ifo WORKING PAPERS

264
2018

October 2018

How efficiently produce elite US universities highly cited papers? A case study based on input and output data

Klaus Wohlrabe, Felix de Moya Anegon, Lutz Bornmann

ifo INSTITUTE
Leibniz Institute for Economic Research
at the University of Munich

Impressum:

ifo Working Papers
Publisher and distributor: ifo Institute – Leibniz Institute for Economic Research at the University of Munich
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49(0)89 9224 0, Telefax +49(0)89 985369, email ifo@ifo.de
www.cesifo-group.de

An electronic version of the paper may be downloaded from the ifo website:
www.cesifo-group.de

ifo Working Paper No. 264

How efficiently produce elite US universities highly cited papers? A case study based on input and output data

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 should be related to the input. The present study is based on a comprehensive dataset with input- and output-data for 50 US universities. As input, we used the amount of budget, 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-α approach. The results of the DEA and FDH analysis show that Harvard University and Rice University 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 Rice University is its small input. In the order-α and order-m frameworks, Harvard University remains efficient, but Rice University 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.

JEL Code: I21, I23, D61

Keywords: University; efficiency analysis; regression analysis; normalized citation impact

Klaus Wohlrabe
ifo Institute – Leibniz Institute for Economic Research
at the University of Munich
Poschingerstr. 5,
81679 Munich, Germany
wohlrabe@ifo.de

Felix de Moya Anegon
CSIC, Institute of Public Goods and Policies (IPP)
Consejo Superior de Investigaciones Científicas C/Albasanz, 26–28
Madrid 28037, Spain

Lutz Bornmann
Division for Science and Innovation Studies
Administrative Headquarters of the Max Planck Society
Hofgartenstr. 8,
80539 Munich, Germany

1 Introduction

The science system has been characterised by the transition from academic science to post-academic science for a number of 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’” (Ziman, 2000, 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, Bornmann, 2011), post-academic science is characterised by the use of quantitative methods of research evaluation. According to Wilsdon et al. (2015) 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 (National Research Council, 2014).

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 (Hazelkorn, 2011). 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 (Wilsdon et al., 2015). However, a lot of critiques have been published in recent years (see e.g. Schmoch, 2015) that focus on the methods and arbitrary weightings used to combine different metrics. Daraio, Bonaccorsi, and Simar (2015) cite 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 study, we pick up the last point “lack of consideration of the input-output structure” and set a possible approach of input-output consideration in institutional evaluation to discuss (in scientometrics). Since positions in rankings depend on certain context factors (Bornmann, Stefaner, de Moya Anegón, & Mutz, 2014; Safón, 2013), rankings should not only offer information on the output, but also the relation of input and output. Moed and

Halevi (2015) 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 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 (2016) 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 Approach of this study

This study follows the call of Bornmann and Haunschild (2016) and Waltman, van Eck, Visser, and Wouters (2016) who propose in a comment on the paper by Abramo’s and D’Angelo’s (2016) 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 budget figures. The output concerns the 1% most frequently cited publications in a specific field and given year (Ptop 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 which we used is standard in the literature, our output variable has never been used before (to the best of our knowledge).

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” (Abramo & D’Angelo, 2014, p. 1134). The most serious drawback of the DEA in its simplest form is that it is extremely vulnerable to outliers and measurement errors.[1]

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 (Cazals, Florens, & Simar, 2002) and order-α (Aragon, Daouia, & Thomas-Agnan, 2005) 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 (1982) and Bessent, Bessent, Charnes, Cooper, and Thorogood (1983). Worthington (2001) and more recently Rhaiem (2017) as well as De Witte and López-Torres (2017) provide comprehensive surveys of the literature. In the efficiency analyses of higher education institutions, PFA has rarely been used. Bonaccorsi, Daraio, and Simar (2006) and Bonaccorsi, Daraio, Raty, and Simar (2007) applied the order-m approach to study 45 universities in Italy and 261 universities across four European countries, respectively. De Witte, Rogge, Cherchye, and Van Puyenbroeck (2013) used DEA and PFA to study the performance of 155 professors working at a Business & Administration department of a Brussels university college. Bruffaerts, Rock, and Dehon (2013) 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 (in press) used partial frontier analysis to identify super-efficient economics departments. There are a number of studies available in the literature which have investigated efficiency aspects in the US higher education system (Agasisti & Johnes, 2015; Cohn, Rhine, & Santos, 1989; Harter, Wade, & Watkins, 2005; Laband & Lentz, 2004; Sav, 2012; Titus, Vamosiu, & McClure, 2017).

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 a first step, we calculate efficiency scores for the universities. In a second step, we calculate adjusted efficiency

[1] 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.

