex EAT (28) is smaller than for the convex DEA (97).

The main conclusion from this section is that standard envelopment techniques like FDH and DEA are unable to discriminate Colombian private HEIs according to their performance. On the contrary, tree methods can unveil existing inefficiencies and provide relevant strategic and policy guidelines aimed at improving the performance of universities, colleges, business schools, etc. For this reason, in the next section we only focus on the results obtained from the tree methodology to identify relevant benchmarks and the input and output adjustments that can lead to efficiency enhancement strategies.

4. Benchmarking: how can Colombian HEIs improve performance?

In this section we show how to practice benchmarking using EAT and CEAT. We first identify the HEIs that can potentially serve as benchmarks by defining the performance frontier at each node. The relationship between the levels of outputs produced and input usage at each node is also discussed, interpreting the size of operations in the different regions of the tree. Afterwards, for the least and most inefficient HEIs at each node, we show how individual observations can learn from their benchmark peers in order to reduce their inefficiency. The benchmark peer is identified as the one presenting the closest input-output structure of the HEI under evaluation, once it has been projected to the frontier radially (i.e., once the outputs have been increased proportionally according to the HEI's efficiency score). The criterion is that of the minimum (Euclidian) distance. We then suggest individual managerial recommendations to improve HEI performance by exploiting not only the equiproportional gains in output production that can be achieved by solving radial inefficiencies, but also additional changes in the output mix. We end this section offering summary statistics of the HEIs defining the production frontier; namely their frequency when serving as benchmarks on the frontier in terms of their specific node and the whole tree.

4.1. Identifying potential benchmarks and relevant inputs on the performance frontier

Following the structure of the tree presented in Fig. 2, the first task is to identify, for each of the 13 leaf nodes, the HEIs that can serve as potential benchmarks. We denote by $\widehat{\boldsymbol{y}}_i^{*EAT} = \boldsymbol{y}_i^* \cdot \varphi_i^{EAT}$ the vector of output values for the HEI's identified as peers in the EAT analysis, and by $\widehat{\boldsymbol{y}}_i^{*CEAT} = \boldsymbol{y}_i^* \cdot \varphi_i^{CEAT}$ those under CEAT. Obvious candidates to become actual peers are those observations that define the benchmark frontier in each node. These observations have an efficiency score of one under the EAT or CEAT approach, $\varphi_i^{EAT} = 1$ or $\varphi_i^{CEAT} = 1$, and therefore their observed values are also the optimal ones. However, as we discuss in what follows, there are nodes that do not harbor efficient observations, because the performance frontier across their inputs space is defined by units belonging to other nodes. In this case, the criterion for selecting the reference peer within the node is the one with the lowest inefficiency, whose efficient projection is defined as above, but with $\varphi_i^{EAT} > 1$ or $\varphi_i^{CEAT} > 1$.

4.1.1. Non-convex EAT analysis

Table 3 presents the collection of best performing peers per node in the EAT analysis. Next to the number identifying the node according to Fig. 2 (first column), we report the total number of HEIs included in the node (second column), and differentiate between inefficient and efficient HEIs (third and fourth columns, respectively). Then, in the fifth column we show all HEI's that are candidates to become peers for inefficient HEIs within the node.

We now discuss the characteristic of the main nodes identified in the analysis in terms of the inputs responsible for their formation when growing the tree, and their output levels.

Table 3
Efficiency analysis of Colombian HEIs by nodes. Identified peers. EAT.

Node	HEIs	Inefficient HEIs	Efficient HEIs	Peer HEIs		Efficient Output						Peer statistics		
Nº	Nº	Nº	Nº	Nº	Score φ^{EAT}	$\widehat{y}_1^{*EAT}$	$\widehat{y}_2^{*EAT}$	$\widehat{y}_3^{*EAT}$	$\widehat{y}_4^{*EAT}$	$\widehat{y}_5^{*EAT}$	$\widehat{y}_6^{*EAT}$	Nº	Node %	Tree %
3	25	20	5	1	1.00	19.34	43.52	30.49	65.43	100.00	100.00	16	64.00	11.11
				5	1.00	5.01	100.00	3.05	8.64	72.12	27.60	4	16.00	2.78
				89	1.00	14.26	13.95	100.00	41.98	70.91	24.40	3	12.00	2.08
				90	1.00	8.92	13.31	76.22	100.00	21.21	24.40	1	4.00	0.69
				101	1.00	100.00	51.63	11.59	16.05	78.79	40.00	1	4.00	0.69
5	18	14	4	34	1.00	8.54	8.10	48.17	14.81	23.64	9.60	2	11.11	1.39
				67	1.00	2.70	15.80	3.66	27.16	10.30	23.60	4	22.22	2.78
				77	1.00	46.06	29.92	4.88	6.17	8.48	16.80	6	33.33	4.17
				79	1.00	22.09	59.28	0.61	4.94	18.18	17.60	6	33.33	4.17
7	25	21	4	4	1.00	4.71	10.46	13.41	4.94	9.09	40.00	7	28.00	4.86
				18	1.00	5.94	5.04	7.32	2.47	26.67	40.00	6	24.00	4.17
				45	1.00	5.39	19.70	1.83	1.43	1.82	14.40	10	40.00	6.94
				109	1.00	1.64	1.03	11.59	3.70	78.79	6.80	2	8.00	1.39
9	19	14	5	52	1.00	2.79	7.09	3.05	6.17	20.61	34.00	7	36.84	4.86
				59	1.00	2.35	12.78	10.37	18.52	6.67	12.80	3	15.79	2.08
				69	1.00	2.78	6.71	23.46	2.47	10.91	8.40	3	15.79	2.08
				74	1.00	1.88	18.38	0.61	1.14	7.27	8.40	5	26.32	3.47
				102	1.00	13.94	8.73	0.06	1.43	3.64	5.60	1	5.26	0.69
10	15	13	2	124	1.00	8.09	5.19	11.59	3.70	30.91	7.60	4	26.67	2.78
				126	1.00	0.53	2.79	3.05	4.94	2.42	14.00	11	73.33	7.64
13	10	6	4	40	1.00	1.49	8.42	13.41	0.15	5.45	3.60	2	20.00	1.39
				61	1.00	2.70	2.30	18.90	4.94	23.03	6.00	1	10.00	0.69
				62	1.00	5.57	3.16	2.44	7.41	7.27	12.00	3	30.00	2.08
				97	1.00	12.26	0.59	4.88	3.70	23.03	8.00	4	40.00	2.78
14	2	2	0	110	2.00	0.86	0.23	0.24	4.94	7.27	7.20	2	100.00	1.39
16	1	1	0	141	3.29	0.78	5.19	10.03	0.49	5.98	1.32	1	100.00	0.69
19	1	0	1	98	1.00	4.46	0.23	4.88	6.17	8.48	8.80	1	100.00	0.69
20	2	1	1	131	1.00	0.35	0.02	0.18	0.69	47.27	1.20	2	100.00	1.39
22	2	2	0	125	2.71	3.60	0.41	11.58	0.80	11.51	4.34	2	100.00	1.39
24	2	1	1	113	1.00	2.34	3.02	6.71	4.94	2.42	4.80	2	100.00	1.39
25	22	21	1	76	1.00	2.67	1.84	7.93	4.94	7.27	10.00	25	100.00	17.36

