




Article

# How Efficiently Do Elite US Universities Produce Highly Cited Papers?

Klaus Wohlrabe <sup>1,\*</sup>, Félix de Moya Anegón <sup>2</sup> and Lutz Bornmann <sup>3</sup><sup>1</sup> Ifo Institute, Poschingerstr. 5, 81679 Munich, Germany<sup>2</sup> CSIC, Institute of Public Goods and Policies (IPP) Consejo Superior de Investigaciones Científicas C/Albasanz, 26-28 28037 Madrid, Spain; felix.moya@scimago.es<sup>3</sup> Division for Science and Innovation Studies, Administrative Headquarters of the Max Planck Society Hofgartenstr. 8, 80539 Munich, Germany; lutz.bornmann@gv.mpg.de

\* Correspondence: wohlrabe@ifo.de; Tel.: +49-89-9224-1229

Received: 2 October 2018; Accepted: 7 January 2019; Published: 10 January 2019



**Abstract:** While output and impact assessments were initially at the forefront of institutional research evaluations, efficiency measurements have become popular in recent years. Research efficiency is measured by indicators that relate research output to input. The additional consideration of research input in research evaluation is obvious, since the output depends on the input. The present study is based on a comprehensive dataset with input and output data for 50 US universities. As input, we used research expenses, and as output the number of highly-cited papers. We employed Data Efficiency Analysis (DEA), Free Disposal Hull (FDH) and two more robust models: the order- $m$  and order- $\alpha$  approaches. The results of the DEA and FDH analysis show that Harvard University and Boston College can be called especially efficient compared to the other universities. While the strength of Harvard University lies in its high output of highly-cited papers, the strength of Boston College is its small input. In the order- $\alpha$  and order- $m$  frameworks, Harvard University remains efficient, but Boston College becomes super-efficient. We produced university rankings based on adjusted efficiency scores (subsequent to regression analyses), in which single covariates (e.g., the disciplinary profile) are held constant.

**Keywords:** university; efficiency analysis; partial frontier analysis; regression analysis; normalized citation impact

## 1. Introduction

The science system has been characterised by the transition from academic science to post-academic science for several years. “Bureaucratization” is the term used to describe most of the processes connected with post-academic science: “The transition from academic to post-academic science is signaled by the appearance of words such as management, contract, regulation, accountability, training, employment, etc. which previously had no place in scientific life. This vocabulary did not originate inside science, but was imported from the more ‘modern’ culture which emerged over several centuries in Western societies—a culture characterized by Weber as essentially ‘bureaucratic’” ([1], p. 82). As an important part of universities’ commitments to accountability (against the government), research evaluation has assumed a steadily growing importance in the science system. While academic science (since its beginnings) has been characterised by the use of the peer-review system to assess single outcomes of science (e.g., manuscripts, [2]), post-academic science is characterised by the use of quantitative methods of research evaluation. According to Wilsdon, et al. [3], there are currently “three broad approaches to the assessment of research: a metrics-based model; peer review; and a mixed model, combining these two approaches. Choosing between these remains contentious” (p. 59).

Typical metrics are publications and citations [4]. For example, the government of Mexico follows a metrics-based model by allocating funds to higher education institutions with several indicators [5].

A special characteristic of research evaluation in the area of post-academic science is the emergence of university rankings. Here, metrics are used to rank the universities in a country or worldwide [6]. University rankings have some obvious advantages. They offer, for example, a quick, simple, and easy way of comparing universities (worldwide). The most interested groups in the rankings are students, the public and governments [3]. However, a lot of critiques have been published in recent years (see e.g., Reference [7]) that focus on the methods and arbitrary weightings used to combine different metrics. Daraio, et al. [8] cited four points summarizing the main criticisms aimed at rankings: mono-dimensionality, statistical robustness, dependence on university size and subject mix, and lack of consideration of the input–output structure.

In this scientometric study, we pick up the last point “lack of consideration of the input–output structure” and set a possible approach for input–output consideration in institutional evaluation to discuss (in scientometrics). Since positions in rankings depend on certain context factors [9,10], rankings should not only offer information on the output, but also the relation of input to output. Moed and Halevi [11] define input indicators as follows: “indicators that measure the human, physical, and financial commitments devoted to research. Typical examples are the number of (academic) staff employed or revenues such as competitive, project funding for research” (p. 1990).

If metrics are used that relate output to input (e.g., the number of papers per full time equivalent researcher), research efficiency is measured. Thus, this study is intended to explore approaches of measuring the efficiency of universities. The study follows on from a recent discussion in the *Journal of Informetrics*, which started with Abramo’s and D’Angelo’s [12] doubts about the validity of established bibliometric indicators and the comments that ensued. Instead, they plead in favor of measuring scientific efficiency. For example, they proposed the Fractional Scientific Strength (FSS) indicator, which is a composite indicator that considers the total salary of the research staff and the total number of publications weighted with citation impact (when used on the university level).

## 2. Conceptual Framework

This study follows the call of Bornmann and Haunschild [13] and Waltman, et al. [14] who propose in a comment on the paper by Abramo’s and D’Angelo’s [12] that scientometricians should try to explore methods and available data to measure the efficiency of research. We do this both by using a unique data set and applying approaches rarely used in academic efficiency analysis. The former comprises information for the top 50 US American universities from the Times Higher Education (THE) Ranking 2015. Input is defined by annual research expenses. The output concerns the 1% most frequently cited publications in a specific field and given year ( $P_{\text{top } 1\%}$ ). The focus on these top publications is derived from the fact that we focus on elite universities represented by the 50 best ranked universities in the THE. Whereas the input we used is standard in the literature, our output variable has never been used before (to the best of our knowledge)—although it is very suitable to study the efficiency of top institutions. Other output variables such as total number of publications, number of graduates or third-party funding could have equally been considered. However, we had three reasons to focus on highly-cited papers: (1) Reputable universities should be evaluated with indicators which focus on excellent research. (2) Gralka, et al. [15] and other studies have shown that conclusions are very similar if (top-cited) publications, third-party funding or other indicators are used in the efficiency analysis. (3) We wanted to trace out the ‘pure’ research effect in efficiency analysis. Including, for instance, teaching-oriented variables would divert from this goal.

University rankings identify top universities from around the world using various indicators. We initially had the idea to undertake a global university analysis by including not only US universities, but also top universities from other countries. This would require an international database with comparable input and output data. Whereas these data are available on the publication output side (with, e.g., the Scopus database from Elsevier), they are not available on the input side.

We focus, therefore, in this study on US universities for which comparable data are available in a national database.