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 constant returns to scale (CSR). 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 the outputs. In our estimation framework we cannot test the kind of economies of scale. The partial frontier approaches used in this paper assume CSR. 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-α. 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 one. 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, Cooper, and Rhodes (1979). 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:

$$\min_{e,\lambda} e \quad \text{subject to}$$

$$e \cdot x_{mi} - \sum_{j=1}^{N} \lambda_j x_{mj} \geq 0 \;\; m = 1, \dots, M$$

$$\sum_{j=1}^{N} \lambda_j y_{lj} - y_{li} \geq 0 \quad l = 1, \dots, L$$

$$\lambda_j \geq 0 \ \forall j$$

where λ 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 (2011) and Wilson and Clemson (2013).

We compare each university i with every other university in the data set ($j = 1 \dots N$). The set of peer universities that satisfies the condition $y_{lj} \geq y_{li}\ \forall l$ 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

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

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, Simar, and Tulkens (2006).

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 (2007) proposed the following four-step procedure:

Draw from B_i a random sample of m peer universities with replacement.

A pseudo-FDH efficiency score ($\hat{e}_{mi}^{\widehat{FDH}_d}$) is calculated using the artificial drawn data.

Repeat steps 1 and 2 D times.

Order-m efficiency is calculated as the average of the pseudo-FDH scores

$$\hat{e}_{mi}^{OM} = \frac{1}{D} \sum_{d=1}^{D} \hat{e}_{mi}^{\widehat{FDH}_d}$$

A potential result of this procedure is that the order-m efficiency scores exceed the value of one. 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 $\hat{\mathrm{e}}_{\mathrm{mi}}^{\mathrm{OM}} > 1$) 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 $\mathrm{m} \rightarrow \infty$ the approach converges to the FDH results.

3.1.3 Order-α efficiency

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

$$\hat{\mathrm{e}}_{\alpha\mathrm{i}}^{\mathrm{OA}} = \mathrm{P}_{\substack{(100-\alpha)\\ \mathrm{j}\in\mathrm{B_i}}}\left\{\max_{\mathrm{k}=1,\dots,\mathrm{K}}\left(\frac{\mathrm{x_{kj}}}{\mathrm{x_{ki}}}\right)\right\}$$

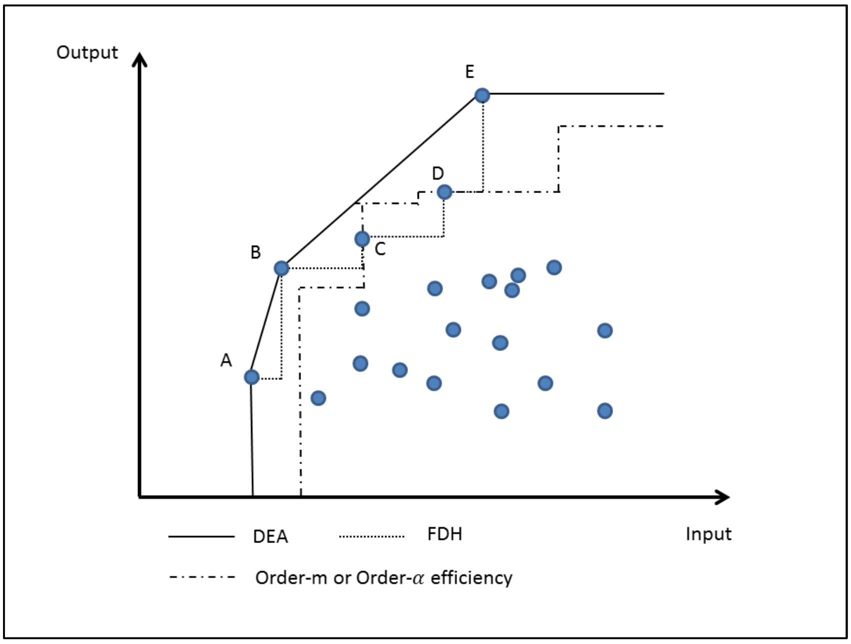
When $\alpha = 100$, the approach replicates the FDH results. In case of $\alpha < 100$, some universities may be classified as super-efficient. As m is the approach explained in section 3.1.2, α can be considered as a tuning parameter: the smaller α is, the larger is the share of the super-efficient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

3.1.4 A simple example for explaining the approaches

Figure 1 illustrates the outlined full and PFA approaches. We plotted input-output combinations for various artificial 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.

Applying the partial frontier approaches – order-m or order-α – 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.

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



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 Agasisti & Wolszczak-Derlacz, 2015). We used Stata (StataCorp., 2015) 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, de Moya Anegón, and Mutz (2013) 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 (http://www.scimagoir.com). 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 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 (Angeles, Cronin, Guilkey, Lance, & Sullivan, 2014).