Source: own elaboration.

Node 3. Number of full-time lecturers with PhD degree, x_2. Looking at the first identified node 3, obtained after the split of the root node, we see in Fig. 2 that it includes HEIs with $x_2 \geq 9.63$. A total of 25 HEIs belong to it, out of which 20 are inefficient and 5 efficient (25%). These efficient HEI are #1, #5, #89, #90 and #101, with an efficiency score $\varphi_i^{EAT} = 1$. Next, we report the optimal output values $\widehat{y}_i^{*EAT}$ for each one of these five HEIs. We observe that this node includes efficient observations that coincide with the maximum estimated output values: $\mathbf{y}(t) = 100.00$. For instance, the first observation (HEI #1): "Pontificia Universidad Javeriana", presents the highest value in the last two outputs; i.e., it is at the top of the GNC-Sapiens ranking related to the generation of new knowledge, and the ASC-Sapiens ranking related to the appropriation of own knowledge. The defining input for this node is the number full-time lecturers with PhD degree (x_2). This identifies such staff as a key threshold characterizing the highest output achievements (effectiveness) in the Colombian higher education system. This allows us to conclude that having many faculty members with PhDs characterizes this region of the frontier T obtaining the largest output levels, which provides reassurance for the current human resource policies in favor of recruiting faculty with the highest educational degree. Indeed, since only 25 HEIs have a proportion of faculty members larger than 9.63 percentage points of the maximum observed value, the remaining 119 should look at this input as the most defining one when attempting to 'climb up' the tree towards larger output values. Naturally, achieving large output values is not the same as performing efficiently with the available resources, i.e., once inputs are considered. This can be seen as we move on to other regions of the tree characterizing the frontier.

Node 5. Number of full professors, x_1. The second node split, identified as node 5, includes 18 observations characterized (after the first split) by a lower number of full-time faculty with PhDs, $x_2 < 9.63$, but exhibiting a large proportion of full professors in their staff, $x_1 \geq 14.20$. Although in this node the level of both the highest estimated output values $\mathbf{y}(t)$ and optimal projected values is significantly smaller than in the first node: $[y_1(t) = 46.06, y_2(t) = 59.28, y_3(t) = 48.17, y_4(t) = 27.16, y_5(t) = 8.48, y_6(t) = 40.0]$, we still find a similar number of efficient HEIs, 4, and proportion, 22.2%, than in the first node. These potential peers are HEIs #34, #67, #77, and #79. This reflects that the size of the HEIs in terms of output levels gets smaller as we descend the tree, yet the method is capable of finding a relevant number of comparable HEIs in terms of their inputs and outputs sizes, while ensuring that the efficiency comparison is performed against benchmark peers that exhibit the same characteristics. This node shows that the educational level (PhD), jointly with a higher academic experice and remuneration (full professor), are key attributes when identifying the production frontier, and lead to high teaching and research levels.

Node 7. Number of full-time lecturers per *active program, x_3.* We proceed similarly by descending along the tree and find that the number of nodes with a relatively large number of HEIs continues, i.e., node 7 includes 25 observations, with 4 HEIs being efficient: #4, #18, #45 and #109. As before, this node has HEIs that are fully efficient and therefore efficient output coincide with observed outputs, $\widehat{y}_i^{*EAT} = y_i^*$. The key input value defining this node, number of full-time lecturers per active program, $x_3 \geq 28.12$, stresses once again the stability of faculty members over part-time hires, showing that less expensive pay-roll schemes, associated to part-time lectures, do not attain large output levels.

Node 9. Number of students enrolled in master and PhD programs, x_7. Similar reasoning can be applied to discuss the defining inputs for the ninth node. Specifically, this node includes 19 HEIs characterized, once again, by high quality education since the defining input is the number of students enrolled in master and PhD programs, $x_7 \geq 3.95$. Here we find 5 HEIs that are fully efficient (i.e., #52, #59, #69, #74, #102). However, the output levels are relatively lower than in the previous nodes, showing that these HEIs are characterized by being specialized in niche markets such as MBAs and other executive education (see also

[60]).

Similar reasoning can be applied to discuss the defining inputs and efficient benchmarks for subsequent nodes, whose size in terms of both output values and number of HEIs diminishes as we move down the tree, except for the last sweeping node collecting those universities, colleges and business schools that have not previously classified in previous nodes. Also, there are two nodes, 16 and 19, whose characteristics are so singular in terms of their input values that the methodology splits individual nodes for both of them. In the case of node 16, we find a university specialized in business and health sciences (CORSALUD) with a small number students enrolled in undergraduate programs, $x_6 < 0.31$. For node 19, we find a small HEI (Corporación Universitaria Rafael Nuñez), $27.56 \leq x_3 < 28.12$. In the first case, the university is inefficient, with a score equal to 3.29, showing that such specialization does not result in efficient outcomes from which the Colombian HEIs can learn best practices, while the second one is efficient.

It is worth remarking that in the non-convex EAT approach, almost all nodes have efficient observations. The only exceptions are the previously commented node 16 and nodes 14 and 22 (shaded in grey in Table 3). In these latter two nodes the best performing observations have efficiency scores equal to 2.00 and 2.71. In the following subsection, these units are used as reference peers for the remaining observations included in these nodes.

