Research Policy 46 (2017) 1595–1605

Contents lists available at ScienceDirect

Research Policy

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

Research paper

An evaluation and explanation of (in)efficiency in higher education institutions in Europe and the U.S. with the application of two-stage semi-parametric DEA

Joanna Wolszczak-Derlacz

Gdańsk University of Technology, Faculty of Management and Economics, Narutowicza 11/12, 80-233 Gdańsk, Poland

ARTICLE INFO

JEL classifications:
I23
C14
I22

Keywords:
Higher education institutions
Efficiency
Two-stage DEA
European-US comparison

ABSTRACT

In this study the technical efficiency of number of public European and American HEIs is assessed over a decade. Efficiency scores are determined using nonparametric DEA with different input-output sets and considering different frontiers: global frontier (all HEIs pooled together), regional frontier (Europe and the U.S. having their own frontiers) and country-specific ones. The external factors affecting the degree of HEI inefficiency are also investigated, e.g. institutional settings (size and department composition), location and funding structure. Specifically, the results indicate a positive association between both regional GDP per capita and number of departments and an institution's efficiency (for both the European and U.S. samples). On average, older European HEIs are more efficient, but this is not confirmed for American ones. Finally, government funding seems to have a negative effect on the efficiency of universities in Europe, which again is not confirmed for the U.S. However, some country heterogeneity at the European level is found through intensive sensitivity analysis.

1. Introduction

Numbers are meaningful: according to the Academic Ranking of World Universities[1] 2016 fifteen of the top twenty universities were in the U.S., Americans published 23% of the total number of scientific articles in the period 1996–2015, counting 33% of the total citations.[2] This is perceived in the literature as the transatlantic gap – referring to the differences between Europe and the U.S. in the quality of academic research (Bonaccorsi et al., 2017). Because of this, it is not surprising that the American system of higher education is perceived to be pre-eminent and when higher education institutions (hereafter, HEIs) around the world are searching to improve their performance they look to universities in the U.S. as their benchmark model, while scholars from the whole world are attracted to American academia (Clotfelter, 2010). However, from the internal American perspective, the higher education sector is not free of problems, and its worldwide dominance has also recently been challenged (Altbach et al., 2011). Nowadays, HEIs in both continents are under pressure due to declining public support, resulting in the need to seek external resources and to provide first-class teaching and research in order to survive amid local and global competition.[3]

This study has three main aims: firstly, to compare the technical efficiency of European and U.S. higher education institutions. Secondly, to evaluate the main factors that determine the efficiency of HEIs and to test whether these factors might have varying impacts on the European and U.S. efficiency. Thirdly, to address an evaluation problem, introducing DEA techniques as an analytic tool which can serve both HEI's managers and policymakers.

Data envelopment analysis (DEA) is used in this study – a methodology which constructs a production frontier in the multi-input/multi-output case – in order to evaluate the relative efficiency of a sample of 500 higher education institutions (in ten European countries and the U.S.) for the period between 2000 and 2012. Different models are estimated for different input-output sets and assumed frontier: global, regional and country-specific ones.

The research is motivated by the fact that most previous studies

E-mail address: jwo@zie.pg.gda.pl.

[1] http://www.shanghairanking.com/ARWU2016.html. It should be underlined that university rankings (among others, ARWU) are a different concept to efficiency analysis based on purely scientific methodology such as DEA or other nonparametric methods as used in our paper. Daraio et al. (2015b) discuss the main criticisms addressed to university rankings more thoroughly (e.g. monodimensionality, lack of statistical robustness etc.) and propose a new generation of rankings based on new ranking techniques. However, despite their methodological shortcomings global rankings are of great importance to university prestige as they receive a great deal of attention in media.

[2] http://www.scimagojr.com/countryrank.php?min=0&min_type=it.

[3] This can be also analysed from the cross-sectoral perspective of increasing competition for public resources between higher education and other public sector services (e.g. healthcare and public pensions, see Kwiek, 2015).

http://dx.doi.org/10.1016/j.respol.2017.07.010
Received 8 August 2016; Received in revised form 14 July 2017; Accepted 26 July 2017
Available online 14 August 2017
0048-7333/