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 the amount of budget. The data source is the National Center for Education Statistics (NCES).1 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 (Bonaccorsi, 2014; Eumida, 2009). The data refer to the academic year, which starts on June, 1st and ends on May, 31st. 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 7/12 of the data from the academic year 2013/14 and 5/12 from 2012/2013. This approach might introduce some unknown biases as we assume that the budget is spent evenly across the year. So, we cannot assure that the budget represents 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.

[1] The data can be downloaded from http://nces.ed.gov/ipeds/datacenter/InstitutionProfile.aspx?unitid=adafaeb2afaf.

Table 1. Descriptive statistics over time

	2011		2012		2013	
	Budget	$P_{top\ 1\%}$	Budget	$P_{top\ 1\%}$	Budget	$P_{top\ 1\%}$
Mean	2,511,431,540	254	2,625,579,734	277	2,744,293,230	225
Median	2,338,410,542	213	2,362,932,171	228	2,485,447,347	197
Standard Deviation	1,354,727,092	160	1,427,057,141	182	1,505,658,689	151
Minimum	538,295,750	35	560,340,667	19	584,337,500	24
Maximum	5,736,790,750	1002	5,883,354,333	1198	6,160,311,750	977

Notes. Descriptive statistics for the input and output are reported.

As we focus on the best US universities, we use as output the number of papers that belong to the 1% (Ptop 1%) most frequently cited papers in the corresponding fields and publication years (Bornmann, de Moya Anegon, & Leydesdorff, 2012). The typical output variables in efficiency analysis are students, graduates, and funding; publications are used rather seldom (Abramo, D'Angelo, & Pugini, 2008; Warning, 2004). To the best of our knowledge this is the first study using top-cited publications as output variable. The data were obtained from the SCImago Institutions Ranking, which is based on Scopus data (see http://www.scimagoir.com). The output data refer to the publication period from 2011 to 2013 with a citation window from publication until the end of 2015. In section 4, we focus on the results for 2013. Both other publication years allowed us to take a look on 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 the budget figures increase whereas the average Ptop 1% peaked in 2012 and dropped considerably in 2013. Table 2 reports the correlation coefficients between the budget and Ptop 1% over time. All coefficients are about 0.6 implying a moderate positive relationship.

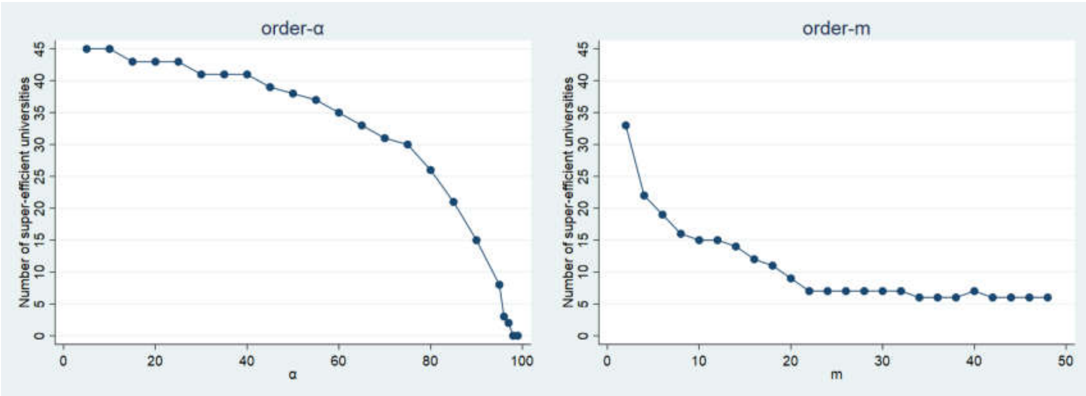
Table 2. Correlations between budget and Ptop 1% over time

	2011	2012	2013
Coefficient	0.587	0.572	0.610

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.

Figure 2. Number of super-efficient universities for different values of m and α



In contrast, PFA requires the specification of parameters, which eventually influence the amount of super-efficient universities. The order-α approach requires α, the percentile of the set of peer universities used as benchmark. Order-m requires m, the number of peer universities randomly drawn from the initial set of universities. Unless we set m = 50 or α = 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 α. We used data for the year 2013. It is clearly visible that with higher m or α values, respectively, the number of super-efficient universities becomes lower. Concerning m, the figure is quite stable beyond 25. We opted to set m = 40 and α = 95%, which yield on 11 and 8 super-efficient universities, respectively.

4.1 Baseline results

4.1.1 Results for 2013

Table 3 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 36 universities using the FDH approach. This is because both Harvard and Rice University dominate the estimated academic efficiency frontier: Harvard University due to its very high output values and Rice University due to its small input values relative to the outputs. With respect to the order-α framework, there are 8 universities with a score larger than one. This number is increased to 11 based on the order-m approach. In the majority of cases, the order-α approach yields higher scores compared to the order-m approach.