4.1.2. Convex EAT analysis

We now turn to the efficiency analysis of the convexified frontier that is generated from the tree, whose results are reported in Table 4. The interpretation of the nodes in terms of the input and output levels, as well as the characteristics of the HEIs, does not differ from the ones commented above. However, as emphasized in section 3.3, only six observations are efficient. Table 4 shows that five out the six fully efficient observations cluster in the first split, node 3, reinforcing the idea that the second input, 'number of full-time lecturers with PhD degree, x_2' is key to achieve best performance. Indeed, the HEIs identified as efficient coincide with those of the non-convex DEA approach: #1, #5, #89, #90 and #101, with efficiency scores $\varphi_i^{CEAT} = 1$ — further discussion of this result is presented in what follows. Then, the next efficient HEI can be found in node 20, which includes only two observations.

The remaining nodes do not include efficient observations, which, on the one hand, shows the discriminatory power of the CEAT approach, but, on the other, renders the benchmarking analysis dependable on units that, even if they are high achievers within their nodes, may exhibit relatively large inefficiency values. Again, to identify efficient benchmark peers, we could 'climb up' the tree until an efficient peer can be found. For instance, it would be possible to compare all inefficient HEIs belonging to node 5, to the efficient peers of node 3. However, when the nodes are very far apart, e.g. node 19 from node 3, this would imply that observations are compared to completely dissimilar peers in terms of their inputs and output levels, questioning the validity of the benchmarking exercise from a practical perspective. For example, it is questionable to compare HEIs specialized in tertiary and executive education — node 9 — with general universities like those included in nodes 3 and 5. Here, as anticipated, we favor a homogeneity criterion when comparing HEIs and restrict the set of possible benchmarks to those included in the node. Among these, we choose the HEI that performs better by exhibiting the lowest inefficiency score within the node. In Table 4, these observations are identified as peers, whose output projections on the frontier, also reported, serve as benchmark for the projection of the remaining inefficient observations.

4.2. Individual benchmarking: identifying peers and output changes

Once we have studied the features and characteristics of the nodes generated by the tree analysis and identified the set of potential peers in each of them, we show how to use this information to prescribe strategic managerial guidelines aimed at improving the efficiency of individual units. We start with the non-convex EAT analysis. Table 5 presents the least and most inefficient HEIs for each node and their reference peers. We see that in node 3, HEI #104 is the least inefficient observation with an efficiency score of $\varphi^{EAT}(\boldsymbol{x}_{104}, \boldsymbol{y}_{104}) = 1.13$, while HEI #55 is the most inefficient one with $\varphi^{EAT}(\boldsymbol{x}_{55}, \boldsymbol{y}_{55}) = 6.23$. In case there are several candidates for peers, as in this node, we adopt the principle of least action or minimum adjustment cost, represented by the minimum distance, to identify the corresponding peer. Hence, after calculating the Euclidean distance between the optimal — radially projected — output values of unit o under evaluation and its $i = 1, \ldots, I$ efficient candidates, we choose the closest one. That is, $\min\{d(HEI_o, HEI_1), \ldots, d(HEI_o, HEI_I)\} = \min\{\sqrt{\sum_{r=1}^{S}(\widehat{y}_{r1}^{*EAT} - \widehat{y}_{ro}^{EAT})^2}, \ldots, \sqrt{\sum_{r=1}^{S}(\widehat{y}_{rI}^{*EAT} - \widehat{y}_{ro}^{EAT})^2}\}$. Peer #5 is the closest efficient HEI for #104, with a minimum distance of $d(HEI_{104}, HEI_5) = 75.41$, while the peer corresponding to #55 is #1 with $d(HEI_5, HEI_1) = 68.65$. In this case, since both peers are efficient, the reference output values for benchmarking, $\widehat{y}_{r1}^{*EAT}$, coincide with the observed ones.

The last six columns of Table 5 indicate the amount in which each radially projected individual output of HEI_o must be adjusted to reach

Table 4
Efficiency analysis of Colombian HEIs by nodes. Identified peers. CEAT.

Node	HEIs	Inefficient HEIs	Efficient HEIs	Peer HEIs		Efficient Output						Peer statistics		
Nº	Nº	Nº	Nº	Nº	Score	$\widehat{y}_1^{*CEAT}$	$\widehat{y}_2^{*CEAT}$	$\widehat{y}_3^{*CEAT}$	$\widehat{y}_4^{*CEAT}$	$\widehat{y}_5^{*CEAT}$	$\widehat{y}_6^{*CEAT}$	Nº	% Node	% Tree
				1	1,00	19,34	43,52	30,49	65,43	100,00	100,00	16	64.00	11.11
				5	1.00	5.01	100.00	3.05	8.64	72.12	27.60	4	16.00	2.78
3	25	20	5	89	1.00	14.26	13.95	100.00	41.98	70.91	24.40	3	12.00	2.08
				90	1.00	8.92	13.31	76.22	100.00	21.21	24.40	1	4.00	0.69
				101	1.00	100.00	51.63	11.59	16.05	78.79	40.00	1	4.00	0.69
5	18	18	0	77	1.21	55.59	36.12	5.89	7.45	10.24	20.28	18	100.00	12.50
7	25	25	0	109	1.01	1.65	1.03	11.67	3.73	79.34	6.85	25	100.00	17.36
9	19	19	0	102	1.97	27.48	17.20	0.12	2.82	7.17	11.04	19	100.00	13.19
10	15	15	0	116	2.49	4.20	0.05	12.15	0.49	36.22	30.88	15	100.00	10.42
13	10	10	0	61	2.50	6.76	5.75	47.22	12.34	57.53	14.99	10	100.00	6.94
14	2	2	0	110	6.58	2.82	0.75	0.80	16.25	23.93	23.69	2	100.00	1.39
16	1	1	0	141	7.84	1.86	12.36	23.89	1.16	14.25	3.13	1	100.00	0.69
19	1	1	0	98	6.24	27.81	1.42	30.41	38.49	52.90	54.87	1	100.00	0.69
20	2	1	1	131	1.00	0.35	0.02	0.18	0.69	47.27	1.20	2	100.00	1.39
22	2	2	0	125	5.34	7.09	0.81	22.79	1.58	22.65	8.54	2	100.00	1.39
24	2	2	0	113	4.64	10.84	14.02	31.12	22.91	11.25	22.27	2	100.00	1.39
25	22	22	0	32	4.00	7.93	17.26	31.70	3.75	31.51	23.99	22	100.00	15.28

Source: own elaboration.

