ASSESSING THE QUALITY OF BRAZILIAN UNDERGRADUATE COURSES: A SLACK-BASED "BENEFIT OF THE DOUBT" COMPOSITE INDICATOR

Lívia Mariana Lopes de Souza Torres
Graduate Program in Management Engineering - PPGEP
Federal University of Pernambuco
Architecture Avenue, s/n, University City, Recife, Pernambuco, Brazil.
livia.mariana@ufpe.br

Felipe Alves Mendes da Silva
Graduate Program in Management Engineering - PPGEP
Federal University of Pernambuco
Architecture Avenue, s/n, University City, Recife, Pernambuco, Brazil.
felipe.alvess@ufpe.br

Francisco de Sousa Ramos
Graduate Program in Management Engineering - PPGEP
Graduate Program in Economics- PIMES/ LabGRC
Federal University of Pernambuco
Architecture Avenue, s/n, University City, Recife, Pernambuco, Brazil.
Professor Moraes Rego, 1235, University City, Recife, Pernambuco, Brazil.
francisco.ramos@ufpe.br

ABSTRACT

The present work aims to evaluate the quality of undergraduate courses in Brazilian higher education institutions. The current methodology used by the Ministry of Education, Preliminary Course Concept (CPC), considers eight indicators and uses pre-established weights to obtain a continuous CPC value. The arbitrary definition of these weights is a significant problem in the current method. Thus, to mitigate this limitation and provide a fairer assessment for the courses, the present study proposes the construction of a composite indicator using a slack-based "benefit of the doubt" model. The construction of the approach is divided in two steps, first all courses are ranked and in a second moment, a tier procedure is used to group the courses. The management engineering courses evaluated in 2019 were used as an object of study to illustrate the procedure. The results show the applicability and relevance of the proposed method.

KEYWORDS. Data Envelopment Analysis (DEA). Benefit of the doubt (BoD). Composite Indicator. Management Engineering.

Paper topics (indicate in order of PRIORITY the paper topic(s)) DEA - Data Envelopment Analysis; AdP&ED – OR in Public Administration and Education.

1. Introduction

The quality of education systems is crucial for any country, making assessment essential to assist in developing education policies. However, this evaluation involves a complex process due to the different indicators used to obtain an overview of the performance of an educational system [Stumbriene et al. 2019]. In this sense, the composite indicator (CI) emerges as one of the most popular techniques to incorporate several indicators, with vast applications in the literature in areas such as health, environment, economics, and technology development [Albo et al. 2017]. CI is a mathematical aggregation of a set of individual indicators that measure multidimensional concepts and generally do not have a similar measurement unit [Shi and Land 2020].

Traditional methods for building a CI consist of assigning weights to each indicator, and these assignments are directly related to the quality and reliability of the calculated index [Stumbriene et al. 2019; Babaee et al. 2021]. Thus, Data Envelopment Analysis (DEA) has been a promising method for assigning the best set of weights to indicators and later aggregating them to maximize the composite index score without a priori knowledge of such indicators' weights [Babaee et al. 2021].

In this context, Benefit-of-the-Doubt (BoD) consists of an approach based on DEA for CI construction, resulting in endogenous weights to aggregate indicators that vary both between such indices and between decision-making units (DMUs) evaluated. The BoD model is rooted in DEA, and it is formally tantamount to the original input-oriented CCR DEA model, with all sub-indicators considered as outputs and a 'dummy input' equal to one for all the DMUs [Cherchye et al. 2007].

In this sense, the evaluated DMUs can choose a set of weights that maximizes their performance in terms of the resulting value of the CI under the constraint that if any other evaluated DMU uses the same set of weights, it will not result in an indicator value compound greater than one [Karagiannis e Karagiannis 2018].

The Preliminary Course Concept (CPC) is a quality that aims to evaluate Brazilian undergraduate courses. It is an important parameter for Brazilian universities to obtain an overview of what is correct and what needs to be improved [Ikuta and Barreyro 2020]. In addition, CPC provides information for the elaboration of educational policies by the government, serves as a subsidy for the supervision and regulation of higher education, and is used by the federal government as a criterion in the construction of the budget distribution matrix [INEP 2020].

However, despite the importance of the indicator and being in progress for years, its formulation is still the focus of debate and criticism by the academic community, due to the lack of consensus that resides mainly in the attribution of weights in the different components that make up the CPC [Zanella and Oliveira 2021]. Therefore, these discussions are fueled by the fact that small changes in component weights can have a significant impact on the performance measurement of a course [Ikuta and Barreyro 2020].

The present study aims to evaluate the quality of undergraduate courses in Brazilian higher education institutions through the construction of a composite indicator using a slack based "benefit of the doubt" model. The proposed methodology can be applied for all the Brazilian undergraduate courses. To illustrate the proposed procedure, a sample of public management engineering undergraduate courses was selected. The used data corresponds to 2019, since it corresponds to the last evaluation of these courses.

The current text is divided into sections for a better understanding. The first briefly introduces the context of evaluating the quality of Brazilian undergraduate courses and justifies the development of the work. The second details aspects of the current model used by the Ministry of Education. The third presents conceptual aspects of Data Envelopment Analysis and its models for creating composite indicators. The fourth elucidates the proposed method. The fifth discusses the results, while the last one focuses on the final considerations.

2. Quality Assessment of Undergraduate Courses in Brazil

In Brazil, the Anísio Teixeira National Institute of Educational Studies and Research (INEP) evaluates undergraduate courses. INEP consists of a federal autarchy and represents a national reference in three main areas of activity: a) educational assessments and exams; b) statistical research and educational indicators; and c) knowledge management and educational studies. It is also vital to highlight that INEP is attached to the Ministry of Education.