In the order-α framework and the order-m approach, Harvard University remains efficient with a score of 1.00 and is not denoted as super-efficient. However, Rice University is super-efficient. The highest efficiency scores in the order-α approach obtain the Case Western Reserve University and the University of California; Santa Barbara has the highest score in the order-m model.

Based on a Stochastic Frontier analysis, Agasisti and Johnes (2015) also report 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 and Rice University among the 20 best universities.

Table 3. Efficiency scores and the corresponding rankings based on different approaches for measuring efficiency in 2013 (sorted by THE Ranking)

		DEA		FDH		Order-α		Order-m	
THE	University	Score	Rank	Score	Rank	Score	Rank	Score	Rank
1	California Institute of Technology	0.505	21	0.844	17	0.859	20	0.851	18
2	Harvard University	1.000	1	1.000	1	1.000	9	1.000	12
3	Stanford University	0.554	20	1.000	1	1.000	9	1.000	12
4	Massachusetts Institute of Technology	0.640	13	1.000	1	1.000	9	1.000	12
5	Princeton University	0.616	15	0.938	15	1.000	9	0.972	15
6	University of California, Berkeley	0.659	12	1.000	1	1.000	9	1.003	11
7	Yale University	0.437	27	0.599	32	0.599	35	0.603	34
8	University of Chicago	0.321	38	0.545	35	0.554	38	0.549	35
9	University of California, Los Angeles	0.320	39	0.510	37	0.510	42	0.510	40
10	Columbia University	0.452	24	0.822	18	0.822	22	0.822	19
11	Johns Hopkins University	0.387	32	0.605	30	0.605	34	0.605	33
12	University of Pennsylvania	0.280	41	0.473	44	0.473	45	0.473	45
13	University of Michigan, Ann Arbor	0.280	42	0.478	43	0.478	44	0.479	44
14	Duke University	0.302	40	0.517	36	0.517	40	0.519	37
15	Cornell University	0.724	8	1.000	1	1.000	9	1.004	10
16	North-western University, Evanston	0.697	10	1.000	1	1.000	9	1.011	8
17	Carnegie Mellon University	0.587	17	0.722	24	0.869	19	0.782	21
18	University of Washington	0.418	29	0.668	28	0.668	30	0.669	29
19	Georgia Institute of Technology	0.623	14	1.000	1	1.319	3	1.092	4
20	University of Texas, Austin	0.393	31	0.703	25	0.716	28	0.709	26
21	University of Illinois at Urbana-Champaign	0.353	35	0.483	41	0.555	37	0.512	38
22	University of Wisconsin, Madison	0.485	22	0.729	23	0.729	27	0.736	24
23	University of California, Santa Barbara	0.873	4	1.000	1	1.292	5	1.126	1
24	New York University	0.261	46	0.418	47	0.418	48	0.422	47
25	University of California, San Diego	0.429	28	0.666	29	0.666	31	0.670	28
26	Washington University in Saint Louis	0.484	23	0.773	20	0.773	24	0.781	22
27	University of Minnesota, Twin Cities	0.381	34	0.605	31	0.615	33	0.609	32
28	University of North Carolina, Chapel Hill	0.439	26	0.671	27	0.671	29	0.676	27
29	Brown University	0.812	6	1.000	1	1.205	6	1.099	3
30	University of California, Davis	0.274	43	0.480	42	0.489	43	0.484	43
31	Boston University	0.566	19	0.821	19	1.000	9	0.888	17
32	Pennsylvania State University	0.196	50	0.397	48	0.404	49	0.402	48
33	Ohio State University, Columbus	0.216	49	0.361	49	0.368	50	0.363	49
34	Rice University	1.000	1	1.000	1	1.294	4	1.082	5
35	University of Southern California	0.270	44	0.507	38	0.516	41	0.512	39
36	Michigan State University	0.331	37	0.436	45	0.646	32	0.494	41
37	University of Arizona	0.449	25	0.680	26	0.780	23	0.725	25
38	University of Notre Dame	0.589	16	0.592	33	0.765	25	0.645	30
39	Tufts University	0.751	7	0.750	21	0.970	18	0.816	20
40	University of California, Irvine	0.338	36	0.499	39	0.573	36	0.531	36
41	University of Pittsburgh	0.583	18	1.000	1	1.018	8	1.005	9
42	Emory University	0.219	47	0.435	46	0.443	47	0.439	46
43	Vanderbilt University	0.262	45	0.487	40	0.522	39	0.494	42
44	University of Colorado, Boulder	0.687	11	1.000	1	1.148	7	1.063	7
45	Purdue University	0.414	30	0.555	34	0.823	21	0.637	31
46	University of California, Santa Cruz	0.925	3	1.000	1	1.348	2	1.123	2
47	Case Western Reserve University	0.722	9	1.000	1	1.362	1	1.082	6
48	University of Rochester	0.218	48	0.302	50	0.447	46	0.344	50
49	Boston College	0.840	5	0.848	16	1.000	9	0.911	16
50	University of Florida	0.382	33	0.733	22	0.746	26	0.740	23

Table 4 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 and the scores 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.