Table 5
Actual peers for inefficient HEIs. Selected HEIs. EAT.

Node	Ineffic. HEIs		Peer HEIs		Minimum Distance	Output slacks					
	Nº	Score	Nº	Score		$\widehat{y}_1^* - \widehat{y}_1$	$\widehat{y}_2^* - \widehat{y}_2$	$\widehat{y}_3^* - \widehat{y}_3$	$\widehat{y}_4^* - \widehat{y}_4$	$\widehat{y}_5^* - \widehat{y}_5$	$\widehat{y}_6^* - \widehat{y}_6$
3	104	1.13	5	1.00	75.41	−2.92	0.02	−35.54	−16.47	63.22	−12.18
	55	6.23	1	1.00	68.65	−10.32	−43.43	−22.70	−34.57	24.47	20.24
5	48	1.40	77	1.00	17.81	0.02	13.17	−1.97	0.98	−0.87	11.75
	135	7.31	77	1.00	35.92	0.00	29.79	0.42	−8.26	−0.37	−18.28
7	29	1.18	18	1.00	17.00	−0.37	−7.26	6.60	−10.60	−8.97	0.02
	132	7.50	45	1.00	27.92	2.81	17.13	−1.83	−17.09	−7.27	11.40
9	37	1.24	52	1.00	31.05	−6.17	−11.30	−13.55	1.59	−19.13	15.69
	65	3.37	74	1.00	14.50	−3.50	0.00	0.38	−12.69	2.75	5.41
10	116	1.13	126	1.00	15.29	−1.38	2.77	−2.46	4.72	−14.00	0.00
	129	8.33	126	1.00	6.29	−2.95	2.64	1.02	0.00	−2.62	4.01
13	23	1.03	62	1.00	13.37	0.77	−1.72	−8.23	4.87	−8.94	−1.99
	21	2.06	97	1.00	11.32	7.57	−5.67	−1.40	−0.57	0.57	−6.00
14	138	9.29	110	2.00	9.58	−7.22	−2.07	−3.16	2.64	1.64	−3.95
20	120	13.26	131	1.00	40.61	−7.73	−3.26	−4.67	−0.62	39.24	−4.10
22	68	2.82	125	2.71	16.01	−0.85	−4.78	11.41	−0.72	9.81	−2.41
24	134	4.33	113	1.00	46.25	−2.70	1.95	4.07	4.30	−44.84	−9.07
25	139	1.11	76	1.00	12.68	−0.42	−3.34	6.57	3.29	5.92	7.77
	128	7.00	76	1.00	7.18	−2.82	0.51	3.66	2.17	3.03	−4.00

Source: own elaboration.

the optimal value of the selected benchmark. For the least inefficient unit #104, the number of students graduated in undergraduate programs, y_1, should be reduced by 2.92 percentage points while that of students graduated in master and PhD programs must be about the same. Interestingly, we also see that there is a trade-off in the ranking positions with respect to the efficient benchmarks. While the GNC ranking y_5 on generation of new knowledge should be increased substantially in 63.22 percentage points, the positions in the other ranking are less relevant, with the reference peer performing at a lower level in technological development and innovation, y_3, production of scientific articles, y_4, and appropriation of knowledge (IPRs), y_6. Similar analysis can be made for the most inefficient HEI #55, whose peer is HEI #1. In this case the first four outputs should be reduced to match the reference benchmark, while the last two rankings should be increased. We do not pursue any further individual benchmarking for the remaining nodes and inefficient units as they follow a similar pattern.

As for the results concerning the convexified approach CEAT, Table 6 reports the same fields as Table 5. We observe first that for the first node 3, the information is equal in both tables. This shows that, for this dataset, the convexification of the frontier does not affect the support facets of the specific region of the frontier corresponding to this node. However, results largely change in the remaining nodes, with higher inefficiency scores as reported in Section 3.3. We see that, within each node, the least and most inefficient HEI are different from Table 5; e.g. for node 5 the least inefficient firm is HEI #79, while it was HEI #48 in the EAT approach. Similarly, the most inefficient HEI is now #7, while it was HEI #135 previously. For each of these HEIs, once the equi-proportional inefficiency has been solved, it is possible to advise individual output changes that would bring their performance closer to that of their peers.

We can identify the most relevant peers by looking at their frequency when serving as benchmark for the remaining observations. Tables 7 and 8 report these frequencies for EAT and CEAT, respectively. To ease understanding, we replicate the first columns of Tables 3 and 4 above, identifying for each node the number of inefficient and efficient peers, as well as the individual HEIs that serve as benchmark. Then, in the last three columns, we report the number of times that the peer is identified as benchmark for inefficient HEIs (including itself), and the percentage that it represents within the node and the whole tree. As before, we start with the non-convex EAT frontier and the first split corresponding to node 3. We see that HEI #1, "Pontificia Universidad Javeriana", serves 16 times as reference benchmark within this node, which represents

Table 6
Actual peers for inefficient HEIs. CEAT. Selected HEIs.