Among its responsibilities concerning Higher Education, the National High School Exam (Enem) and the assessment of undergraduates, undergraduate courses, and universities must be emphasized. Enem corresponds to the most relevant evaluation test to enter Brazilian universities. Its grade can be used in public and private institutions and abroad since universities in Portugal are using the grade as an admission criterion.

INEP provides annually a set of quality indicators regarding higher education. The one discussed in the current study is the Preliminary Course Concept (CPC). CPC is a quality indicator that combines, in a single measure, different aspects related to undergraduate courses. It consists of eight components, grouped into four dimensions that are intended to assess the quality of undergraduate courses [INEP, 2020]. Table 1 details the dimensions and the indicators used in the CPC calculation.

Table 1: Descriptive statistics of the indicators

Dimension	Indicators	Weight
Student's performance	Grade of senior students in Enade (CE)	0,2
Added value by the education provided by the undergraduate course	Difference Indicator Score between Observed and Expected Performance (IDD)	0,35
Faculty	Masters Proportion Score (M)	0,075
	Doctor Proportion Score (D)	0,15
	Work regime Proportion Score (WRP)	0,075
Student perception of the undergraduate course conditions	Grade regarding the didactic-pedagogical organization (DPO)	0,075
	Grade regarding infrastructure and Physical Facilities (NF)	0,05
	Grade regarding Opportunity to Expand Vocational Training (OEVT)	0,025

All the indicators in the student perception of the undergraduate course conditions are obtained in a questionnaire that all undergraduates must fill out before the Enade exam. The value of each indicator consists of an average of the values returned by the students. The indicators regarding the faculty represent the proportion of the faculty that possesses a master's and a Ph.D. title. In contrast, the work regime proportion indicates the percentage of undergraduate course professors with partial or full-time work.

Two indicators represent the first two dimensions and reflect how well students performed. CE indicated the grade obtained in Enade and this test evaluates the main topics a student should know when nearly concluding their undergraduate course. On the other hand, the IDD measures the value added by the undergraduate course to the development of graduating students, considering their performance in the Enade [INEP, 2020]. In Table 1, it is also possible to verify the weights used by INEP to obtain the CPC value. It is essential to mention that in the technical report Nº 58/2020/CGCQES/DAES, in which the calculation of the CPC is detailed, there is no justification or explanation for the selection of such values.

CPC consists of a weighted sum of eight indicators; therefore, it is a composite indicator. Due to its relevance to the evaluation of undergraduate courses and the universities in which the course is located, this study proposes an alternative proposition to calculate the CPC. A method based on DEA "benefit of the doubt" is the suggested mathematical formulation to obtain a new CPC, because it does not require pre-defined weights, and only uses the indicators values to perform the evaluation. The following section presents the mathematical formulation for the DEA BoD model in detail.

3. Data Envelopment Analysis (DEA) for building composite indicators

Data Envelopment Analysis (DEA) is one of the most effective techniques to measure the efficiency of a set of decision-making units (DMUs) with multiple inputs and outputs [Kao 2016]. Developed by Charnes, Cooper, and Rhodes [1978], this technique has become one of the main non-parametric techniques, with models consolidated over more than forty years. In this sense, there are specific input-oriented cases, where only the outputs are evaluated, with an input dummy equal to one for all DMUs, such an approach is labeled as Benefit-of-the-Doubt (BoD), being fundamental in the context of CI construction, with all sub-indicators considered as outputs [Van Puyenbroeck 2017].

Cherchye et al. [2006] was the pioneer to propose a benefit of the doubt methodology to build CIs. The authors started from the assumption of the vulnerable credibility of CIs, affected by possible divergences between specialists or stakeholders on the weighting scheme used to aggregate sub-indicators. The methodology is proposed to enable a flexible weighting scheme to compare DMUs. This seminal study was expanded in several directions, weight-constrained models [Van Puyenbroeck and Rogge 2017], robust models [Dardha and Rogge 2020], directional models [Färe et al. 2019], conditional models [Fusco et al. 2018] and slack based models (SBM) [Mariano et al. 2021].

Given the relevance involved in the construction of CIs, the Organization for Economic Co-operation and Development (OECD) has developed a manual for the construction of CIs with applications in diverse areas, such as politics, communication, economics, education, and the environment [OECD 2008]. As such, there are more than 150 composite indicators, including the Human Development Index (HDI), Quality of Life, Competitiveness, Technological Achievement, Health System Performance, Environmental Performance, and various subjective well-being indexes [Karagiannis and Karagiannis 2020].

Through the various areas of applicability of CIs, education has received attention in this context. [Rogge 2011] proposed the application of a benefit of the doubt approach in teacher's performance evaluation. [Maričić et al. 2016] developed a I-distance contrained BoD model to discuss the weight scheme impact on the Quacquarelli Symonds (QS) university index. [De Witte and Schiltz 2018] applied a robust and conditional BoD approach to investigate school districts effectiveness. [Szuwarzyński 2019] addressed research performance of Australian universities with the aid of a constrained BoD model.

Aiming to contribute to the literature on educational assessments with BoD models, the present study is focused on quality of higher education, namely the evaluation of undergraduate courses at public Brazilian universities. Brazil is an intercontinental country with socioeconomic problems, such as social inequality, lack of qualified human capital, and increasing poverty. Thus, overcoming such problems involves the consolidation of a qualified higher education system committed to the interests of society [Bertolin 2011]. Therefore, the performance evaluation and monitoring of the national higher education system is an essential public service analysis to ensure greater parity between the regions of the country.