Table 4. Spearman rank correlation coefficients for 2013

	THE	DEA	FDH	order-α	order-m
THE	1.000				
DEA	0.031	1.000			
FDH	0.169	0.926	1.000		
order-α	-0.014	0.959	0.950	1.000	
order-m	0.079	0.947	0.984	0.978	1.000

4.1.2 Stability of the results over time

Table 5 reports the rank correlations across time (2011, 2012, and 2013) for each approach of the efficiency analysis. They are all above 0.9 suggesting that the results are quite stable over the observed time period.

Table 5. 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	0.98	1.00		2012	0.92	1.00	
2013	0.98	0.99	1.00	2013	0.93	0.89	1.00
	order-α				order-m		
2011	1.00			2011	1.00		
2012	0.95	1.00		2012	0.92	1.00	
2013	0.95	0.93	1.00	2013	0.93	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 6. As dependent variables, the efficiency scores from Table 3 are used (results from the DEA, FDH, order-α approach, and order-m approaches). We performed 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 – with the largest coefficients – is related to lower efficiency scores. Furthermore, the results in Table 6 demonstrate that private universities are more efficient than public universities. Most of the coefficients in the models are statistically not significant (which might be the result of the low numbers of universities in the study).

Table 6. Beta coefficients and t statistics of the regression models with the efficiency scores as dependent variable for 2013

	DEA	FDH	order-α	order-m
Life sciences	-0.34	-0.86*	-0.32	-0.73
	(-1.17)	(-2.40)	(-0.69)	(-1.89)
Physical sciences	-0.99	-2.51*	-0.80	-2.09
	(-1.17)	(-2.39)	(-0.58)	(-1.84)
Social sciences	-0.29	-0.76*	-0.27	-0.63
	(-1.27)	(-2.62)	(-0.71)	(-2.02)
Health sciences	-0.78	-1.84*	-0.66	-1.55
	(-1.31)	(-2.47)	(-0.66)	(-1.92)
Private state	0.11	0.05	0.09	0.06
	(1.38)	(0.69)	(0.91)	(0.74)
Constant	0.45***	0.69***	0.74***	0.71***
	(11.66)	(17.19)	(14.46)	(16.48)
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 7 besides the initial ranking positions. Although both ranking positions are highly correlated (DEA: rs=0.87, FDH: rs=.88, order-α: rs=.85, order-m: rs=.89), there are significant rank changes for some universities. For example, Rice University shows a perfect rank position in the DEA; but if the score is adjusted by the independent variables in the regression model, its score decreases, leading to the fifth ranking position.

Table 7. Initial efficiency rank positions and adjusted rank positions in 2013 (sorted by adjusted DEA scores)

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

4.2.2 Stability of the results over time

Table 8 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 0.9, which demonstrate that the results are stable over the publication years considered.

Table 8. 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	0.96	1.00		2012	0.81	1.00	
2013	0.96	0.96	1.000	2013	0.89	0.88	1.00
	order-α				order-m		
2011	1.00			2011	1.00		
2012	0.89	1.00		2012	0.84	1.00	
2013	0.91	0.92	1.00	2013	0.91	0.92	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 (Bornmann, 2011). Whereas output and impact assessments were initially at the forefront of assessments, efficiency measurements have become popular in recent years (Rhaiem, 2017). According to Moed and Halevi (2015), 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 (Rhaiem, 2017; Waltman & van Eck, 2016). 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 number of staff and amount of budget and as output the number of (highly-cited) papers. The results of the DEA and FDH analysis show that Harvard

University and Rice University 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 Rice University is its small input. In the order-α and order-m frameworks, Harvard University remains efficient, but Rice University becomes super-efficient. Although Harvard University and Rice University are well-known as the best universities in the world, the correlations between the ranking positions of the universities in the THE Ranking 2015 and the results of the 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, Rice 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 (Rhaiem, 2017), 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. (2014). 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. (2014) 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 taken into account in future studies.

The first limitation is related to the numbers of indicators used. We included only two input- and output-indicators, 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. The efficiency study of Bruffaerts et al. (2013), 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 (2017) 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 (Waltman et al., 2016). "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" (Glänzel, Thijs, & Debackere, 2016, p. 659). 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. (2016) 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 (2016) highlight many difficulties in explaining research performance (output and impact) as a linear function of input indicators. Bornmann and Haunschild (2016) see efficiency in research as diametric to creativity and faulty incrementalism, which are basic elements of each (successful) research process. According to Ziman (2000) "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).