Node	Ineffic. HEIs		Peer HEIs		Minimum Distance	Output slacks					
	Nº	Score	Nº	Score		$\widehat{y}_1^* - \widehat{y}_1$	$\widehat{y}_2^* - \widehat{y}_2$	$\widehat{y}_3^* - \widehat{y}_3$	$\widehat{y}_4^* - \widehat{y}_4$	$\widehat{y}_5^* - \widehat{y}_5$	$\widehat{y}_6^* - \widehat{y}_6$
3	104	1.13	5	1.00	75.41	−2.92	0.02	−35.54	−16.47	63.22	−12.18
	55	6.23	1	1.00	68.65	−10.32	−43.43	−22.70	−34.57	24.47	20.24
5	79	1.26	77	1.21	49.54	27.74	−38.63	5.12	1.22	−12.69	−1.92
	7	11.87	77	1.21	130.92	−38.21	−27.09	−66.51	−21.87	−83.31	−55.72
7	18	1.47	109	1.01	66.27	−7.05	−6.35	0.95	0.11	40.27	−51.75
	49	12.23	109	1.01	107.00	−45.98	−10.82	−47.97	−26.46	71.93	−32.28
9	52	2.50	102	1.97	90.01	20.49	−0.56	−7.51	−12.63	−44.43	−74.10
	65	10.29	102	1.97	50.62	12.66	−33.47	−0.51	−35.29	−5.31	2.80
10	124	2.51	116	2.49	51.24	−16.05	−12.95	−16.88	−8.79	−41.21	11.84
	129	18.02	116	2.49	30.01	−3.32	−0.30	7.75	−10.18	25.30	9.26
13	111	2.56	61	2.50	56.35	−9.98	−8.77	0.35	6.01	54.42	0.64
	21	7.37	61	2.50	52.46	−10.01	−16.66	24.76	−2.94	−22.83	−35.10
14	138	20.48	110	6.58	23.37	−14.98	−4.31	−6.69	11.19	11.52	−0.87
20	120	23.95	131	1.00	38.18	−14.25	−5.90	−8.58	−1.67	32.76	−8.38
22	68	16.02	125	5.34	52.42	−18.25	−28.74	21.82	−7.12	12.94	−29.91
24	113	5.97	134	4.64	65.73	3.91	12.55	27.48	22.02	−53.83	3.18
25	94	4.14	32	4.00	29.47	3.93	7.73	16.55	2.32	−3.63	−22.39
	88	16.78	32	4.00	29.13	−11.35	−3.42	6.44	2.14	25.73	0.49

Source: own elaboration.

Table 7
Peers statistics. EAT.

Node	HEIs	Inefficient HEIs	Efficient HEIs	Peer HEIs		Peer statistics		
Nº	Nº	Nº	Nº	Nº	Score	Nº	Node %	Tree %
3	25	20	5	1	1.00	16	64.00	11.11
				5	1.00	4	16.00	2.78
				89	1.00	3	12.00	2.08
				90	1.00	1	4.00	0.69
				101	1.00	1	4.00	0.69
5	18	14	4	34	1.00	2	11.11	1.39
				67	1.00	4	22.22	2.78
				77	1.00	6	33.33	4.17
				79	1.00	6	33.33	4.17
7	25	21	4	4	1.00	7	28.00	4.86
				18	1.00	6	24.00	4.17
				45	1.00	10	40.00	6.94
				109	1.00	2	8.00	1.39
9	19	14	5	52	1.00	7	36.84	4.86
				59	1.00	3	15.79	2.08
				69	1.00	3	15.79	2.08
				74	1.00	5	26.32	3.47
				102	1.00	1	5.26	0.69
10	15	13	2	124	1.00	4	26.67	2.78
				126	1.00	11	73.33	7.64
13	10	6	4	40	1.00	2	20.00	1.39
				61	1.00	1	10.00	0.69
				62	1.00	3	30.00	2.08
				97	1.00	4	40.00	2.78
14	2	2	0	110	2.00	2	100.00	1.39
16	1	1	0	141	3.29	1	100.00	0.69
19	1	0	1	98	1.00	1	100.00	0.69
20	2	1	1	131	1.00	2	100.00	1.39
22	2	2	0	125	2.71	2	100.00	1.39
24	2	1	1	113	1.00	2	100.00	1.39
25	22	21	1	76	1.00	22	100.00	17.36

Source: own elaboration.

Table 8
Peers statistics. CEAT.

Node	HEIs	Inefficient HEIs	Efficient HEIs	Peer HEIs		Peer statistics		
Nº	Nº	Nº	Nº	Nº	Score	Nº	% Node	% Tree
				1	1,00	16	64.00	11.11
				5	1.00	4	16.00	2.78
3	25	20	5	89	1.00	3	12.00	2.08
				90	1.00	1	4.00	0.69
				101	1.00	1	4.00	0.69
5	18	18	0	77	1.21	18	100.00	12.50
7	25	25	0	109	1.01	25	100.00	17.36
9	19	19	0	102	1.97	19	100.00	13.19
10	15	15	0	116	2.49	15	100.00	10.42
13	10	10	0	61	2.50	10	100.00	6.94
14	2	2	0	110	6.58	2	100.00	1.39
16	1	1	0	141	7.84	1	100.00	0.69
19	1	1	0	98	6.24	1	100.00	0.69
20	2	1	1	131	1.00	2	100.00	1.39
22	2	2	0	125	5.34	2	100.00	1.39
24	2	2	0	113	4.64	2	100.00	1.39
25	22	22	0	32	4.00	22	100.00	15.28

Source: own elaboration.

64% of the 25 observations included in it, and 11.11% of the total 144 observations in the tree. This HEI emerges as the most relevant university within the Colombian higher education system, and therefore, the ideal peer for inefficient HEIs, offering an output-mix that includes the maximum values in several dimensions. This first observation is followed by HEI #5, with less importance in terms of practical benchmarking.

Consequently, it is now possible to draw more general conclusions at the node level, highlighting the relevance of the HEI #1 regarding its inputs, outputs and its organizational settings. For example, taking the average of the output adjustments of the inefficient units that have this observation as peer we note that these HEIs have an excess in the number of undergraduate and graduate students, because y_1 and y_2 should be reduced by -3.3 and -8.65 percentage points, respectively. As for the ranking positions, y_3 thru y_6, they need to improve their standings in most of them, with average values of -0.20, 10.30, 22.14 and 17.14 percentage points. Even if some individual HEIs may need to adjust their outputs in opposite directions, these recommendations hold

for the majority of HEIs that identify the first observation as a reference. Note that the prescribed reduction in the number of students graduated in undergraduate programs - *y1*, and master and PhD programs - *y2*, to reach the levels of the benchmark peers, may be impractical or undesirable for university administrators. However, as in this node the level of the values for these output variables are among the highest for the entire higher education system, their reduction is marginal. Moreover, enrolling less students with the same faculty, implies that the ratio of students/faculty reduces, leading, in principle to a higher educational quality, which should show up in an increase in the position in the rankings as prescribed by the benchmark analysis. In the mid-term, this strategy resulting in an increase in the reputation of universities, may open the possibility to implement price-revenue management strategies, increasing the tuition fees for some student segments, and resulting in higher profitability. In principle, this reallocation strategy, freeing up resources by having less students, may prompt university administrators to encourage faculty to engage in projects and contracts with third party stakeholders (e.g., joint ventures, industry consulting, technical advice), fostered by technology transfer offices, which may ultimately result in a further net profitability. Therefore, this HEI #1 should be considered as a case study, given its relevance from a managerial and policy perspective, and worth to pay a visit. This idea is reinforced by the results obtained under the convexified CEAT frontier, which are the same for node 3.