The most frequently used tool in the academic efficiency literature is the Data Efficiency Analysis (DEA) and variations of this non-parametric approach. The DEA yields an institutional efficiency score between 0 and 1, where 1 means efficient. However, these non-parametric approaches have several shortcomings. There is no well-defined data-generating process, and a deterministic approach is assumed: “Any deviation from the frontier is associated with inefficiency, and it is not possible to take into consideration casual elements or external noise which might have affected the results” [16]. The most serious drawback of the DEA in its simplest form is that it is extremely vulnerable to outliers and measurement errors.<sup>1</sup>

Thus, we further employ the Free Disposal Hull (FDH), which is less prone to outliers, and apply the partial frontier analysis (PFA), which nests FDH and DEA. Specifically, we employ the order- $m$  [17] and order- $\alpha$  [18] approaches. Here, the sensitivity to outliers and measurement errors is reduced by allowing for super-efficient universities with efficiency scores larger than 1. To this end, sub-samples of the data are used and resampling techniques are employed. The use of four different approaches allows us to validate the robustness of our conclusions. Finally, we calculate efficiency scores adjusted for institutional background and research focus.

There is plenty of literature examining the efficiency of (higher) education institutions. Early examples are Lindsay [19] and Bessent, et al. [20]. Worthington [21] and more recently Rhaïem [22] as well as De Witte and López-Torres [23] provide comprehensive surveys of the literature. In the efficiency analyses of higher education institutions, PFA has rarely been used. Bonaccorsi, et al. [24,25] applied the order- $m$  approach to study 45 universities in Italy and 261 universities across four European countries, respectively. De Witte, et al. [26] used DEA and PFA to study the performance of 155 professors working at a Business & Administration department of a Brussels university college. Bruffaerts, et al. [27] used FDH and PFA to study the efficiency of 124 US universities. The authors tried to explain which factors drove the efficiency scores. However, they do not provide scores for each university. Gnewuch and Wohlrabe Gnewuch and Wohlrabe [28] used partial frontier analysis to identify super-efficient economics departments. There are several studies available in the literature which have investigated efficiency aspects in the US higher education system [29–34].

The paper is organized as follows: It starts by explaining the four statistical approaches used in this study for calculating the efficiency scores of the universities. The paper subsequently describes the data set and provides some descriptive statistics. In the first step, we calculate efficiency scores for the universities. In the second step, we calculate adjusted efficiency scores. These scores are adjusted to the different profiles of the universities (e.g., their disciplinary profiles). After presenting our results, we discuss the implications of our analysis.

### 3. Methods

#### 3.1. (Partial) Academic Production Frontier Analysis

The main goal of efficiency measurement is to calculate an efficiency score for each unit (here: each university). There are two main concepts: (1) input-orientated efficiency, where the output is set constant and the inputs are adjusted accordingly; (2) output-orientated efficiency, where for a given input the output is maximized. These concepts differ in terms of the direction in which the distance of a university from the efficiency frontier is measured. In this paper, we resort to input-efficiency and variable returns to scale (VSR). With respect to the former point, we could also consider output-efficiency as US universities may have control over both the inputs (the acquired budget) and

---

<sup>1</sup> There are also parametric approaches available (e.g., the stochastic frontier analysis, SFA), which have several disadvantages too. One disadvantage is that they rely on distributional assumptions; a specific functional form is required. The potential endogeneity of inputs cannot be accounted for.

the outputs. In our estimation framework, we cannot test the nature of economies of scale. The partial frontier approaches used in this paper assume constant returns to scale. Furthermore, the results do not point to evidences of how the production process with respect to top-cited publications works.

We start this section by describing two full production frontier approaches for elicitation of academic efficiency scores: The most commonly used DEA and the less known FDH approach. We subsequently outline two PFA: order- $m$  and order- $\alpha$ . Both techniques are generalizations of the FDH approach, as they nest it. Both approaches allow for the existence of super-efficient universities, i.e., universities with efficiency scores larger than 1. In Section 3.1.4, we illustrate the four approaches with a simple example.

We denote the input and output of a university  $i$  with  $x_i$  and  $y_i$ , respectively. We consider  $N$  universities. The corresponding efficiency score is given by  $e_i$ .

### 3.1.1. Data Envelopment Analysis (DEA)

DEA was introduced by Charnes, et al. [35]. It is a linear programming approach which envelopes the data by a piecewise-linear convex hull. The DEA efficiency score  $e_i^{DEA}$  solves the following optimization problem:

$$\begin{aligned} & \min_{e, \lambda} e \text{ subject to} \\ & e \cdot x_{mi} - \sum_{j=1}^N \lambda_j x_{mj} \geq 0 \quad m = 1, \dots, M \\ & \sum_{j=1}^N \lambda_j y_{lj} - y_{li} \geq 0 \quad l = 1, \dots, L \\ & \lambda_j \geq 0 \quad \forall j \end{aligned} \tag{1}$$

where  $\lambda$  is a weighting parameter that maximizes the productivity. In this paper, we focus on the basic version of the DEA. With respect to outliers, sampling and measurement data issues we focus on the later introduced partial frontier analysis. For (robust) extensions of the DEA, we refer to Bogetoft and Otto [36] and Wilson and Clemson [37].

We compare each university  $i$  with every other university in the data set ( $j = 1 \dots N$ ). The set of peer universities that satisfy the condition  $y_{lj} \geq y_{li} \quad \forall l$  is denoted by  $B_i$ . Among the peer universities, the one that exhibits the minimum input serves as a reference to  $i$ , and  $e_i^{FDH}$  is calculated as the relative input use

$$\hat{e}_i^{FDH} = \min_{j \in B_i} \left\{ \max_{k=1, \dots, K} \left( \frac{x_{kj}}{x_{ki}} \right) \right\} \tag{2}$$

Universities that exhibit the minimum input–output serve as references. For these universities, the efficiency score  $e_i^{FDH}$  is 1. The FDH approach was introduced by Deprins, et al. [38].

### 3.1.2. Order- $m$ Efficiency

In case of order- $m$  efficiency, the partial aspect comes in by departing from the assumption that the universities are benchmarked on the basis of the best-performing universities in the sample. Instead, the best performance of a sample including  $m$  peers is considered. Daraio and Simar [39] proposed the following four-step procedure:

1. Draw from  $B_i$  a random sample of  $m$  peer universities with replacement.
2. A pseudo-FDH efficiency score ( $\hat{e}_{mi}^{FDH_d}$ ) is calculated using the artificially drawn data.
3. Repeat steps 1 and 2  $D$  times.
4. Order- $m$  efficiency is calculated as the average of the pseudo-FDH scores

$$\hat{e}_{mi}^{OM} = \frac{1}{D} \sum_{d=1}^D \hat{e}_{mi}^{FDH_d} \tag{3}$$

A potential result of this procedure is that the order- $m$  efficiency scores exceed the value of 1. This is due to the resampling: In each replication  $d$ , university  $i$  may or may not be used for its own comparison. Therefore, this procedure allows for super-efficient universities (with  $\hat{e}_{mi}^{OM} > 1$ ) located beyond the estimated production-possibility frontier. There are two parameters that need to be determined beforehand:  $m$  and  $D$ .  $D$  is just a matter of accuracy. The higher  $D$  is, the more accurate are the results. It prolongs the computational time only. The choice of  $m$  is more critical. The smaller  $m$  is, the larger is the share of super-efficient universities. For  $m \rightarrow \infty$  the approach converges to the FDH results.