4. Research Method

This section details the research method employed in the paper. First, there is a brief discussion of descriptive statistics of the eight indicators employed. Second, the SBM BoD model is used to aggregate the indicators into a composite one. And at last, the procedure used to obtain tier to identify similarities among the DMUs is described.

4.1 Data

The current study aims to compare the results of the current model used by the Ministry of Education to evaluate Brazilian undergraduate courses with an SBM DEA BoD Model. Thus, selecting the current eight indicators used in the evaluation process is necessary to ensure comparability of results. The source of collected data used in the study is the website of INEP. The data corresponds to 2019, the last year that the Ministry of Education evaluated Management Engineering courses. It is also relevant to mention that all indicators are normalized in a range of 0 to 5. This fact helps the development of this study since the use of normalized data aids in avoiding outliers and in finding a more homogenous relative contribution between the variables [Mariano et al. 2021].

In 2019, 654 courses of Management Engineering were evaluated. One hundred thirteen of these courses correspond to public institutions. Since, in Brazil, public HEIs are funded with public resources and students do not pay any tuition fees, it is vital to analyze the quality of the education provided by these institutions. Six courses were removed due to the lack of information. Our sample comprises 107 DMUs. 79 are Federal institutions, 27 are State institutions, and 7 are Municipal institutions. Table 2 shows the descriptive statistics of the indicators. As previously mentioned, all indicators are ranged between 0 to 5 and all indicators selected are the same used by the government in the CPC analysis to obtain a directly comparison.

It is also relevant to verify the distribution across the macro-regions of the country. The Brazilian Midwest (CO) contemplates eight undergraduate courses, and the North (N) provides six. The Southeast (SE) has the higher concentration, 48, and the northeast (NE) and south (SUL) have 22 and 23, respectively.

Table 2: Descriptive statistics of the indicators

Indicator	Mean	SD	Median	Max	Min
Grade of senior students in Enade (CE)	3,289	0,979	3,26	5	0,755
Difference Indicator Score between Observed and Expected Performance (IDD)	2,566	0,614	2,485	5	0,776
Doctors (D)	3,371	1,152	3,548	5	0,394
Masters (M)	4,422	0,705	4,642	5	1,818
Infrastructure and Physical Installations (IPI)	2,3	0,861	2,322	4,782	0,249
Opportunity to Expand Vocational Training (OEVT)	2,901	0,66	2,973	4,319	0,877
Didactic-pedagogical organization (DPO)	2,203	0,761	2,209	4,518	0,596
Work regime Proportion (WRP)	4,871	0,569	5	5	1,225

4.2 BoD SBM Model

As previously mentioned, the BoD models provide a way to aggregate multiple indicators without the pre-definition of weights. In this study, the DEA BoD model is based on the SBM hypothesis. In this study, the model used is detailed in (1). Let's consider *n* DMUs under evaluation and its respective indicators, in our case, eight are chosen (*CE, IDD, D, M, IPI, OEVT, DPO, WRP*). The linear programming presented in (1) corresponds to the envelopment form of an SBM BoD DEA model. CE_O, IDD_O, D_O, M_O, IPI_O, $OEVT_O$, DPO_O and WRP_O are the indicators of the DMU under observation, while s_{CE}, s_{IDD}, s_D, s_M, s_{IPI}, s_{OEVT}, s_{DPO} and s_{WRP} are the corresponding indicators' slacks, while λ_k corresponds to the importance level of benchmark k for the target of the course under analysis.

$$max\ \frac{1}{CPC_{BOD}} = 1 + \frac{1}{8}\left(\frac{s_{CE}}{CE_O} + \frac{s_{IDD}}{IDD_O} + \frac{s_D}{D_O} + \frac{s_M}{M_O} + \frac{s_{IPI}}{IPI_O} + \frac{s_{OEVT}}{OEVT_O} + \frac{s_{DPO}}{DPO_O} + \frac{s_{WRP}}{WRP_O}\right)$$

$Subject\ to$

$$\sum_{k=1}^{n} CE_k * \lambda_k - s_{CE} = CE_O$$

$$\sum_{k=1}^{n} IDD_k * \lambda_k - s_{IDD} = IDD_O$$

$$\sum_{k=1}^{n} D_k * \lambda_k - s_D = D_O$$

$$\sum_{k=1}^{n} OEVT_k * \lambda_k - s_{OEVT} = OEVT_O$$

$$\sum_{k=1}^{n} DPO_k * \lambda_k - s_{DPO} = DPO_O$$

$$\sum_{k=1}^{n} WRP_k * \lambda_k - s_{WRP} = WRP_O$$

$$\sum_{k=1}^{n} M_k * \lambda_k - s_M = M_O \qquad \sum_{k=1}^{n} \lambda_k = 1$$

$$\sum_{k=1}^{n} IPI_k * \lambda_k - s_{IPI} = IPI_O \qquad \lambda_k, s_{WRP}, s_{DPO}, s_{OEVT}, s_{IPI}, s_M, s_D, s_{IDD}, s_{CE} \geq 0 \tag{1}$$

When model (1) is solved, the inverse of the objective function corresponds to the composite indicator, and this value will allow ranking the DMU's. This model must be run *n* times, and it will also be used as a base for the clustering procedure presented in the next section.