References

Aagaard, K., & Schneider, J. W. (2016). Research funding and national academic performance: Examination of a Danish success story. *Science and Public Policy, 43*(4), 518-531. doi: 10.1093/scipol/scv058.

Abramo, G., D'Angelo, C. A., & Pugini, F. (2008). The measurement of Italian universities' research productivity by a non parametric-bibliometric methodology. *Scientometrics, 76*(2), 225-244. doi: 10.1007/s11192-007-1942-2.

Abramo, G., & D'Angelo, C. (2014). How do you define and measure research productivity? *Scientometrics, 101*(2), 1129-1144. doi: 10.1007/s11192-014-1269-8.

Abramo, G., & D'Angelo, C. A. (2016). A farewell to the MNCS and like size-independent indicators. *Journal of Informetrics, 10*(2), 646-651.

Agasisti, T., & Johnes, G. (2015). Efficiency, costs, rankings and heterogeneity: the case of US higher education. *Studies in Higher Education, 40*(1), 60-82. doi: 10.1080/03075079.2013.818644.

Agasisti, T., & Wolszczak-Derlacz, J. (2015). Exploring efficiency differentials between Italian and Polish universities, 2001–11. *Science and Public Policy*. doi: 10.1093/scipol/scv026.

Angeles, G., Cronin, C., Guilkey, D. K., Lance, P. M., & Sullivan, B. A. (2014). *Guide to Longitudinal Program Impact Evaluation*. Chapel Hill, NC, USA: Measurement, Learning & Evaluation Project.

Aragon, Y., Daouia, A., & Thomas-Agnan, C. (2005). Nonparametric frontier estimation: A conditional quantile-based approach. *Econometric Theory, 21*(2), 358-389. doi: 10.1017/S0266466605050206.

Bessent, A. M., Bessent, E. W., Charnes, A., Cooper, W. W., & Thorogood, N. C. (1983). Evaluation of Educational-Program Proposals by Means of Dea. *Educational Administration Quarterly, 19*(2), 82-107. doi: Doi 10.1177/0013161x83019002006.

Bogetoft, P., & Otto, L. (2011). *Benchmarking with DEA, SFA and R*. New York, NY, USA: Springer.

Bonaccorsi, A. (2014). *Knowledge, Diversity and Performance in Europe-an Higher Education: A Changing Landscape*. Cheltenham, UK: Edvard Elgar.

Bonaccorsi, A., Daraio, C., Raty, T., & Simar, L. (2007). *Efficiency and University Size: Discipline-wise Evidence from European Universities, MPRA Paper 10265*. Munich, Germany: University Library of Munich.

Bonaccorsi, A., Daraio, C., & Simar, L. (2006). Advanced indicators of productivity of universities. An application of robust nonparametric methods to Italian data. *Scientometrics, 66*(2), 389-410. doi: 10.1007/s11192-006-0028-x.

Bornmann, L. (2011). Scientific peer review. *Annual Review of Information Science and Technology, 45*, 199-245.

Bornmann, L., de Moya Anegon, F., & Leydesdorff, L. (2012). The new Excellence Indicator in the World Report of the SCImago Institutions Rankings 2011. *Journal of Informetrics, 6*(2), 333-335. doi: 10.1016/j.joi.2011.11.006.

Bornmann, L., de Moya Anegón, F., & Mutz, R. (2013). 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. *Journal of the American Society for Information Science and Technology, 64*(11), 2310-2316.

Bornmann, L., & Haunschild, R. (2016). Efficiency of research performance and the glass researcher. *Journal of Informetrics, 10*(2), 652-654.

Bornmann, L., Stefaner, M., de Moya Anegón, F., & Mutz, R. (2014). 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. *Journal of Informetrics, 8*(3), 581-593. doi: 10.1016/j.joi.2014.04.008.

Bruffaerts, C., Rock, B. D., & Dehon, C. (2013). *The Research Efficiency of US Universities: a Nonparametric Frontier Modelling Approach, Working Papers ECARES ECARES 2013-31*. Bruxelles, Belgium: Universite Libre de Bruxelles.

Cazals, C., Florens, J. P., & Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics, 106*(1), 1-25. doi: Doi 10.1016/S0304-4076(01)00080-X.

Charnes, A., Cooper, W. W., & Rhodes, E. (1979). Measuring the Efficiency of Decision-Making Units. *European Journal of Operational Research, 3*(4), 339-339. doi: Doi 10.1016/0377-2217(79)90229-7.