### 3.1.3. Order- $\alpha$ Efficiency

The order- $\alpha$  approach generalizes the FDH otherwise. Instead of searching for the minimum input-output relationship among the available peer universities (the benchmark), order- $\alpha$  uses the  $(100 - \alpha)$ th percentile

$$\hat{e}_{\alpha i}^{OA} = P_{(100-\alpha)} \left\{ \max_{j \in B_i} \left( \frac{x_{kj}}{x_{ki}} \right) \right\} \tag{4}$$

When  $\alpha = 100$ , the approach replicates the FDH results. In case of  $\alpha < 100$ , some universities may be classified as super-efficient. As  $m$  is the approach explained in Section 3.1.2,  $\alpha$  can be considered as a tuning parameter: the smaller  $\alpha$  is, the larger is the share of the super-efficient universities.

### 3.1.4. A Simple Example for Explaining the Approaches

To illustrate the outlined full and PFA approaches, we sketched them out in Figure 1. We plotted input-output combinations for various universities. The results of the DEA are given by the straight line. Universities A, B and E define the academic production frontier. These universities have an efficiency score of 1, i.e., an optimal input-output combination. The other universities on the right of or below the frontier are considered as inefficient. In case of the FDH, the outer hull is spanned in more explicitly by also considering universities that are not on the DEA curve. In Figure 1, universities C and D are also efficient now. Since the frontier has shifted towards the right, the efficiency scores for all other universities slightly increase. The distance to the frontier is smaller than for the DEA.

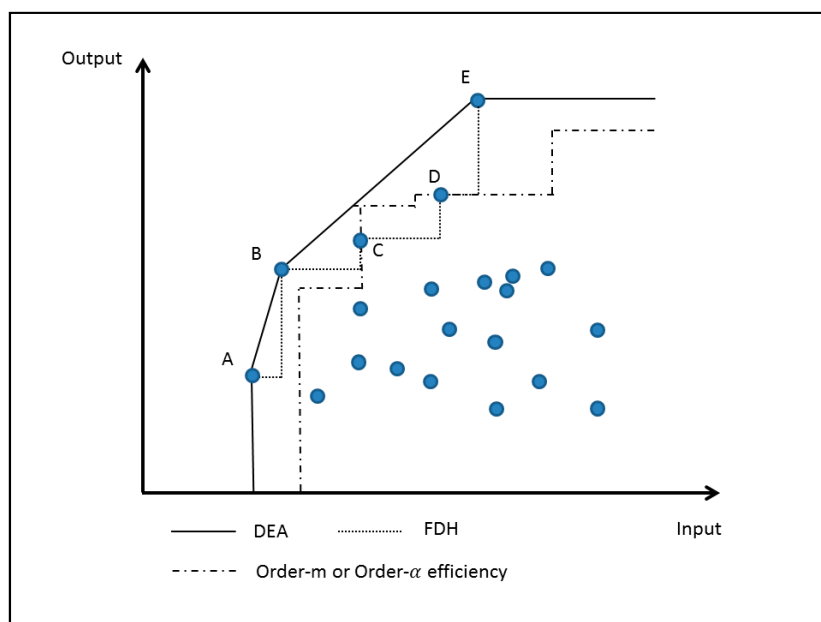


Figure 1. Graphical exposition of full and partial frontier efficiency analysis.

Applying the partial frontier approaches, order-m or order- $\alpha$ , we get a different picture. Only universities C and D are efficient with a corresponding score of 1; universities A, B and E are considered as super-efficient with a score larger than 1. Of course, both approaches do not necessarily yield the same results, as the figure might suggest.

### 3.1.5. Regression Analyses and Adjusted Efficiency Scores

We performed regression analyses to produce adjusted efficiency scores for the universities. Since the universities have different profiles, the scores from the regression analyses are adjusted to these differences. Thus, the focus of the regression analyses is not on explaining the variance of the scores (as done, e.g., by Reference Agasisti and Wolszczak-Derlacz [40]). We used Stata [41] to compute the regression analyses.

The efficiency scores from the four approaches (explained above) are the dependent variable in the model. Four indicators are included as independent variables in the models, which reflect the disciplinary profile of the university. We expect that the disciplinary profile is related to the efficiency of a university. The results of Bornmann, et al. [42] show that the field-normalized citation impact of universities depends on the disciplinary profile. For each university, we searched for the number of publications in four broad disciplines and the multidisciplinary field in the SCImago Institutions Ranking.<sup>2</sup> For each institution, the percentages of publications that belong to the four disciplines were calculated and included in the regression model (mean centered). As a further independent variable, the binary information is considered for whether the institution is a public (0) or private (1) university. Private universities tend to be elite research institutions. More than these two indicators are not available in the SCImago Institutions Ranking, which was used in the regression analyses to reflect the profiles of universities.

We used the cluster option in Stata to consider in the regression analysis that the universities are in different US states. With 10 universities, the most universities are located in California. The different regulations and financial opportunities in the states probably lead to related efficiency scores for universities within one state. The cluster option corrects the standard errors for the fact that there are up to 10 universities in each state. Although the point estimates of the coefficients are the same as in the regression model without the option, the standard errors are typically larger [43].

### 3.2. Data

For our case study, we gathered input and output data for the 50 best performing US universities as listed in the THE Ranking 2015. As input we used research expenses. The data source is the National Center for Education Statistics (NCES).<sup>3</sup> The NCES gathers data from universities by applying uniform data definitions. This ensures the comparability of inputs across universities, which is an important requirement of efficiency studies [44,45]. The expenses are self-reported data by the universities. The category includes institutes and research centers, as well as individual and project research. Information technology expenses related to research activities are also considered if the institution separates budgets and expenses information technology resources. Universities are asked to report actual or allocated costs for operation and maintenance of plant, interest and depreciation. The data refer to the academic year, which starts on 1 July and ends on 30 June. As we needed information for three calendar years (the output data refer to the calendar years 2011, 2012 and 2013), we transformed the data. As an example, we obtained the input data for 2013 by taking the mean of the data from the

---

<sup>2</sup> See <http://www.scimagoir.com>.

<sup>3</sup> The data are from <http://nces.ed.gov/ipeds/datacenter/InstitutionProfile.aspx?unitid=adafaeb2afaf>. The database provides also research staff figures, which could have been considered additionally in our study. However, the reported figures do not seem to be consistent. In some cases, the reported research staff were far too low compared to the overall staff of a university. For other universities, numbers varied substantially over time.

academic year 2013/14 and 2012/2013.<sup>4</sup> This approach might introduce some unknown biases as we assume that the expenses are being spent evenly across the year. So, we cannot assure that the research expenses represent correctly the production process of a university. Potential measurement errors are further reasons to employ PFA. In the best case, biases cancel out across the sample.