Subsequently, looking at the highest frequencies in the remaining nodes, we observe that HEI #45 (Corporación Universidad Piloto de Colombia) in node 7 has frequency of 10, representing 40% of the observations included in this node and 6.94% of the whole sample. Nevertheless, the remaining peers in that node are not far from this value, e.g., HEI #4 and HEI #18 follow with a frequency of 7 and 6, respectively. Again, these observations are worth studying for the 21 inefficient HEI included in node 7. Similar conclusions can be drawn for other nodes, although the relevance of the benchmarks is much lower in nodes including a limited number of observations, where the only efficient peer normally serves as reference for its inefficient counterparts (i.e., thereby scoring a frequency of 100% within the node). Once again, the last node 25, represents a particular case, as all HEIs not previously classified in the preceding nodes are swept into this last part of the tree. Here, HEI #22 represents the benchmark for the remaining 21 HEIs, which are characterized by their small scale, making them peripheral within the system as a whole. In this last node it is difficult to draw general conclusions, given the diversity of managerial models included in it, even if HEI #22 is the most frequent peer in the whole tree (17.36% of observations).

Finally, Table 8 presents the analysis of benchmarks for the convexified frontier, CEAT. Besides the results corresponding to node 3, which are equal to the EAT results as explained above, we see that since most of the nodes do not harbor efficient observations, there is one single peer — the one with the least inefficiency — that serves as reference for all remaining observations. In this case, the benchmarking process simplifies because it does not require applying the least distance criterion, as is the case when there is a multiplicity of efficient observations. Then, individual adjustments can be clearly defined in the way already explained in Tables 5 and 6

5. Conclusions

The benchmarking analysis using machine learning techniques carried out in this study is expected to contribute to the literature with new evidence on how to carry out systemic evaluations of the performance of HEIs. With its application to the Colombian higher education system, it evidences how these methods have the potential to provide insights to managers and policymakers in charge of the decision making of HEIs, thereby supporting the design of strategies aimed at improving efficiency. In particular, the created tree and its nodes segment observations according to their input-output structural characteristics, to which we associate specific managerial models, and the different HEIs are grouped. In each node of the tree we identify specific benchmarks and define the output adjustments that are needed so inefficient HEIs can match the performance of their peers.

From a methodological perspective, the results achieved in the application of EAT and CEAT in the context of Colombian private HEIs reveal how standard envelopment techniques like FDH and DEA are unable to discriminate HEIs according to their efficiency. In parallel, tree methods cannot only unveil the magnitude of existing inefficiencies, but are also instrumental in providing strategic managerial and policy guidelines aimed at improving performance. Hence, the paper offers a response to the challenges associated to effective ranking and benchmarking. In this regard, machine learning methods represent a promising avenue for research, that can complement the path that has been undertaken in traditional efficiency studies, following FDH, DEA or other frontier-based analyses.

Finally, we are aware that this paper presents some limitations. EAT, as happens with DEA, represent a non-parametric technique that does not require the assumption of a specific mathematical expression for the efficient frontier. However, just like DEA, EAT is a deterministic methodology. Consequently, EAT assumes that the gap between any observation and the frontier of the technology is uniquely attributed to technical inefficiency, thus failing to capture stochastic variations affecting inputs and outputs (i.e., random noise). Also, from the boxplots associated with the empirical results, we can highlight that EAT signals some HEIs as outliers. In this regard, the treatment of this type of units with respect to the global results as well as data uncertainty could represent new lines of research to be addressed in the future. Additionally, we resorted to the radial models to measure technical efficiency. These types of measures do not prevent the existence of slacks in inputs and outputs after projecting the assessed unit onto the efficient frontier. That is why we need to apply the principle of least action when identifying the closest benchmarks. Using other efficiency measures that prevent the existence of slacks (as the Russell measures or the weighted additive models, see Ref. [61]) could be seen as a fruitful line of research as it will help to solve this drawback. Finally, another relevant dimension of analysis would be to determine the scale efficiency of HEIs by comparing their EAT/CEAT efficiency scores calculated under the assumptions of constant and variable returns to scale.

CRediT authorship contribution statement

Jose Luis Zofio: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Juan Aparicio:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Javier Barbero:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Jon Mikel Zabala-Iturriagagoitia:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Resources, Supervision, Visualization, Writing – original draft, Writing – review & editing.

Data availability

Data will be made available on request.

Acknowledgements

The authors are grateful to the editors of the special issue and to the reviewers for the comments provided in an earlier version of this paper. The authors thank the grants PID2022-138212NA-I00 and PID2022-136383NB-I00 funded by Ministerio de Ciencia e Innovación/Agencia

Estatal de Investigación/10.13039/501100011033. Juan Aparicio thanks the grant PROMETEO/2021/063 funded by the Valencian Community (Spain). Jon Mikel Zabala-Iturriagagoitia acknowledges financial support from the Basque Government Department of Education, Language Policy and Culture (IT 1429–22) and from the 2021 Membership Research Grant Scheme of the Regional Studies Association.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.seps.2024.101845.