Cohn, E., Rhine, S. L. W., & Santos, M. C. (1989). Institutions of Higher-Education as Multi-Product Firms - Economies of Scale and Scope. *Review of Economics and Statistics, 71*(2), 284-290. doi: Doi 10.2307/1926974.

Daraio, C., Bonaccorsi, A., & Simar, L. (2015). Rankings and university performance: A conditional multidimensional approach. *European Journal of Operational Research, 244*(3), 918-930. doi: 10.1016/j.ejor.2015.02.005.

Daraio, C., & Simar, L. (2007). *Advanced Robust and nonparametric methods in efficiency analysis: methodology and applications*. Heidelberg, Germany: Springer.

De Witte, K., & López-Torres, L. (2017). Efficiency in education: a review of literature and a way forward. *Journal of the Operational Research Society, 68*(4), 339-363. doi: 10.1057/jors.2015.92.

De Witte, K., Rogge, N., Cherchye, L., & Van Puyenbroeck, T. (2013). Accounting for economies of scope in performance evaluations of university professors. *Journal of the Operational Research Society, 64*(11), 1595-1606. doi: 10.1057/jors.2012.115.

Deprins, D., Simar, L., & Tulkens, H. (2006). Measuring Labor-Efficiency in Post Offices. In P. Chander, J. Drèze, C. K. Lovell & J. Mintz (Eds.), *Public goods, environmental externalities and fiscal competition* (pp. 285-309). Boston, MA: Springer US.

Eumida. (2009). *Final Study Report: Feasibility Study for Creating a European University Data Collection*. Brussels, Belgium: European Commission, Research Directorate-General C-European Research Area Universities and Researchers.

Glänzel, W., Thijs, B., & Debackere, K. (2016). 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. *Journal of Informetrics, 10*(2), 658-660. doi: http://dx.doi.org/10.1016/j.joi.2016.04.008.

Gnewuch, M., & Wohlrabe, K. (in press). Super-efficiency of education institutions: an application to economics departments. *Education Economics*.

Harter, J. F. R., Wade, J. A., & Watkins, T. G. (2005). An examination of costs at four-year public colleges and universities between 1989 and 1998. *Review of Higher Education, 28*(3), 369-+. doi: DOI 10.1353/rhe.2005.0011.

Hazelkorn, E. (2011). *Rankings and the reshaping of higher education. The battle for world-class excellence*. New York, NY, USA: Palgrave Macmillan.

Laband, D. N., & Lentz, B. F. (2004). Do costs differ between for-profit and not-for-profit producers of higher education? *Research in Higher Education, 45*(4), 429-441. doi: Doi 10.1023/B:Rihe.0000027394.33115.71.

Lindsay, A. W. (1982). Institutional Performance in Higher-Education - the Efficiency Dimension. *Review of Educational Research, 52*(2), 175-199.

Moed, H. F., & Halevi, G. (2015). Multidimensional assessment of scholarly research impact. *Journal of the Association for Information Science and Technology, 66*(10), 1988-2002. doi: 10.1002/asi.23314.

National Research Council. (2014). *Furthering America's Research Enterprise*. Washington, DC: The National Academies Press.

Rhaiem, M. (2017). Measurement and determinants of academic research efficiency: a systematic review of the evidence. *Scientometrics, 110*(2), 581–615. doi: 10.1007/s11192-016-2173-1.

Safón, V. (2013). What do global university rankings really measure? The search for the X factor and the X entity. *Scientometrics, 97*(2), 223-244. doi: 10.1007/s11192-013-0986-8.

Sav, G. T. (2012). Stochastic Cost Inefficiency Estimates and Rankings of Public and Private Research and Doctoral Granting Universities. *Journal of Knowledge Management, Economics and Information Technology, 4*(3), 11-29.

Schmoch, U. (2015). The Informative Value of International University Rankings: Some Methodological Remarks. In I. M. Welpe, J. Wollersheim, S. Ringelhan & M. Osterloh (Eds.), *Incentives and Performance* (pp. 141-154): Springer International Publishing.

StataCorp. (2015). *Stata statistical software: release 14*. College Station, TX, USA: Stata Corporation.

Titus, M. A., Vamosiu, A., & McClure, K. R. (2017). Are Public Master's Institutions Cost Efficient? A Stochastic Frontier and Spatial Analysis. *Research in Higher Education, 58*(5), 469-496. doi: 10.1007/s11162-016-9434-y.

Waltman, L., & van Eck, N. J. (2016). The need for contextualized scientometric analysis: An opinion paper. In I. Ràfols, J. Molas-Gallart, E. Castro-Martínez & R. Woolley (Eds.), *Proceedings of the 21 ST International Conference on Science and Technology Indicator* (pp. 541-549). València, Spain: Universitat Politècnica de València.