As we focus on the best US universities, we use as output the number of papers that belong to the 1% ( $P_{\text{top 1\%}}$ ) most frequently cited papers in the corresponding fields and publication years. The use of this indicator ensures that the citation impact of all papers is standardized with respect to the year and subject area of publication. The typical output variables in efficiency analysis are students, graduates and funding; publications are used rather seldom [46,47]. The data were obtained from the SCImago Institutions Ranking, which is based on Scopus data.<sup>5</sup> The output data refer to the publication period from 2011 to 2013 with a citation window from publication until the end of 2015. We did not use data later than 2013 since it is standard to use a citation window of at least 3 years in bibliometrics [48]. In Section 4, we focus on the results for 2013. Both other publication years allowed us to look at the stability of the results.

Table 1 shows the descriptive statistics both for the input and output from 2011 to 2013. The dataset is fairly heterogeneous as the difference between minimum and maximum indicates. Furthermore, the standard deviation is quite large compared to the mean. The distributions of the variables are not significantly skewed as mean and median are very close together. The development over time points out that research expenses increase whereas the average  $P_{\text{top 1\%}}$  peaked in 2012 and dropped considerably in 2013. The correlation coefficients between research expenses and  $P_{\text{top 1\%}}$  are relatively constant over time. All coefficients are about 0.6 implying a moderate positive relationship.

**Table 1.** Descriptive statistics over time.

	2011		2012		2013	
	Research Expenses	$P_{\text{top 1\%}}$	Research Expenses	$P_{\text{top 1\%}}$	Research Expenses	$P_{\text{top 1\%}}$
Mean	514	254	521	277	523	225
Median	482	213	483	228	477	197
Standard Deviation	289	160	298	182	303	151
Minimum	38	35	37	19	36	24
Maximum	1265	1002	1321	1198	1372	977

Notes. Descriptive statistics for the input and output are reported. Research expenses in Million Dollars.

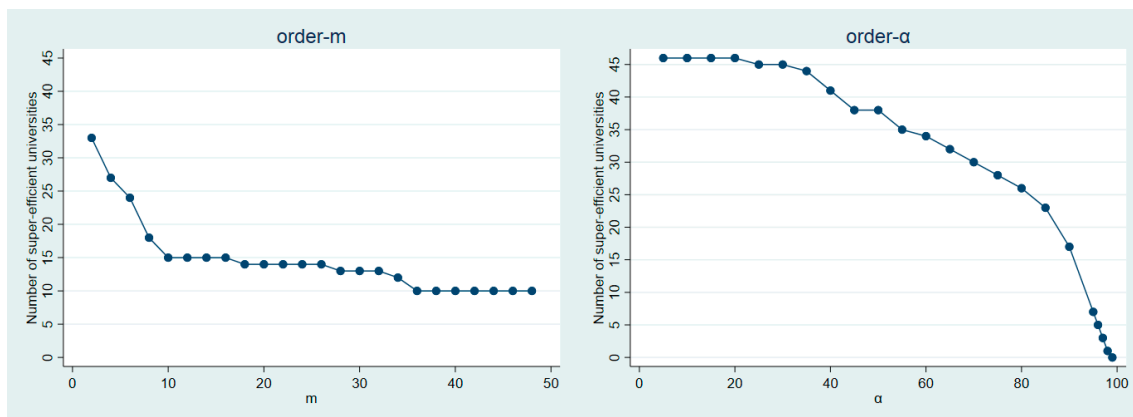
#### 4. Results

Following the methods as outlined in Section 3.1, we estimated four efficiency scores for each university and year in our data set and obtained the corresponding efficiency rankings for 2011 to 2013.

In contrast, PFA requires the specification of parameters, which eventually influence the amount of super-efficient universities. The order- $\alpha$  approach requires  $\alpha$ , the percentile of the set of peer universities used as the benchmark. Order- $m$  requires  $m$ , the number of peer universities randomly drawn from the initial set of universities. Unless we set  $m = 50$  or  $\alpha = 100$ , where the partial frontier approaches converge to FDH, we find super-efficient universities by construction. Figure 2 shows the number of super-efficient universities for different values of  $m$  and  $\alpha$ . We used data for the year 2013. It is clearly visible that with higher  $m$  or  $\alpha$  values, respectively, the number of super-efficient universities becomes lower. Concerning  $m$ , the figure is quite stable beyond 35. We opted to set  $m = 40$  and  $\alpha = 95\%$ , which yield 10 and 7 super-efficient universities, respectively.

<sup>4</sup> There are a few exceptions ( $n = 7$ ) where the academic year differs slightly across years. We adjusted the figures accordingly.

<sup>5</sup> See <http://www.scimagoir.com>. We preferred Scopus over Web of Science data as the coverage of the Scopus database is much broader.



**Figure 2.** Number of super-efficient universities for different values of  $m$  and  $\alpha$ .

4.1. Baseline Results

4.1.1. Results for 2013

Table 2 reports the efficiency scores with their corresponding rankings based on the data from 2013. The universities are sorted by their ranking positions in the THE Ranking 2015. In 2013, 48 universities are not efficient according to the DEA results. This number drops to 38 universities using the FDH approach. This is because both Harvard and Boston College dominate the estimated academic efficiency frontier: Harvard University due to its very high output values and Boston College due to its small input values relative to the outputs. With respect to the order- $\alpha$  framework, there are seven universities with a score larger than 1. This number is increased to 10 based on the order- $m$  approach. In the majority of cases, the order- $\alpha$  approach yields higher scores compared to the order- $m$  approach.

In the order- $\alpha$  framework and the order- $m$  approach, Harvard University remains efficient with a score of 1.00 and is not denoted as super-efficient. However, Boston College is super-efficient with the highest corresponding score both for the order- $\alpha$  and order- $m$  model.

Based on a Stochastic Frontier analysis, Agasisti and Johnes [33] also reported efficiency scores for various US universities, but with the number of bachelors and postgraduate degrees on the output side. Similar to our results, the authors found Harvard University at the top but Boston College is not among their 20 best universities.

**Table 2.** Efficiency scores and the corresponding rankings based on different approaches for measuring efficiency in 2013 (sorted by THE Ranking).

THE	University	DEA		FDH		Order- $\alpha$		Order- $m$	
		Score	Rank	Score	Rank	Score	Rank	Score	Rank
1	California Institute of Technology	0.680	6	1.000	1	1.264	6	1.059	5
2	Harvard University	1.000	1	1.000	1	1.000	8	1.000	11
3	Stanford University	0.382	28	0.765	20	0.765	30	0.765	23
4	Massachusetts Institute of Technology	0.268	43	0.619	31	0.619	38	0.619	35
5	Princeton University	0.473	15	0.614	32	0.866	24	0.704	29
6	University of California, Berkeley	0.414	23	1.000	1	1.000	8	1.000	9
7	Yale University	0.444	18	0.720	25	0.720	34	0.732	27
8	University of Chicago	0.577	11	1.000	1	1.077	7	1.023	6
9	University of California, Los Angeles	0.409	24	0.897	14	0.897	20	0.897	16
10	Columbia University	0.475	14	1.000	1	1.000	8	1.000	10
11	Johns Hopkins University	0.249	45	0.584	34	0.584	41	0.584	37
12	University of Pennsylvania	0.460	17	1.000	1	1.000	8	1.000	11
13	University of Michigan, Ann Arbor	0.359	32	0.794	17	0.794	27	0.794	19
14	Duke University	0.309	35	0.766	19	0.766	29	0.766	22
15	Cornell University	0.656	7	1.000	1	1.000	8	1.014	8



Table 2. Cont.