have only considered one or a limited number of countries, mainly due to the fact that micro data on HEIs (at the level of individual institutions) are not easily obtainable and comparable across countries and time periods. Few studies have looked at the efficiency and productivity of HEIs from the international perspective. In particular, the efficiency of Italian universities has been compared to that of those in the U.K. (Agasisti and Johnes, 2009), Spain (Agasisti and Pérez-Esparrells, 2010), Germany (Agasisti and Pohl, 2012) and Poland (Agasisti and Wolszczak-Derlacz, 2016). However, as these authors admit, general conclusions cannot be drawn on the basis of comparisons between the performances of HEIs in only two countries. Some recent papers utilise European Tertiary Education Register (ETER)[4] database and its ancestors, the Aquameth and Eumida. Bonaccorsi et al. (2007a) cover universities in Italy, Spain, Portugal, Norway, Switzerland and the UK. Bonaccorsi et al. (2007b) compare universities by research field in four European countries. Still, they concentrate mainly on testing economies of scale and scope. Similarly, Daraio et al. (2015a) conduct the analysis of 400 HEIs from 16 European countries but only for the single year 2008/2009 and Daraio et al. (2015b) using the same data underline the aspect of country's differentiation affecting university efficiency. Finally, Bolli et al. (2016) examine the role of competitive funding on both the production frontier and university efficiency.

However, unlike the present paper, none of these studies compare the efficiency of European HEIs with their U.S. counterparts or examine differences in performance measured over a decade taking into account cross-country and cross-unit heterogeneity.[5]

In the present paper following the bootstrap procedures proposed by Simar and Wilson (2000, 2007) we calculate bias-corrected DEA scores and in a second stage the relationship between a given external variable and previously estimated efficiency scores is verified. The results of this quantitative exercise are tested in the numbers of robustness checks.

The results indicate that European and U.S. institutions are relatively inefficient, with a high heterogeneity of efficiency scores both between and within countries. The inefficiency is lower for U.S. institutions compared to the mean value for the whole Europe, although higher in relation to some specific examples of European countries (e.g. the U.K.) what is confirmed in the model with country-specific frontier. The main findings of the second-stage analysis are: (a) universities located in wealthier regions of Europe and the U.S. are more efficient; (b) the number of different departments is positively associated with efficiency – indicating the presence of economies of scope and/or economies of scale; (c) funding structure matters for technical efficiency but the direction of the effect varies between the European and U.S. sample; (d) a greater inefficiency of universities with a larger proportion of revenue obtained from government resources is confirmed only in the case of the European sample with some cross-country heterogeneity.

The remainder of this paper is structured as follows: in Section 2 the methodological basis for the non-parametric analysis of technical efficiency is briefly presented together with literature review of empirical studies in which DEA has been applied to evaluating the efficiency of HEIs in cross-country studies. Next, in Section 3, we describe the panel and data, along with key descriptive statistics on the HEIs in the sample. In Section 4, different versions of DEA models are evaluated for different input-output sets and assumed frontiers. In Section 5, the second-step analysis is conducted, in which we treat the (previously estimated) efficiency scores as dependent variable in a regression equation. Finally, Section 6 is dedicated to the discussion of the findings from a policy perspective and conclusion.

We argue that DEA techniques (with full knowledge of the methodology utilized e.g. its limitations) can be used as an additional tool to help strategic planning and/or evaluation of HEIs. The results of the second step of our analysis where we look for the determinants of the HEI's inefficiency can be informative both to management and policymakers. Specifically, it is shown that funding mechanisms (e.g. through pressure on the competitive resources) have the potential to significantly alter the nature and efficiency of higher education providers.

2. Using two-stage DEA to evaluate technical efficiency and its determinants – method and literature review

In the empirical part of this study the technical efficiency of HEIs will be evaluated through non-parametric DEA analysis, and then by regressing efficiency scores on potential covariates. There is much support for DEA methodology for the empirical evaluation of the production of multi-input/multi-output units, which is in fact a characteristic of the activities carried out by HEIs (Bougnol and Dula, 2006). The formal presentation of the method following closely the notation of Simar and Wilson (2000, 2007) is presented in the Appendix A in Supplementary material. First, we calculate DEA efficiency scores ($\hat{\lambda}$) by maximizing achievable output for a given level of the inputs. If the DEA efficiency score is , then the DMU is said to be efficient, if $\hat{\lambda} > 1$ (or 100%) then the unit is inefficient and the magnitude of the inefficiency is determined by the distance to the benchmark units called frontier (the greater the difference between the DEA score and 1, the greater the inefficiency).

The second step of our analysis involves examination of (the direction and magnitude of) the potential determinants (Z) of the previously estimated bias-corrected efficiency scores $\left(\hat{\hat{\lambda}}_i\right)$.

$$\hat{\hat{\lambda}}_i = \alpha + Z_i\beta + \varepsilon_i, \tag{1}$$

where ε_i is a statistical noise with distribution restricted by $\varepsilon_i \geq 1 - \alpha - z_i\beta$. The bootstrap procedure is employed to obtain bias-corrected beta coefficients to overcome the problems arising from the serial correlation of previously estimated scores and a possible correlation of the error term (ε_i) with environmental variables (Z_i) – see Appendix A in Supplementary material.