The choice of a SBM model aims to obtain a non-radial evaluation. That is, in this efficiency analysis it is not necessary to consider a proportional increase in outputs for improvements in the efficiency score. In this way, modeling allows a more realistic assessment of the situation.

4.3 Tier DEA

After defining the BoD SBM model as the chosen model to aggregate the indicators of the study, we applied the procedure developed by [Barr et al., 2000] to obtain tiers of DMUs. The choice of the Tier DEA procedure is based on the identification of similar groups, which allows the development of specific actions for each group, directly attacking the problems of each group.

The method developed by the authors is referred to as "Peeling the DEA Onion" and consists of serial steps to represent successive layers of relative efficient production surfaces. The DMUs at any giver tier are less productively efficient than those of outer tiers and more efficient than DMUs at inner tier [Barr et al., 2000].

The procedure starts with a traditional DEA. Then, a new layer of efficiency is obtained when efficient DMUs are suppressed from the data to uncover the next level of efficient DMUs. The removal of the efficient DMUs in each step reveals a series of frontiers of decreasing productivity and separates the DMUs into efficient and inefficient groups, therefore creating the tier levels [Barr et al., 2000; Bougnol and Dulá, 2006].

The creation of tiers is significant because it allows for identifying managerial practices that enhance or reduce the efficiency when analyzing top-ranked and bottom-ranked DMUs. Since the institutions evaluated in this study are funded with public resources, it is crucial to improve their performance, and public policies can be used to guide these courses to enhance the quality that they provide to their undergraduate students. Also, considering that the number of courses that the Ministry of Education evaluates is vast, the identification of similar groups allows for proposing more accurate directions to improve performance and define targets to be archived by such parties.

Let's consider *t* as a tier index, $E^*_{[t]}$ and $I^*_{[t]}$ corresponds to the sets of efficient and inefficient DMUs on tier *t*, respectively, relative to set $D^{[t]}$. The Tier DEA is described as follows.

Tier DEA Procedure

$1 - Initialize: t \leftarrow 1, D^{[1]} \leftarrow D.$
$2 - While\ D^{[t]} \neq \emptyset\ do:$
$a)\ Apply\ a\ DEA\ model\ to\ the\ DMUs\ in\ set\ D^{[t]}\ to\ identify\ E^*_{[t]}$
$b)\ I^*_{[t]} = D^{[t]} - E^*_{[t]}$
$c)\ t \leftarrow t + 1$
$d)\ D^{[t]} = I^*_{[t]}$

5. Results

This section presents three levels of results as discussed in the research method. First, the indicators are presented, and a brief discussion regarding their dispersion across the Brazilian