THE	University	DEA		FDH		Order- $\alpha$		Order-m	
		Score	Rank	Score	Rank	Score	Rank	Score	Rank
16	North-western University, Evanston	0.545	13	1.000	1	1.000	8	1.019	7
17	Carnegie Mellon University	0.381	29	0.486	40	0.892	21	0.639	34
18	University of Washington	0.354	33	0.749	23	0.749	33	0.749	25
19	Georgia Institute of Technology	0.208	49	0.361	48	0.456	47	0.389	49
20	University of Texas, Austin	0.313	34	0.476	42	0.602	39	0.507	42
21	University of Illinois at Urbana-Champaign	0.261	44	0.349	49	0.492	45	0.406	48
22	University of Wisconsin, Madison	0.225	48	0.402	46	0.402	50	0.413	47
23	University of California, Santa Barbara	0.618	8	0.902	13	1.000	8	0.974	14
24	New York University	0.273	41	0.472	43	0.472	46	0.484	43
25	University of California, San Diego	0.306	36	0.761	21	0.761	31	0.762	24
26	Washington University in Saint Louis	0.439	19	0.760	22	0.760	32	0.779	20
27	University of Minnesota, Twin Cities	0.240	46	0.416	44	0.448	48	0.425	46
28	University of North Carolina, Chapel Hill	0.364	31	0.594	33	0.594	40	0.605	36
29	Brown University	0.799	4	1.000	1	1.834	2	1.329	2
30	University of California, Davis	0.276	40	0.512	38	0.551	42	0.524	40
31	Boston University	0.782	5	1.000	1	1.410	3	1.156	3
32	Pennsylvania State University	0.192	50	0.329	50	0.416	49	0.354	50
33	Ohio State University, Columbus	0.386	26	0.713	26	0.879	22	0.738	26
34	Rice University	0.805	3	1.000	1	1.383	4	1.125	4
35	University of Southern California	0.468	16	0.742	24	0.939	18	0.799	18
36	Michigan State University	0.299	37	0.477	41	0.672	36	0.521	41
37	University of Arizona	0.283	39	0.388	47	0.547	43	0.450	44
38	University of Notre Dame	0.603	10	0.856	15	1.000	8	0.937	15
39	Tufts University	0.554	12	0.689	27	0.953	17	0.776	21
40	University of California, Irvine	0.414	22	0.578	35	0.815	25	0.680	30
41	University of Pittsburgh	0.289	38	0.534	37	0.658	37	0.554	38
42	Emory University	0.399	25	0.625	29	0.790	28	0.667	32
43	Vanderbilt University	0.431	21	0.793	18	0.977	16	0.818	17
44	University of Colorado, Boulder	0.432	20	0.572	36	0.806	26	0.657	33
45	Purdue University	0.383	27	0.661	28	0.932	19	0.721	28
46	University of California, Santa Cruz	0.618	9	0.826	16	1.368	5	0.982	13
47	Case Western Reserve University	0.272	42	0.495	39	0.698	35	0.544	39
48	University of Rochester	0.368	30	0.620	30	0.874	23	0.677	31
49	Boston College	1.000	1	1.000	1	3.018	1	1.856	1
50	University of Florida	0.237	47	0.405	45	0.511	44	0.435	45

Table 3 shows the coefficients for the correlation between the ranking positions of the universities in the THE Ranking 2015 and the results of the efficiency analyses. The results point out that the ranking positions from the efficiency analysis are correlated at a (very) low level compared to the correlations among the different results of the efficiency analyses. The results of the four efficiency approaches are highly correlated, implying that one can derive similar conclusions.

Table 3. Spearman rank correlation coefficients for 2013.

	THE	DEA	FDH	order- $\alpha$	order-m
THE	1.000				
DEA	0.073	1.000			
FDH	0.299	0.840	1.000		
order- $\alpha$	0.035	0.927	0.890	1.000	
order-m	0.205	0.899	0.980	0.942	1.000

#### 4.1.2. Stability of the Results over Time

Table 4 reports the rank correlations across time (2011, 2012 and 2013) for each approach of the efficiency analysis. They are all above 0.8, suggesting that the results are quite stable over the observed time period.

**Table 4.** Spearman rank correlations for each approach of efficiency analysis across time.

	DEA			FDH			
	2011	2012	2013	2011	2012	2013	
2011	1.00			2011	1.00		
2012	0.95	1.00		2012	0.84	1.00	
2013	0.93	0.96	1.00	2013	0.85	0.84	1.00
	order- $\alpha$			order-m			
2011	1.00			2011	1.00		
2012	0.91	1.00		2012	0.86	1.00	
2013	0.88	0.91	1.00	2013	0.87	0.89	1.00

## 4.2. Adjusted Scores and Ranking Positions

### 4.2.1. Results for 2013

The results of the regression analyses are shown in Table 5. As dependent variables, the efficiency scores from Table 2 are used (results from the DEA, FDH, order- $\alpha$  approach, and order-m approaches). We estimated linear regressions because the residuals were approximately normally distributed (as tested with the *sktest* in Stata). The coefficients for all disciplines point out that a decrease in the share of publications is associated with higher efficiency scores. If expensive research is done by the university, its efficiency is decreasing. Thus, a high share of paper output especially in physical and health sciences—having the largest coefficients—is related to lower efficiency scores. Furthermore, the results in Table 5 demonstrate that private universities are more efficient than public universities. Many coefficients in the models are statistically not significant (which might be the result of the low numbers of universities in the study).

**Table 5.** Beta coefficients and t statistics of the regression models with the efficiency scores as dependent variable for 2013.

	DEA	FDH	order- $\alpha$	order-m
Life sciences	−0.45 (−01.90)	−01.20 * (−02.41)	−00.46 (−00.65)	−00.99 (−01.83)
Physical sciences	−01.43 * (−02.26)	−03.65 * (−02.52)	−01.39 (−00.67)	−02.99 (−01.91)
Social sciences	−00.35 (−01.90)	−01.02 * (−02.36)	−00.17 (−00.27)	−00.74 (−01.54)
Health sciences	−01.08 * (−02.51)	−02.60 * (−02.53)	−01.08 (−00.73)	−02.16 (−01.95)
Private state	0.18 *** (3.94)	0.14 (1.77)	0.33 ** (2.89)	0.21 * (2.48)
Constant	0.34 *** (10.36)	0.63 *** (10.38)	0.69 *** (7.65)	0.65 *** (8.82)
Universities	50	50	50	50

Notes. *t* statistics in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Subsequent to the regression models, we calculated efficiency scores for every university, which are adjusted by the influence of the independent variables. Thus, the scores are adjusted to the different institutional and field-specific profiles of the universities. It is worth noting that the adjusted scores are not predicted values, but institutional values for which the residuals from the regression analyses were added to the mean initial efficiency scores.