Since the 80 s the DEA method has been applied to assess the efficiency of entities operating in various sectors of the economy. In this steam of the literature, examination of the higher education sector is also present, albeit with a quantitatively lower representation.[6] Due to the nature of the present empirical analysis, the following literature review is restricted to works considering the evaluation of the efficiency of HEIs in more than one country (Table B1 in the Appendix B in Supplementary material).

In particular, Agasisti and Johnes (2009) examine universities in Italy and the UK between the years 2002/2003 and 2004/2005, finding that UK universities were more efficient, but the Italian ones were improving their technical efficiency. Italian universities have also been compared to Spanish universities (Agasisti and Pérez-Esparrells, 2010) and to German ones (Agasisti and Pohl, 2012) In the latter publications, the authors conduct also a second-stage analysis employing tobit regression and find evidence that medical faculties and operating in regions with a higher unemployment rate were negatively associated with efficiency and the regional share of employees working in science and technology was positively related.

[4] Aquameth and Eumida were projects funded by the European Commission with intention to create the foundations of a regular data collection on individual HEIs in the EU-27 Member States. As far as the author is aware, these datasets were not freely available to researchers outside the consortium (for a detail description of these databases see e.g. Bonaccorsi et al., 2010 and Daraio et al., 2011). The following project ETER (https://eter.joanneum.at/) give open access to the data at the level of individual HEIs. Currently data are available for 2465 HEIs in 32 countries and for three academic years: 2011, 2012 and 2013. Its detail coverage and comparison with our data will be discussed more thoroughly in the section dedicated to data collection.

[5] Wolszczak-Derlacz (2016) use the analogous data for European and American universities, but her analysis is focused on the productivity changes measured by Malmquist indices.

[6] Emrouznejad and Yang (2017) cover DEA-related studies for the period 1978–2016. They refer to more than 10 000 studies with only about 150 dedicated to education sector.

The next publications examine the Aquameth/Eumida dataset. A group of 79 universities in four European countries is examined by Bonaccorsi et al. (2007a). They focus in particular on the relationship between the size of the unit and its efficiency: finding economies of scale for the efficiency of education, decreasing economies of scale for research efficiency and lack of the relationship between the size of the unit and efficiency when the latter accounts both for education and research activity. In Bonaccorsi et al. (2007b), this time the level of analysis is four different disciplines confirming and a positive association between the size of a unit and efficiency. Finally, in a recent paper by Daraio et al. (2015a), the analysis is enlarged to 400 universities in 16 European countries, but refers only to a single year (2008/2009). This confirms that the size (economy of scale) and specialization (economy of scope) of a given university have a statistically significant impact both jointly and separately, showing an inverted u-shape effect on efficiency.

A two-stage analysis is performed by Wolszczak-Derlacz and Parteka (2011) on a set of 259 universities in seven European countries for the period 2001–2005. They show that more efficient universities have a higher number of different departments, a larger proportion of females among the academic staff, a higher percentage of funds from external sources and are older. In their next paper (Parteka and Wolszczak-Derlacz, 2013), they utilise the same set of units to calculate Malmquist indexes and find an average annual growth of 4%.

Entire higher education sectors (where the units of analysis are whole countries) are analysed by Agasisti (2011) and Aubyn et al. (2009). In the first-mentioned publication, an analysis of the performance of 18 OECD countries is conducted and on the basis of a tobit regression, the author postulates a positive correlation between the GDP per capita of a given country and the efficiency of its higher education system only when other control variables are included. In some specifications, the percentage of public funding on tertiary education is negatively correlated with its efficiency. Aubyn et al. (2009) show that a good-quality secondary system, output-based funding rules, independent evaluation of institutions and staff policy autonomy are positively related to efficiency.

In contrast, Joumady and Ris (2005) examine the efficiency of universities at the lowest level of aggregation – based on the level of generic and vocational competencies acquired by graduates from 209 HEIs in 8 countries.

To the best of the author's knowledge, there are only three studies concerned with intercontinental (Europe versus the U.S.) analysis of HEI efficiency using the DEA approach. However, they are related to specific cases and no general conclusion can be drawn. Reichmann and Sommersguter-Reichmann (2006) evaluate the efficiency of 118 university libraries in Australia, Austria, Canada, Germany, Switzerland and the United States, utilising a specific library-related input/output mix (see Table B1 in the Appendix B in Supplementary material). They find that non-European libraries are more efficient. Colbert et al. (2000) determine the relative efficiency of three foreign MBA programmes as compared to seven top-ranking U.S. MBA programmes. However, they find that only one programme is inefficient, which is probably due to the low discriminatory power of such a small number of analysed units in relation to the number of inputs and outputs. Finally, Wolszczak-Derlacz (2016) performs the analogous analysis to ours, but with the focus on productivity changes between European and American institutions (measured by Malmquist indices). She shows that, a rise in TFP is registered for the whole European sample (strongest for Dutch and Italian HEIs), while the productivity of American HEIs suffered a slight decline over the period 2000–2010.