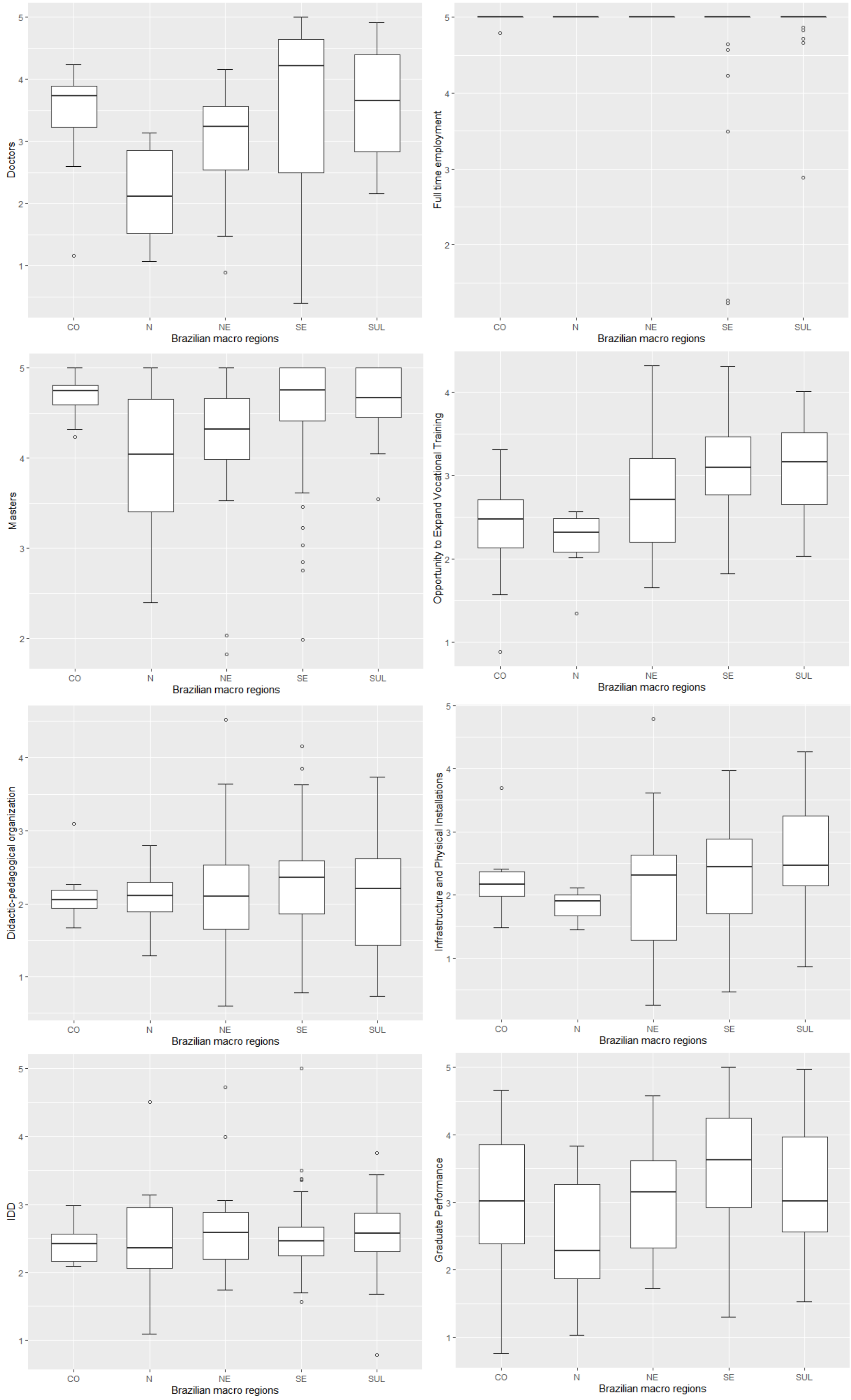

Figure 1: Dispersion of indicators in Brazilian macro-regions

macro-region is presented. After, the results obtained by applying the DEA BoD SBM model are detailed. And the last level of analysis consists of the identification of similar courses with the DEA cluster procedure. Figure 1 illustrates the boxplots created for analyzing the eight indicators. It is essential to notice that some variability was expected since the Brazilian macro-regions are very distinct.

The indicator's values confirmed this fact, but on average, all regions obtained similar values for the full-time employment indicator. This can be explained by the fact that the institutions are public, and most hired teachers must work full time for the institution, but this pattern does not apply to private institutions in Brazil. Regarding faculty qualification, the percentage of masters hired is higher than doctors, and although there is variability among the regions, the North performs worse in both cases.

Table 3- SBM BoD Ranking

Rank	University	Score	Rank	University	Score	Rank	University	Score
1	UNESP-Itapeva	1	37	UERJ-Rio de Janeiro	0,8163	73	UFPB-João Pessoa	0,6681
1	UNESP-Itapeva	1	38	UTFPR-Londrina	0,8156	74	FEMASS-Macaé	0,6663
1	UFF-Niterói	1	39	UFF-Petrópolis	0,8145	75	IFMG-Congonhas	0,6592
1	UFMG-Belo Horizonte	1	40	UFF-Volta Redonda	0,8034	76	UEM-Goioerê	0,6592
1	UFRGS-Porto Alegre	1	41	UFV-Rio Paranaíba	0,7980	77	UFERSA-Angicos	0,6464
1	UFC-Russas	1	42	UFF-Volta Redonda	0,7959	78	IFSP-São Paulo	0,6461
1	UFSC-Florianópolis	1	43	UDESC-São Bento do Sul	0,7954	79	UEM-Maringá	0,6439
1	UFSC-Florianópolis	1	44	UFPE-Caruaru	0,7937	80	UFGD-Dourados	0,6438
1	UFRJ-Rio de Janeiro	1	45	UNB-Brasília	0,7879	81	UFBA-Salvador	0,6412
1	UFRJ-Macaé	1	46	UFU-Ituiutaba	0,7835	82	UFCG-Campina Grande	0,6270
1	UTFPR-Apucarana	1	47	UFAL-Penedo	0,7817	83	UEA-Manaus	0,6241
1	CEFET/RJ-Nova Iguaçu	1	48	UFMS-Campo Grande	0,7793	84	UFAM-Itacoatiara	0,6210
1	CEFET/RJ-Itaguaí	1	49	UNESPAR-Campo Mourão	0,7779	85	FESURV-Rio Verde	0,6133
1	UENF-Campos dos Goytacazes	1	50	UFTM-Uberaba	0,7766	86	USCS-São Caetano do Sul	0,6014
1	IFCE-Quixadá	1	51	IFES-Cariacica	0,7666	87	UNIR-Cacoal	0,5868
1	UNESP-Guaratinguetá	1	52	UNIPAMPA-Bagé	0,7664	88	UEM-Maringá	0,5823
17	UFSM-Santa Maria	0,9382	53	UFS-São Cristóvão	0,7644	89	UFES-Vitória	0,5823
18	UFOP-Ouro Preto	0,9303	54	UFCA-Juazeiro do Norte	0,7626	90	UFPR-Curitiba	0,5740
19	UFABC-São Bernardo do Campo	0,9249	55	UFVJM-Teófilo Otoni	0,7573	91	UFMS-Três Lagoas	0,5722
20	UNIFEI-Itajubá	0,9079	56	FURB-Blumenau	0,7570	92	IFMG-Governador Valadares	0,5509
21	UFOP-João Monlevade	0,8974	57	UNIRIO-Rio de Janeiro	0,7549	93	UEM-Maringá	0,5150
22	UFF-Rio das Ostras	0,8881	58	UFG-Catalão	0,7479	94	UFES-São Mateus	0,5082
23	UFV-Viçosa	0,8710	59	UFPI-Teresina	0,7476	95	UFOB-Luís Eduardo Magalhães	0,5038
24	CEFET/RJ-Rio de Janeiro	0,8668	60	UFSCAR-Sorocaba	0,7312	96	UERJ-Resende	0,5011
25	UTFPR-Medianeira	0,8626	61	UFPE-Recife	0,7252	97	UFAL-Delmiro Gouveia	0,4958
26	CEFET/MG-Belo Horizonte	0,8613	62	UNIFEI-Itabira	0,7222	98	UEZO-Rio de Janeiro	0,4753
27	UNICAMP-Limeira	0,8526	63	UNIVASF-Juazeiro	0,7204	99	UFPEL-Pelotas	0,4668
28	UFG-Aparecida de Goiânia	0,8515	64	UEMA-São Luís	0,7162	100	UEM-Maringá	0,4629
29	UNESP-Bauru	0,8500	65	IFMG-Bambuí	0,7104	101	UNEMAT-Barra do Bugres	0,4620
30	UFSJ-São João del Rei	0,8421	66	UDESC-Joinville	0,7045	102	UNITAU-Taubaté	0,4614
31	UNICAMP-Limeira	0,8364	67	UFPB-João Pessoa	0,7037	103	UEMG-Passos	0,4540