The adjusted ranking positions (based on the adjusted scores) are listed in Table 6 besides the initial ranking positions. Although both ranking positions are (highly) correlated (DEA:  $r_s = 0.71$ , FDH:  $r_s = 0.82$ , order- $\alpha$ :  $r_s = 0.64$ , order-m:  $r_s = 0.79$ ), there are significant rank changes for some universities. For example, Harvard University shows a perfect rank position in the FDH; but if the score is adjusted by the independent variables in the regression model, its score decreases, leading to the 17th ranking position.

**Table 6.** Initial efficiency rank positions and adjusted rank positions in 2013 (sorted by adjusted DEA scores).

University	DEA	DEA Adjust.	FDH	FDH Adjust.	order- $\alpha$	order- $\alpha$ Adjust.	order-m	order-m Adjust.
Harvard University	1	1	1	17	8	17	11	16
Brown University	4	2	1	6	2	1	2	1
Rice University	3	3	1	5	4	5	4	6
Boston University	5	4	1	9	3	6	3	8
University of California, Santa Barbara	8	5	13	3	8	11	14	7
University of California, Santa Cruz	9	6	16	14	5	3	13	3
California Institute of Technology	6	7	1	7	6	4	5	4
Tufts University	12	8	27	23	17	8	21	18
Boston College	1	9	1	19	1	2	1	2
Cornell University	7	10	1	8	8	22	8	14
University of California, Los Angeles	24	11	14	1	20	7	16	5
Ohio State University, Columbus	26	12	26	13	22	9	26	12
University of California, Irvine	22	13	35	31	25	14	30	21
University of Michigan, Ann Arbor	32	14	17	12	27	10	19	10
Northwestern University, Evanston	13	15	1	4	8	21	7	11
University of North Carolina, Chapel Hill	31	16	33	28	40	33	36	30
University of Pittsburgh	38	17	37	24	37	15	38	20
Purdue University	27	18	28	15	19	13	28	15
University of Notre Dame	10	19	15	11	8	47	15	23
University of Colorado, Boulder	20	20	36	35	26	26	33	34
University of Chicago	11	21	1	20	7	32	6	24
University of Pennsylvania	17	22	1	10	8	19	11	13
University of Southern California	16	23	24	22	18	30	18	27
Washington University in Saint Louis	19	24	22	29	32	28	20	25
Vanderbilt University	21	25	18	18	16	18	17	19
University of California, Davis	40	26	38	26	42	20	40	22
Emory University	25	27	29	34	28	25	32	33
University of Arizona	39	28	47	40	43	36	44	38
University of California, Berkeley	23	29	1	2	8	16	9	9
Georgia Institute of Technology	49	30	48	32	47	29	49	32
University of Florida	47	31	45	30	44	27	45	29
University of California, San Diego	36	32	21	21	31	12	24	17
University of Minnesota, Twin Cities	46	33	44	41	48	40	46	41
University of Texas, Austin	34	34	42	38	39	43	42	44
University of Rochester	30	35	30	36	23	23	31	35
Columbia University	14	36	1	16	8	37	10	28
University of Washington	33	37	23	25	33	31	25	26
University of Wisconsin, Madison	48	38	46	42	50	41	47	40
Case Western Reserve University	42	39	39	44	35	24	39	36
Michigan State University	37	40	41	33	36	42	41	43
Yale University	18	41	25	43	34	46	27	45
University of Illinois at Urbana-Champaign	44	42	49	46	45	45	48	46
Carnegie Mellon University	29	43	40	45	21	38	34	42
Johns Hopkins University	45	44	34	39	41	35	37	37
Duke University	35	45	19	27	29	34	22	31
Stanford University	28	46	20	37	30	39	23	39
Princeton University	15	47	32	48	24	48	29	48
Pennsylvania State University	50	48	50	47	49	49	50	49
New York University	41	49	43	50	46	50	43	50
Massachusetts Institute of Technology	43	50	31	49	38	44	35	47

#### 4.2.2. Stability of the Results over Time

Table 7 shows the Spearman rank correlation coefficients across time (2011, 2012 and 2013) for each approach of the efficiency analysis (adjusted scores). The coefficients are above or around 0.8, which demonstrate that the results are (more or less) stable over the publication years considered.

**Table 7.** Spearman rank correlations for the adjusted scores from each approach across time.

		DEA			FDH			
		2011	2012	2013	2011	2012	2013	
2011	1.00				2011	1.00		
2012	0.92	1.00			2012	0.65	1.00	
2013	0.91	0.93	1.000		2013	0.75	0.72	1.00
		order- $\alpha$			order-m			
2011	1.00				2011	1.00		
2012	0.86	1.00			2012	0.76	1.00	
2013	0.88	0.89	1.00		2013	0.82	0.82	1.00

## 5. Discussion

Research evaluation is the backbone of modern science. The emergence of the modern science system is closely related to the introduction of the peer review process in assessments of research results [2]. Whereas output and impact assessments were initially at the forefront of assessments, efficiency measurements have become popular in recent years [22]. According to Moed and Halevi [11], research efficiency or productivity is measured by indicators that relate research output to input. The consideration of research input in research evaluation is obvious, since the output should be directly related to the input. The output is determined by the context in which research is undertaken [22,49]. In this study, we went one step further. We not only related input to output for universities, but also calculated adjusted efficiency scores, which consider the different institutional and field-specific profiles of the universities. For example, it is easily comprehensible that the input–output relations are determined by the disciplinary profiles of the universities.

The present study is based on a comprehensive dataset with input and output data for 50 US universities. As input, we used research expenses, and as output the number of (highly-cited) papers. The results of the DEA and FDH analysis show that Harvard University and Boston College can be called especially efficient—compared with many other universities. Similar results can be found in other efficiency studies including US institutions. Whereas the strength of Harvard University is its high output of (highly-cited) papers, the strength of Boston College is its small input. In the order- $\alpha$  and order-m frameworks, Harvard University remains efficient, but Boston College becomes super-efficient. Although Harvard University is well-known as belonging to the best universities in the world, the correlations between the ranking positions of the universities in the THE Ranking 2015 and the results of our efficiency analyses are at a relatively low level. Thus, the consideration of inputs puts a different complexion on institutional performance.

Besides the university rankings based on the different statistical approaches for efficiency analyses, we produced rankings using adjusted efficiency scores (subsequent to regression analyses). Here, for example, Harvard University's ranking position fell. Although regression analyses have been used in many other efficiency studies, they have been commonly used to explain the differences in efficiency scores [22], but not to generate adjusted scores (for rankings). The adjusted rankings open up new possibilities for institutional performance measurements, as demonstrated by Bornmann, et al. [9]. They produced a covariate-adjusted ranking of research institutions worldwide in which single covariates are held constant. For example, the user of the ranking produced by Bornmann, et al. [9] is able to identify institutions with a very good performance (in terms of highly cited papers), despite a bad financial situation in the corresponding countries.