References

[1] Sánchez-Barrioluengo M. Articulating the "three-missions" in Spanish universities. Res Pol 2014;43(10):1760–73.
[2] Benneworth P, Pinheiro R, Sánchez-Barrioluengo M. One size does not fit all! New perspectives on the university in the social knowledge economy. Sci Publ Pol 2016;43(6):731–5.
[3] Bonaccorsi A, Daraio C. The differentiation of the strategic profile of higher education institutions. New positioning indicators based on microdata. Scientometrics 2008;74(1):15–37.
[4] Sarrico CS, Rosa MJ, Teixeira PN, Cardoso MF. Assessing quality and evaluating performance in higher education: worlds apart or complementary views? Minerva 2010;48(1):35–54.
[5] Daraio C, Bonaccorsi A, Simar L. Efficiency and economies of scale and specialization in European universities: a directional distance approach. Journal of Informetrics 2015;9:430–48.
[6] Aparicio J, Rodríguez DY, Zabala-Iturriagagoitia JM. The systemic approach as an instrument to evaluate higher education systems: opportunities and challenges. Res Eval 2021;30(3):336–48.
[7] de La Torre E, Casani F, Sagarra M. Defining typologies of universities through a DEA-MDS analysis: an institutional characterization for formative evaluation purposes. Res Eval 2018;27(4):388–403.
[8] Navas LP, Montes F, Abolghasem S, Salas RJ, Toloo M, Zarama R. Colombian higher education institutions evaluation. Soc Econ Plann Sci 2020;71:100801.
[9] Warning S. Performance differences in German higher education: empirical analysis of strategic groups. Rev Ind Organ 2004;24(4):393–408.
[10] Nazarko J, Šaparauskas J. Application of DEA method in efficiency evaluation of public higher education institutions. Technol Econ Dev Econ 2014;20(1):25–44.
[11] Park J, Lim S, Bae H. An optimization approach to the construction of a sequence of benchmark targets in DEA- based benchmarking. Journal of the Korean Institute of Industrial Engineers 2015;40(6):628–41.
[12] Ruiz JL, Segura JV, Sirvent I. Benchmarking and target setting with expert preferences: an application to the evaluation of educational performance of Spanish universities. Eur J Oper Res 2015;242:594–605.
[13] Carrington R, Coelli T, Rao DSP. The performance of Australian universities: conceptual issues and preliminary results. Econ Pap: A Journal of Applied Economics and Policy 2005;24(2):145–63.
[14] Glass JC, McCallion G, McKillop DG, Rasaratnam S, Stringer KS. Best-practice benchmarking in UK higher education: new nonparametric approaches using financial ratios and profit efficiency methodologies. Appl Econ 2009;41(2):249–67.
[15] De Witte K, López-Torres L. Efficiency in education: a review of literature and a way forward. J Oper Res Soc 2017;68(4):339–63.
[16] Sarrico C. On performance in higher education: towards performance governance. Tert Educ Manag 2010;16(2):145–58.
[17] Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. Eur J Oper Res 1978;2:429–44.
[18] Torgersen AM, Førsund FR, Kittelsen SAC. Slack-adjusted efficiency measures and ranking of efficient units. J Prod Anal 1996;7:379–98.
[19] Andersen P, Petersen NC. A procedure for ranking efficient units in data envelopment analysis. Manag Sci 1993;39:1261–4.
[20] Aparicio J, Zofío JL. Economic cross-efficiency. Omega 2020;100:102374.
[21] Aparicio J, Zofío JL. New definitions of economic cross-efficiency. In: Aparicio J, Lovell CAK, Pastor JT, Zhu J, editors. *Advances in Efficiency and productivity II*, international series in operations research & management science 281. New York: Springer Science+Business Media; 2021. p. 11–32. 2020.
[22] Yamada Y, Matsui T, Sugiyama M. An inefficiency measurement method for management systems. The Operations Research Society of Japan 1994;37:158–68. 1994.
[23] Chen JX. A comment on DEA efficiency assessment using ideal and anti-ideal decision making units. Appl Math Comput 2012;219:583–91.
[24] Barbero J, Zabala-Iturriagagoitia JM, Zofío JL. Benchmarking innovation systems with DEA-TOPSIS: on the relevance of decreasing returns on waning performance. Technovation 2021;107:102314.
[25] Aldamak A, Zolfaghari S. Review of efficiency ranking methods in data envelopment analysis. Measurement 2017;106:161–72.
[26] Balk BM, de Koster R, Kaps C, Zofío JL. An evaluation of cross-efficiency methods: with an application to warehouse performance. Appl Math Comput 2021;406:126261.
[27] Esteve M, Aparicio J, Rabasa A, Rodriguez-Sala JJ. Efficiency analysis trees: a new methodology for estimating production frontiers through decision trees. Expert Syst Appl 2020;162:113783.
[28] Aparicio J, Esteve M, Rodriguez-Sala JJ, Zofio JL. The estimation of productive efficiency through machine learning techniques: efficiency analysis trees. In: Zhu J, Charles V, editors. Data-enabled analytics: DEA for big data. Cham: Springer International Publishing; 2021. p. 51–92.
[29] Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. Taylor & Francis; 1984.
[30] Teixeira P, Amaral A. Private higher education and diversity: an exploratory survey. High Educ Q 2001;55(4):359–95.
[31] Carpentier V. Expansion and differentiation in higher education: the historical trajectories of the UK, the USA and France. In: Centre for global higher education working paper series No. 33; 2018. April 2018.
[32] Kwiek M. De-privatization in higher education: a conceptual approach. High Educ 2017;74(2):259–81.
[33] Rodríguez Castro DY, Zabala-Iturriagagoitia JM, Aparicio J. Strategic groups in private higher education. Educ XX1 2021;24(1):163–87.
[34] Pineda P, Celis J. ¿Hacia la Universidad Corporativa? Reformas Basadas en el Mercado e Isomorfismo Institucional en Colombia. Archivos Analíticos de Políticas Educativas 2017;25(71):1–30.
[35] Salmi J. The tertiary education imperative knowledge, skills and values for development. In: Global perspectives on higher education. Rotterdam: Sense Publishers; 2017.
[36] Kwiek M. Private higher education in developed countries. In: Shin JC, Teixeira P, editors. The encyclopedia of international higher education systems and institutions. Dordrecht: Springer; 2018.
[37] Banker RD, Charnes A, Cooper WW. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Manag Sci 1984;30(9):1078–92.
[38] Agasisti T. Performances and spending efficiency in higher education: a European comparison through non-parametric approaches. Educ Econ 2011;19(2):199–224.
[39] Agasisti T, Bonomi F. Benchmarking universities' efficiency indicators in the presence of internal heterogeneity. Stud High Educ 2014;39(7):1237–55.
[40] Agasisti T, Johnes G. Efficiency, costs, rankings and heterogeneity: the case of US higher education. Stud High Educ 2015;40(1):60–82.
[41] Hong L, Chen X-F, Wang Y, Li Q. An exploration research of the independent college operation efficiency indicators based on DEA. Grey Syst Theor Appl 2014;4(2):362–9.
[42] Johnes G, Ruggiero J. Revenue efficiency in higher education institutions under imperfect competition. Publ Pol Adm 2017;32(4):282–95.
[43] Johnes G, Tone K. The efficiency of higher education institutions in England revisited: comparing alternative measures. Tert Educ Manag 2017;23(3):191–205.
[44] Cesaroni G, Kerstens K, Van de Woestyne I. Global and local scale characteristics in convex and nonconvex nonparametric technologies: a first empirical exploration. Eur J Oper Res 2017;259(2):5076–586.
[45] Kerstens K, Sadeghi J, Van de Woestyne I. Convex and nonconvex input-oriented technical and economic capacity measures: an empirical comparison. Eur J Oper Res 2019;276(2):699–709.
[46] Deprins D, Simar L, Tulkens H. Measuring labor inefficiency in post offices. Amsterdam: north-holland. In: Marchand M, Pestieau P, Tulkens H, editors. The performance of public enterprises: concepts and measurements; 1984. p. 243–67.
[47] Hastie T, Tibshirani R, Friedman JH, Friedman JH. In: The elements of statistical learning: data mining, inference, and prediction, vol. 2. New York: springer; 2009. p. 1–758.
[48] Vapnik V. The nature of statistical learning theory. In: Springer science & business media; 1999.
[49] Charles V, Aparicio J, Zhu J. The curse of dimensionality of decision-making units: a simple approach to increase the discriminatory power of data envelopment analysis. Eur J Oper Res 2019;279(3):929–40.
[50] Trow M. Problems in the transition from elite to mass higher education. No. 73. 1973. Berkeley, California.
[51] Melo-Becerra LA, Ramos-Forero JE, Hernández-Santamaría PO. Higher education in Colombia: current situation and efficiency analysis. Desarrollo y Sociedad 2017;78:59–111.
[52] Levy D. The decline of private higher education. High Educ Pol 2013;26(1):25–42.
[53] OECD. Private spending on education indicator. Paris: OECD; 2019.
[54] Uribe L. The decline of Colombian private higher education. International Higher Education 2010;61:12–4.
[55] Rossi F. Increased competition and diversity in higher education: an empirical analysis of the Italian University system. High Educ Pol 2009;22(4):389–413.
[56] Huisman J, Lepori B, Seeber M, Frølich N, Scordato L. Measuring institutional diversity across higher education systems. Res Eval 2015;24(4):369–79.
[57] Levy D. Growth and typology. In: Bjarnason S, Cheng K, Fielden J, Lemaitre M, Levy D, Varghese N, editors. A new dynamic: private higher education. UNESCO; 2009. p. 7–27.
[58] Ministerio de Educación Nacional de Colombia. Anuario estadístico educación superior colombiana. Colombia: Bogotá; 2021.
[59] Esteve M, España V, Aparicio J, Barber X. Eat: an R package for fitting efficiency analysis trees. RELC J 2022;14(3):248–80.
[60] OECD, & The World Bank. Reviews of national policies for education: tertiary education in Colombia. Paris: OECD; 2012.
[61] Pastor JT, Lovell CK, Aparicio J. Families of linear efficiency programs based on Debreu's loss function. J Prod Anal 2012;38:109–20.
[62] Thursby JG, Kemp S. Growth and productive efficiency of university intellectual property licensing. Res Pol 2002;31(1):109–24.