32	UFSC-Florianópolis	0,8361	68	UFAM-Manaus	0,7007	104	UESC-Ilhéus	0,4104
33	UFPEL-Pelotas	0,8354	69	UFSCAR-São Carlos	0,6892	105	UEAP-Macapá	0,3960
34	UTFPR-Ponta Grossa	0,8348	70	UFCG-Sumé	0,6850	106	UEMG-Divinópolis	0,2925
35	UFRN-Natal	0,8231	71	UFPA-Abaetetuba	0,6811	107	URCA-Juazeiro do Norte	0,2616
36	UFPR-Jandaia do Sul	0,8192	72	UFERSA-Mossoró	0,6757			

It is also interesting that, on average, all regions performed similarly in the IDD factor, but there is a significant dispersion in students' performance in the ENADE evaluation. When observing students' perception of its courses, there is higher variability in the opportunities to expand Vocation training than in the other two factors.

Table 3 presents the ranking of courses according to the composite indicator obtained with the BoD model. The average performance of Management Engineering undergraduate courses corresponds to 0,7431. Initially, we can observe that the 16 courses got the maximum scores, and the course of URCA-Juazeiro do Norte obtained the minimum performance of 0,2616.

The best performance can explain the maximum score of the 16 efficient courses in at least two indicators. In addition, 4 of them achieved maximum performance in three dimensions. It is also vital to notice that only two efficient courses are not located in the Southeast and South regions.

On the other hand, the course with the lowest performance of the sample was from URCA located in Juazeiro do Norte-CE. The low efficiency is a consequence of the worst performance in IPI, the fourth-worst performance in DPO, and the seventh-worst performance in CE. In addition, it is essential to note that there is minimal difference between the performances of the universities of URCA-Juazeiro do Norte and UEMG-Divinópolis (next to last place in the ranking), due to the performance in similar indicators.

Since 14 of 16 efficient courses were situated in the Southeast and South regions, we investigated the results of the composite indicator aggregated by macro-region to analyze the course performance concerning their geographical position in the country. Figure 2 illustrates the results obtained with the DEA model by macro-regions.

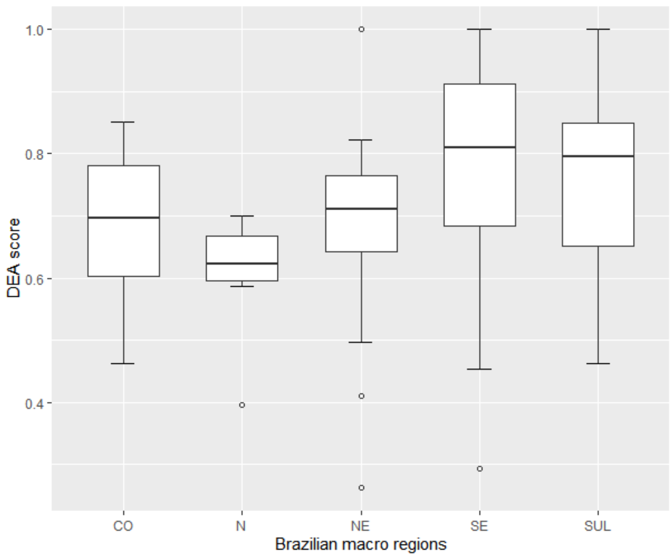

Figure 2: Composite indicators dispersion across Brazilian macro regions

It is possible to verify that, on average, the Southeast region obtained the best performance, followed by the South, the Northeast, and the Midwest, while the North obtained the worst performance. The Kruskal–Wallis non-parametric statistical test was the tool selected to verify whether the median of the Brazilian regions' efficiencies are equal. The test returned a p-value of 0,009802 and a chi-squared of 13,323 considering four degrees of freedom and a significance level of 0,05. This value indicates that at least one of the regions differs significantly from the others and can be comprehended as evidence of a significant difference in performance across the country.

After comparing the macro-regions results, a correlation analysis was performed between the ranking provided by the BoD model and the one provided by the continuous CPC value. Spearman's rho and Kendall's Tau correlation coefficient were calculated at a significance level of 0,05. The significant values of 0,8737and 0,7053 were obtained for the Spearman's and Kendall's coefficients (p < 0,0001). Both values indicate a strong correlation between the two rankings. However, the use of BoD model does not require a pre definition of weights, while the CPC values are obtained with the weights detailed in section 2.

As detailed in section 4.3, the algorithm used to obtain tiers of courses consists of identifying the efficient DMUs in each iteration of the procedure. Table 4 details the 6 tiers identified.