What are the limitations of the current study? Although we tried to realize an advanced design of efficiency analyses, the study is affected by several limitations that should be considered in future studies.

The first limitation is related to the numbers of indicators used. We included only one input and output indicator, respectively. One important reason for this restriction is the focus of this study on efficiency in research. However, many more indicators could be included in future studies (if the focus is broader and not limited to excellent research as in this study). The efficiency study of

Bruffaerts, et al. [27], which also focuses on US universities, additionally included the number of PhD degrees as input indicators, as well as several environmental variables (e.g., university size and teaching load). In an overview of efficiency studies, Rhaiem [22] categorized possible research output indicators for efficiency analyses as follows: research outputs, research productivity indices and quality of research indicators. The categorizations for possible input indicators are: “Firstly, human capital category refers to academic staff and non-academic staff; secondly, physical capital category refers to productive capital (building spaces, laboratories, etc.); thirdly, research funds category encompasses budget funds and research income; fourthly, operating budget refers to income and current expenditures; fifthly, stock of cumulative knowledge regroups three sub-categories: knowledge embedded in human resources, knowledge embedded in machinery and equipment, and public involvement in R&D; sixthly, agglomeration effects category refers to regional effect and entrepreneurial environment” (p. 595).

The second limitation concerns the quality of the input data [14]. “Salary and investment financial structures differ hugely between countries, and salary levels differ hugely between functions, organizations and countries. To paraphrase Belgian surrealism: a salary is not a salary, while a research investment is not a research investment. Comparability (and hence validity) of the underlying data themselves not only is a challenge, it is a problem” [50]. We tried to tackle the problem in this study by using the data for all universities from one source: NCES. However, the comparability of the data for the different universities may remain a problem. Thus, Waltman, et al. [14] recommend that “scientometricians should investigate more deeply what types of input data are needed to construct meaningful productivity indicators, and they should explore possible ways of obtaining this data” (p. 673) in future studies.

The third limitation questions the general implementation of efficiency studies in the practice of research evaluation. The results of the study by Aagaard and Schneider [51] highlight many difficulties in explaining research performance (output and impact) as a linear function of input indicators. Bornmann and Haunschild [13] see efficiency in research as diametric to creativity and faulty incrementalism, which are basic elements of each (successful) research process. According to Ziman [1], “the post-academic drive to ‘rationalize’ the research process may damp down its creativity. Bureaucratic ‘modernism’ presumes that research can be directed by policy. But policy prejudice against ‘thinking the unthinkable’ aborts the emergence of the unimaginable” (p. 330).

**Author Contributions:** Conceptualization, K.W. and L.B.; data collection, F.d.M.A.; formal analysis, K.W. and L.B.; writing—original draft preparation, K.W. and L.B.; writing—review and editing, K.W. and L.B.

**Funding:** This research received no external funding.

**Acknowledgments:** We thank Alexandra Baumann for careful research assistance.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ziman, J. *Real Science. What It Is, and What It Means*; Cambridge University Press: Cambridge, UK, 2000.
2. Bornmann, L. Scientific peer review. *Annu. Rev. Inf. Sci. Technol.* **2011**, *45*, 199–245. [[CrossRef](#)]
3. Wilsdon, J.; Allen, L.; Belfiore, E.; Campbell, P.; Curry, S.; Hill, S.; Jones, R.; Kain, R.; Kerridge, S.; Thelwall, M.; et al. *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*; Higher Education Funding Council for England (HEFCE): Bristol, UK, 2015.
4. National Research Council. *Furthering America’s Research Enterprise*; The National Academies Press: Washington, DC, USA, 2014; p. 216.
5. Tarango, J.; Hernandez-Gutierrez, P.Z.; Vazquez-Guzman, D. Evaluation of Scientific Production in Mexican State Public Universities (2007–2011) Using Principal Component Analysis. *Prof. Inf.* **2015**, *24*, 567–576. [[CrossRef](#)]
6. Hazelkorn, E. *Rankings and the Reshaping of Higher Education. The Battle for World-Class Excellence*; Palgrave Macmillan: New York, NY, USA, 2011.

7. Schmoch, U. The Informative Value of International University Rankings: Some Methodological Remarks. In *Incentives and Performance*; Welpé, I.M., Wollersheim, J., Ringelhan, S., Osterloh, M., Eds.; Springer International Publishing: New York, NY, USA, 2015; pp. 141–154.
8. Daraio, C.; Bonaccorsi, A.; Simar, L. Rankings and university performance: A conditional multidimensional approach. *Eur. J. Oper. Res.* **2015**, *244*, 918–930. [[CrossRef](#)]
9. Bornmann, L.; Stefaner, M.; de Moya Anegón, F.; Mutz, R. What is the effect of country-specific characteristics on the research performance of scientific institutions? Using multi-level statistical models to rank and map universities and research-focused institutions worldwide. *J. Informetr.* **2014**, *8*, 581–593. [[CrossRef](#)]
10. Safón, V. What do global university rankings really measure? The search for the X factor and the X entity. *Scientometrics* **2013**, *97*, 223–244. [[CrossRef](#)]
11. Moed, H.F.; Halevi, G. Multidimensional assessment of scholarly research impact. *J. Assoc. Inf. Sci. Technol.* **2015**, *66*, 1988–2002. [[CrossRef](#)]
12. Abramo, G.; D’Angelo, C.A. A farewell to the MNCS and like size-independent indicators. *J. Informetr.* **2016**, *10*, 646–651. [[CrossRef](#)]
13. Bornmann, L.; Haunschild, R. Efficiency of research performance and the glass researcher. *J. Informetr.* **2016**, *10*, 652–654. [[CrossRef](#)]
14. Waltman, L.; van Eck, N.J.; Visser, M.; Wouters, P. The elephant in the room: The problem of quantifying productivity in evaluative scientometrics. *J. Informetr.* **2016**, *10*, 671–674. [[CrossRef](#)]
15. Gralka, S.; Wohlrabe, K.; Bornmann, L. How to Measure Research Efficiency in Higher Education? Research Grants vs. Publication Output. *J. High. Educ. Policy Manag.* **2018**, in press.
16. Abramo, G.; D’Angelo, C. How do you define and measure research productivity? *Scientometrics* **2014**, *101*, 1129–1144. [[CrossRef](#)]
17. Cazals, C.; Florens, J.P.; Simar, L. Nonparametric frontier estimation: A robust approach. *J. Econom.* **2002**, *106*, 1–25. [[CrossRef](#)]
18. Aragon, Y.; Daouia, A.; Thomas-Agnan, C. Nonparametric frontier estimation: A conditional quantile-based approach. *Econom. Theory* **2005**, *21*, 358–389. [[CrossRef](#)]
19. Lindsay, A.W. Institutional Performance in Higher-Education—The Efficiency Dimension. *Rev. Educ. Res.* **1982**, *52*, 175–199. [[CrossRef](#)]
20. Bessent, A.M.; Bessent, E.W.; Charnes, A.; Cooper, W.W.; Thorogood, N.C. Evaluation of Educational-Program Proposals by Means of Dea. *Educ. Adm. Q.* **1983**, *19*, 82–107. [[CrossRef](#)]
21. Worthington, A.C. An Empirical Survey of Frontier Efficiency Measurement Techniques in Education. *Educ. Econ.* **2001**, *9*, 245–268. [[CrossRef](#)]
22. Rhaïem, M. Measurement and determinants of academic research efficiency: A systematic review of the evidence. *Scientometrics* **2017**, *110*, 581–615. [[CrossRef](#)]
23. De Witte, K.; López-Torres, L. Efficiency in education: A review of literature and a way forward. *J. Oper. Res. Soc.* **2017**, *68*, 339–363. [[CrossRef](#)]
24. Bonaccorsi, A.; Daraio, C.; Simar, L. Advanced indicators of productivity of universities. An application of robust nonparametric methods to Italian data. *Scientometrics* **2006**, *66*, 389–410. [[CrossRef](#)]
25. Bonaccorsi, A.; Daraio, C.; Raty, T.; Simar, L. *Efficiency and University Size: Discipline-Wise Evidence from European Universities*; MPRA Paper 10265; University Library of Munich: Munich, Germany, 2007.
26. De Witte, K.; Rogge, N.; Cherchye, L.; Van Puyenbroeck, T. Accounting for economies of scope in performance evaluations of university professors. *J. Oper. Res. Soc.* **2013**, *64*, 1595–1606. [[CrossRef](#)]
27. Bruffaerts, C.; Rock, B.D.; Dehon, C. *The Research Efficiency of US Universities: A Nonparametric Frontier Modelling Approach*; Working Papers ECARES ECARES 2013-31; Université Libre de Bruxelles: Bruxelles, Belgium, 2013.
28. Gnewuch, M.; Wohlrabe, K. Super-efficiency of education institutions: An application to economics departments. *Educ. Econ.* **2018**, in press. [[CrossRef](#)]
29. Cohn, E.; Rhine, S.L.W.; Santos, M.C. Institutions of Higher-Education as Multi-Product Firms—Economies of Scale and Scope. *Rev. Econ. Stat.* **1989**, *71*, 284–290. [[CrossRef](#)]
30. Harter, J.F.R.; Wade, J.A.; Watkins, T.G. An examination of costs at four-year public colleges and universities between 1989 and 1998. *Rev. High. Educ.* **2005**, *28*, 369–391. [[CrossRef](#)]
31. Laband, D.N.; Lentz, B.F. Do costs differ between for-profit and not-for-profit producers of higher education? *Res. High. Educ.* **2004**, *45*, 429–441. [[CrossRef](#)]