[63] Castano MCN, Cabanda EC. Performance evaluation of the efficiency of philippine private higher educational institutions: application of frontier approaches. Int Trans Oper Res 2007;14:431–44.
[64] Gwendolyn M, Cabanda E. Selected private higher educational institutions in metro manila: a DEA efficiency measurement. J Bus Educ 2009;2(6):97–108.
[65] Said A. Assessing the efficiency of for profit colleges and universities during the period of 2005-2009. International Research Journal of Finance and Economics 2011;70:90–128.
[66] Sav GT. For-profit college entry and cost efficiency: stochastic frontier estimates vs two-year public and non-profit colleges. Int Bus Res 2012;5(3):26–32.

José Luis Zofío is Professor of Economics at the Universidad Autónoma de Madrid (Spain) and Visiting Professor at the Erasmus University Rotterdam (the Netherlands). His research interests are related to measurement theory, and in particular the use of index numbers for efficiency and productivity analysis. He is engaged in multidisciplinary research, related to environmental economics, transportation, innovation, and regional science. His most recent book is "Benchmarking Economic Efficiency", published by Springer in 2022. He has made several contributions in the field of computational economics, with multiple packages for MATLAB and Julia that focus on Data Envelopment Analysis and panel data methods.

Juan Aparicio is Professor at the Department of Statistics, Mathematics and Information Technology at the Miguel Hernandez University (Spain), and head of the Center of Operations Research. His research interests include efficiency and productivity analysis combined with machine learning and data science. He has published and co-edited several books on performance evaluation and benchmarking using Data Envelopment Analysis, as well as more than 100 scientific papers in the most relevant international journals in the field. He is associate editor of Omega, the International Journal of Management Science; and Mathematics.

Javier Barbero is Assistant Professor at the Universidad Autónoma de Madrid (Spain). He has served as an economic analyst at the European Commission's Joint Research Center (JRC) in Seville and is a member of the Oviedo Efficiency Group of the University of Oviedo (Spain). His main research interests are regional economics, territorial development, and spatial economics. He has made multiple contributions to the open-source scientific software community in the previous areas.

Jon Mikel Zabala-Iturriagagoitia is Associate Professor at the University of Deusto in San Sebastian (Spain). His research skills and interests include innovation policy, innovation management, and the use of indicators to inform policy decisions and evaluations related to innovation. His research has received numerous recognitions and awards, having had a strong policy impact. Due to the impact of his research, he has been appointed as a member of the Basque Academy of Sciences, Arts and Letters (Jakiunde), the Spanish foundation for innovation (COTEC), the Science, Technology and Innovation Advisory Council (CACTI) of the Ministry of Science and Innovation of the Spanish Government, and the United Nations Economic Commission for Europe Task Force on Innovation Policy Principles.