Table 4 – Six tier and respective courses

Tier 1	Tier 2	Tier 3	Tier 4	Tier 5	Tier 6
UNESP-Itapeva	UNB-Brasília	UFOP-João Monlevane	UFS-São Cristóvão	UFAM-Itacoatiara	UEM-Maringá
UNESP-Itapeva	UFOP-Ouro Preto	UFSCAR- São Carlos	UFAM-Manaus	UESC-Ilhéus	UFAL-Delmiro Gouveia
UFF-Niterói	UFV-Viçosa	UFV-Rio Paranaíba	UFPI-Teresina	UEM-Maringá	UNITAU-Taubaté
UFMG-Belo Horizonte	UDESC-São Bento do Sul	UDESC-Joinville	UFSCAR-São Carlos	UEM-Maringá	UEMG-Passos
UFRGS-Porto Alegre	UNICAMP-Limeira	UNICAMP-Limeira	UFU-Ituiutaba	UEM-Maringá	UEAP- Macapá
UFC-Russas	UNESP-Bauru	UFSJ-São João del Rei	UEM-Goioerê	UERJ-Resende	
UFSC-Florianópolis	FURB-Blumenau	UERJ-Rio de Janeiro	UEMA-São Luíz	UFES-São Mateus	
UFSC-Florianópolis	UFF-Rio das Ostras	UFPA-Abaetetuba	UFBA- Salvador	UFPEL-Pelotas	
UFRJ-Rio de Janeiro	UFF-Petrópolis	UFRN-Natal	UFPB-João Pessoa	UFMS- Três Lagoas	
UFRJ-Macaé	UFAL-Penedo	UFPR-Curitiba	UFPB-João Pessoa	UNEMA-Barra do Bugres	
UTFPR-Apucarana	UFSM-Santa Maria	UFPR-Jandaia do Sul	UFG-Catalão	URCA-Juazeiro do Norte	
CEFET/RJ-Nova Iguaçu	UFG-Aparecida de Goiana	UFF-Volta Redonda	UTFPR-Londrina	IFMG-Governador Valadares	
CEFET/RJ-Itaguaí	UFSC-Florianópolis	UFF-Volta Redonda	UFERSA-Mossoró	UFGD-Dourados	
UENF-Campos dos Goytacazes	UTFPR-Medianeira	UFES-Vitória	UNIFEI-Itabira		
IFCE-Quixadá	CEFET-Rio de Janeiro	UFPE-Recife	UFMS-Campo Grande		
UNESP-Guaratinguetá	UNIFEI-Itajubá	UFPE-Caruaru	UEMG-Divinópolis		
	USCS-São Caetano do Sul	UTFPR- Ponta Grossa	UFCG-Campina Grande		
	UFPEL-Pelotas	UFERSA	UFCG-Sumé		
	UNIR-Cacoal	CEFET-Minas Gerais	UFOB-Luís Eduardo Magalhães		
	IFMG-Bambuí	UFVJM-Teófilo Otoni			
	FESURV-Rio Verde	UFTM-Uberaba			

	UFABC-São Bernardo do Campo	UFRJ- Rio de Janeiro
	UNESPAR-Campo Mourão	IFES-Cariacica
		IFSP-São Paulo
		UEA-Manaus
		IFMG-Congonhas
		UNIVASF-Juazeiro
		FEMASS-Macaé
		UEZO-Rio de Janeiro
		UNIPAMPA-Bagé
		UFCA-Juazeiro do Norte

The first one corresponds to the 16 DMUs previously identified as efficient, and this value is equivalent to 14.95% of the sample. These DMUs achieved maximum performance on two or more indicators, therefore explaining their position on the first tier of courses.

The second tier comprises 23 universities (21.49% of the sample). This group has a high predominance of DMUs in the South and Southeast regions and a good performance in the master's indicator. The third tier corresponds to 31 universities (28.97% of the sample). The DMUs belonging to this group have an average performance in the indicators, with the exception of indicator OEVT.

The tier comprises 19 universities (17.76% of the sample), with a high predominance of universities in the northeast region and a moderate performance in most indicators (values ranging between 2 and 3). However, on average lower than the values verified in the third tier. The fifth one has 13 universities (12.15% of the sample), and this group has a lower performance in indicators D and M compared to the previous tiers. Finally, the sixth tier contains five universities (4.67% of the sample) with low performance in all indicators, except for indicator OEVT.

The aim of developing tier was to identify DMUs with similar performance characteristics so that it is possible to design similar public policies to improve the performance of a group of universities. This gathering allows verifying where to prioritize actions. And these actions should be different according to each group's performance. Thus, the method used allows for identifying groups of similar DMUs, and its results indicate that tier 6 requires more attention. As for tiers 4 and 5, indicators D and M should be the focus. In contrast, the others have average performance in most indicators. Therefore, the efficient courses should be used as a benchmark to identify best practices to improve them.

6. Conclusion

This paper has outlined a new procedure to assess the quality of Brazilian undergraduate courses with the aid of a slack-based "benefit of the doubt" model. The intent is to present a valid alternative to aggregate the eight indicators considered in the evaluation of the undergraduate courses without a pre definition of weights. The choice of a slack-based approach rather than other DEA models is based on the possibility of a non-radial evaluation of the indicators and since there is no previous opinions of experts, the SBM approach is able to provide a more balanced evaluation as evidenced by the results of [Mariano et al. 2021].

Besides the development of a new scheme to obtain a new CPC, we also proposed the use of a tiering method to gather similar DMUs, and this identification is useful to develop specific policies for each tier to improve their performance based on the best practices. It is essential to mention that the current study investigated public Management Engineering undergraduate courses to illustrate the applicability of the method, but the proposition is not limited to this context and can be used to evaluate all the courses.