32. Sav, G.T. Stochastic Cost Inefficiency Estimates and Rankings of Public and Private Research and Doctoral Granting Universities. *J. Knowl. Manag. Econ. Inf. Technol.* **2012**, *4*, 11–29.
33. Agasisti, T.; Johnes, G. Efficiency, costs, rankings and heterogeneity: The case of US higher education. *Stud. High. Educ.* **2015**, *40*, 60–82. [[CrossRef](#)]
34. Titus, M.A.; Vamosiu, A.; McClure, K.R. Are Public Master’s Institutions Cost Efficient? A Stochastic Frontier and Spatial Analysis. *Res. High. Educ.* **2017**, *58*, 469–496. [[CrossRef](#)]
35. Charnes, A.; Cooper, W.W.; Rhodes, E. Measuring the Efficiency of Decision-Making Units. *Eur. J. Oper. Res.* **1979**, *3*, 338–339. [[CrossRef](#)]
36. Bogetoft, P.; Otto, L. *Benchmarking with DEA, SFA and R*; Springer: New York, NY, USA, 2011.
37. Wilson, P.W.; Clemson, S.C. *FEAR 2.0: A Software Package for Frontier Analysis with R*; Department of Economics, Clemson University: Clemson, SC, USA, 2013.
38. Deprins, D.; Simar, L.; Tulkens, H. Measuring Labor-Efficiency in Post Offices. In *Public Goods, Environmental Externalities and Fiscal Competition*; Chander, P., Drèze, J., Lovell, C.K., Mintz, J., Eds.; Springer US: Boston, MA, USA, 2006; pp. 285–309.
39. Daraio, C.; Simar, L. *Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications*; Springer: Heidelberg, Germany, 2007.
40. Agasisti, T.; Wolszczak-Derlacz, J. Exploring efficiency differentials between Italian and Polish universities, 2001–2011. *Sci. Public Policy* **2015**, *43*, 128–142. [[CrossRef](#)]
41. StataCorp. *Stata Statistical Software: Release 14*; Stata Corporation: College Station, TX, USA, 2015.
42. Bornmann, L.; de Moya Anegón, F.; Mutz, R. Do universities or research institutions with a specific subject profile have an advantage or a disadvantage in institutional rankings? A latent class analysis with data from the SCImago ranking. *J. Am. Soc. Inf. Sci. Technol.* **2013**, *64*, 2310–2316. [[CrossRef](#)]
43. Angeles, G.; Cronin, C.; Guilkey, D.K.; Lance, P.M.; Sullivan, B.A. *Guide to Longitudinal Program Impact Evaluation*; Measurement, Learning & Evaluation Project: Chapel Hill, NC, USA, 2014.
44. Bonaccorsi, A. *Knowledge, Diversity and Performance in Europe—An Higher Education: A Changing Landscape*; Edvard Elgar: Cheltenham, UK, 2014.
45. Eumida. *Final Study Report: Feasibility Study for Creating a European University Data Collection*; European Commission, Research Directorate-General C-European Research Area Universities and Researchers: Brussels, Belgium, 2009.
46. Abramo, G.; D’Angelo, C.A.; Pugini, F. The measurement of Italian universities’ research productivity by a non parametric-bibliometric methodology. *Scientometrics* **2008**, *76*, 225–244. [[CrossRef](#)]
47. Warning, S. Performance differences in German higher education: Empirical analysis of strategic groups. *Rev. Ind. Organ.* **2004**, *24*, 393–408. [[CrossRef](#)]
48. Glänzel, W.; Schöpfli, U. A Bibliometric Study on Aging and Reception Processes of Scientific Literature. *J. Inf. Sci.* **1995**, *21*, 37–53. [[CrossRef](#)]
49. Waltman, L.; van Eck, N.J. The need for contextualized scientometric analysis: An opinion paper. In *Proceedings of the 21st International Conference on Science and Technology Indicator*, Valencia, Spain, 14–16 September 2016; pp. 541–549.
50. Glänzel, W.; Thijs, B.; Debackere, K. Productivity, performance, efficiency, impact—What do we measure anyway?: Some comments on the paper “A farewell to the MNCS and like size-independent indicators” by Abramo and D’Angelo. *J. Informetr.* **2016**, *10*, 658–660. [[CrossRef](#)]
51. Aagaard, K.; Schneider, J.W. Research funding and national academic performance: Examination of a Danish success story. *Sci. Public Policy* **2016**, *43*, 518–531. [[CrossRef](#)]