Waltman, L., van Eck, N. J., Visser, M., & Wouters, P. (2016). The elephant in the room: The problem of quantifying productivity in evaluative scientometrics. *Journal of Informetrics, 10*(2), 671-674. doi: http://dx.doi.org/10.1016/j.joi.2015.12.008.

Warning, S. (2004). Performance differences in German higher education: Empirical analysis of strategic groups. *Review of Industrial Organization, 24*(4), 393-408. doi: DOI 10.1023/B:REIO.0000037538.48594.2c.

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., . . . Johnson, B. (2015). *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. Bristol, UK: Higher Education Funding Council for England (HEFCE).

Wilson, P. W., & Clemson, S. C. (2013). *FEAR 2.0: A Software package for Frontier Analysis with R*. Clemson, SC, USA: Department of Economics, Clemson University.

Worthington, A. C. (2001). An Empirical Survey of Frontier Efficiency Measurement Techniques in Education. *Education Economics, 9*(3), 245-268. doi: 10.1080/09645290110086126.

Ziman, J. (2000). *Real science. What it is, and what it means*. Cambridge, UK: Cambridge University Press.

ifo Working Papers

No. 263 Schwefer, M., Sitting on a Volcano: Domestic Violence in Indonesia Following Two Volcano Eruptions, September 2018.

No. 262 Vandrei, L., Does Regulation Discourage Investors? Sales Price Effects of Rent Controls in Germany, June 2018.

No. 261 Sandkamp, A.-N., The Trade Effects of Antidumping Duties: Evidence from the 2004 EU Enlargement, June 2018.

No. 260 Corrado, L. and T. Schuler, Financial Bubbles in Interbank Lending, April 2018.

No. 259 Löffler, M., A. Peichl and S. Siegloch The Sensitivity of Structural Labor Supply Estimations to Modeling Assumptions, March 2018.

No. 258 Fritzsche, C. and L. Vandrei, Causes of Vacancies in the Housing Market – A Literature Review, March 2018.

No. 257 Potrafke, N. and F. Rösel, Opening Hours of Polling Stations and Voter Turnout: Evidence from a Natural Experiment, February 2018.

No. 256 Hener, T. and T. Wilson, Marital Age Gaps and Educational Homogamy – Evidence from a Compulsory Schooling Reform in the UK, February 2018.

No. 255 Hayo, B. and F. Neumeier, Households' Inflation Perceptions and Expectations: Survey Evidence from New Zealand, February 2018.

No. 254 Kauder, B., N. Potrafke and H. Ursprung, Behavioral determinants of proclaimed support for environment protection policies, February 2018.

No. 253 Wohlrabe, K., L. Bornmann, S. Gralka und F. de Moya Anegon, Wie effizient forschen Universitäten in Deutschland, deren Zukunftskonzepte im Rahmen der Exzellenzinitiative ausgezeichnet wurden? Ein empirischer Vergleich von Input- und Output-Daten, Februar 2018.

No. 252 Brunori, P., P. Hufe and D.G. Mahler, The Roots of Inequality: Estimating Inequality of Opportunity from Regression Trees, January 2018.

No. 251 Barrios, S., M. Dolls, A. Maftei, A. Peichl, S. Riscado, J. Varga and C. Wittneben, Dynamic scoring of tax reforms in the European Union, January 2018.

No. 250 Felbermayr, G., J. Gröschl and I. Heiland, Undoing Europe in a New Quantitative Trade Model, January 2018.

No. 249 Fritzsche, C., Analyzing the Efficiency of County Road Provision – Evidence from Eastern German Counties, January 2018.

No. 248 Fuest, C. and S. Sultan, How will Brexit affect Tax Competition and Tax Harmonization? The Role of Discriminatory Taxation, January 2018.

No. 247 Dorn, F., C. Fuest and N. Potrafke, Globalization and Income Inequality Revisited, January 2018.

No. 246 Dorn, F. and C. Schinke, Top Income Shares in OECD Countries: The Role of Government Ideology and Globalization, January 2018.

No. 245 Burmann, M., M. Drometer and R. Méango, The Political Economy of European Asylum Policies, December 2017.

No. 244 Edo, A., Y. Giesing, J. Öztunc and P. Poutvaara, Immigration and Electoral Support for the Far Left and the Far Right, December 2017.

No. 243 Enzi, B., The Effect of Pre-Service Cognitive and Pedagogical Teacher Skills on Student Achievement Gains: Evidence from German Entry Screening Exams, December 2017.

No. 242 Doerrenberg, P. and A. Peichl, Tax morale and the role of social norms and reciprocity. Evidence from a randomized survey experiment, November 2017.

No. 241 Fuest, C., A. Peichl and S. Siegloch, Do Higher Corporate Taxes Reduce Wages? Micro Evidence from Germany, September 2017.