As future directions to expand the studies, it is necessary to discussed the indicators used in the evaluation and compare the SBM results with other BoD alternatives such as directional models or conditional ones. The conditional BoD can present an interesting alternative because it allows for the considerations of aspects, such as geographical location since the statistical tests indicated discrepancies between the Brazilian macro-regions.

References

Albo, Y; Lanir, J; Rafaeli, S (2017). A Conceptual Framework for Visualizing Composite Indicators. *Social Indicators Research*, *141(1)*:1–30.

Babaee, S.; Toloo, M.; Hermans, E.; Shen, Y. (2021). A new approach for index construction: the case of the road user behavior index. Computers & Industrial Engineering, *152*:106–993.

Barr, R., Durchholz, M., & Seiford, L. (2000). *Peeling the DEA Onion: Layering and Rank-Ordering DMUs using Tiered DEA*. Dallas, TX: Southern Methodist University Technical Report.

Bertolin, J. C. G. (2011). Uma proposta de indicadores de desempenho para a educação superior brasileira. *Estudos em Avaliação Educacional, 22(50)*: 471.

Bougnol, M. L., & Dulá, J. H. (2006). Validating DEA as a ranking tool: An application of DEA to assess performance in higher education. *Annals of Operations Research*, *145*(1): 339–365.

Charnes, A.; Cooper, W. W.; Rhodes, E (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2(6)*:429–444.

Cherchye, L.; Moesen, W.; Rogge, N; Van puyenbroeck, T. (2007). An Introduction to 'Benefit of the Doubt' Composite Indicators. *Social Indicators Research, 82(1)*:111–145.

Dardha, E.; Rogge, N. (2020). How's Life in Your Region? Measuring Regional Material Living Conditions, Quality of Life and Subjective Well-Being in OECD Countries Using a Robust, Conditional Benefit-of-the-Doubt Model. *Social Indicators Research, 151(3)*: 1015–1073.

De Witte, K.; Schiltz,F. (2018). Measuring and explaining organizational effectiveness of school districts: Evidence from a robust and conditional Benefit-of-the-Doubt approach. *European Journal of Operational Research,* 267: 1172-1181.

Färe, R.; Karagiannis, G.; Hasannasab, M.; Margaritis, D. (2019). A benefit-of-the-doubt model with reverse indicators. *European Journal Of Operational Research, 278(2)*: 394–400.

Fusco, E.; Vidoli, F.; Sahoo, B. K. (2018). Spatial heterogeneity in composite indicator: a methodological proposal. *Omega, 77:*1–14.

Ikuta, C. Y. S.; Barreyro, G. B. (2020). análise da qualidade dos cursos do programa universidade para todos (prouni). *Revista da Faeeba - Educação e Contemporaneidade*, *30(61)*: 344–363.

INEP. (2020). CPC 2019 de 91,6% dos cursos está entre as faixas 3 e 5. //https://www.gov.br/inep/pt-br/assuntos/noticias/indicadores-de-qualidade-da-educacao-superior/cpc-2019-de-91-6-dos-cursos-esta-entre-as-faixas-3-a-5 . Acessado: 2022-05-06.

INEP. (2020). Nota Técnica nº 58/2020/CGCQES/DAES. // https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/indicadores-de-qualidade-da-educacao-superior/outros-documentos. Acessado: 200-05-01.

Kao, C. (2016). Efficiency decomposition and aggregation in network data envelopment analysis. *European Journal of Operational Research*, *255(3)*:778–786.

Karagiannis, R.; Karagiannis, G. (2020). Constructing composite indicators with Shannon entropy: the case of human development index. *Socio-Economic Planning Sciences, 70*: 100–701.

Karagiannis, R; Karagiannis, G (2018). Intra- and inter-group composite indicators using the BoD model. *Socio-Economic Planning Sciences, 61*:44-51.

Mariano, E. B., Ferraz, D., & de Oliveira Gobbo, S. C. (2021). The Human Development Index with Multiple Data Envelopment Analysis Approaches: A Comparative Evaluation Using Social Network Analysis. *Social Indicators Research*, *157*(2): 443–500.

Maričić, M., Bulajić, M., Radojičić, Z., Jeremić, V. (2016). Multivariate approach to imposing additional constraints on the Benefit-of-the-Doubt model: The case of QS World University Rankings by Subject. *Croatian Review of Economic, Business and Social Statistics*, *2(1):*1-14.

OCDE. (2008). Centro Comum de Investigação-Comissão Europeia. Manual de construção de indicadores compostos: Metodologia e guia do usuário. Paris: Editora OCDE.

Rogge, N. (2011). Granting teachers the "benefit of the doubt" in performance. *International Journal of Educational Management*, *25(6)*:590-614.

Shi, C.; Land, K. C. (2020). The Data Envelopment Analysis and Equal Weights/Minimax Methods of Composite Social Indicator Construction: a methodological study of data sensitivity and robustness. *Applied Research In Quality Of Life, 16(4):*1689–1716.

Stumbriene, D; Camanho, A. S.; Jakaitiene, A (2019). The performance of education systems in the light of Europe 2020 strategy. *Annals Of Operations Research*, *288(2)*: 577–608.

Van puyenbroeck, T. (2017). On the Output Orientation of the Benefit-of-the-Doubt-Model. *Social Indicators Researc*h, *139(2):* 415–431.

Szuwarzyński, A. (2019). Benefit of the doubt approach to assessing the research performance of Australian universities,*73(2)*:235-250.

Zanella, A; Oliveira, R. M. S (2021). Avaliação de desempenho na educação superior: uma abordagem utilizando a análise envoltória de dados. *Ciência e Natura*, 43: 81.