Critical comments to the previous studies refer mainly to the data *per se* (limited time and country coverage) and utilisation of improper or questionable techniques e.g. tobit model in a second stage regression as in: Agasisti (2011), Agasisti and Pohl (2012) or lack of sensitivity analysis: Veiderpass and McKelvey (2016).

In view of these facts, the present study can be claimed to be the first one considering such a broad cross-country coverage and concentrating on a Europe-US comparison with an intensive sensitivity part, and thus can partly fill a gap in this literature.

3. Data and key characteristic of HEIs

In order to guarantee a relative homogeneity, the country's population of HEIs, is restricted to public (private government dependent in the UK) academic institutions awarding doctoral degree excluding specialist entities such as military, music, sport and theatre academies.[7] The private sector differs considerably, e.g. in terms of the legislation under which it operates, funding etc. and its analysis is beyond the scope of this paper. In the case of binary systems (e.g. German or Austrian Fachhohschule) only general and technical universities are taken into account.

Although the final sample was conditioned by the feasibility of collecting complete data, the representativeness (in terms of enrolled students and number of graduates) with respect to the country's reference population of HEIs (defined as public academic HEIs awarding doctoral degrees excluding specialist entities) is high. Altogether we have information on 348 universities in ten European countries (Austria, Finland, Germany, Italy, the Netherlands, Poland, Spain, Sweden, Switzerland and the United Kingdom) for the years 2000–2012 and 152 in the United States for the years 2000–2010.[8] Table B2 in the Appendix B in Supplementary material shows the coverage for the analysed countries, which goes from 61.5% for Poland to 96% for Spain.[9]

Individual European countries vary considerably in terms of providing information about HEIs.[10] Data come either from national statistical offices (Germany and Switzerland), ministries of education (Austria, Finland, Poland) and/or relevant authorities/agencies (e.g. Spanish Rectors Conference (CRUE), Association of Netherlands Universities (VSNU) etc.). Some of the information was extracted directly from the financial reports of individual institutions. A detailed description of the sources of the data together with definition of all variables is presented in Table B3 in the Appendix B in Supplementary material.

For the U.S. institutions, data come from The Integrated Postsecondary Education Data System (IPEDS). IPEDS covers all higher education institutions in the U.S., but it was decided to limit the population to only those classified by the Carnegie Foundation as public 4-year or above institutions conducting research, in order to guarantee

[7] We exclude also distance learning universities e.g.: Università Telematica in Italy, Open Universiteit in Netherlands and The Open University in the UK.

[8] A detailed list of all the universities covered by this study is available from the author on request. Since the DEA methodology requires the same number of institutions with a complete set of variables for every year of the analysis, in the case of missing values a regression imputation procedure was employed (e.g. data for Spanish universities were available only for every second year). In the part dedicated to the sensitivity analysis we present whether the data imputation alter the results. We thank the anonymous referee for pointing this out.

[9] The student coverage is calculated with respect to the total number of students enrolled (ISCIC 5–7), while the population of students refers to the sum of all public (and private government dependent in the UK) academic HEIs awarding doctoral degrees as present in ETER database. The same applies to graduates. An alternative measure of coverage (the sum of the students and graduates of all HEIs as reported in the ETER database) goes from 40% for Poland to 85% for Sweden.

[10] In view of these facts, the creation of one common publicly available database of HEIs in Europe such as ETER is a milestone of data collection. Unfortunately, when we started our project, ETER was not freely available. In most cases our data sources are the same as those of the ETER. The pairwise correlations between the variables in our data and ETER for 2011 and 2012 (there is overlap only for these two year in both data sources) are very high (0.97 for numbers of students, 0.94 for academic staff, 0.98 for total staff, 0.96 for graduates, 0.82 for total revenue). The main difference refers to the completeness of data. Our data is a balanced panel (no missing observations in any of the inputs and outputs), while in the ETER, despite its enormous coverage, the level of completeness greatly varies by country, domain and variable, e.g. it is least complete for financial data, missing more than 50% of cases for some financial categories (Daraio et al., 2016 p. 4).

comparability with the European sample.

Despite extreme efforts put into ensuring the correspondence and reliability of the data (in particular by following the data collection manuals: UNESCO-UIS/OECD/Eurostat (2010), Frascati Manual: OECD (2015), some noteworthy differences across countries emerge, e.g. the total number of academic staff is expressed either in full time equivalent or in full time employment (in the cases of Germany and Poland), the total number of students in Poland is recorded without foreigners, and in case of part-time student enrolment the total number of students is estimated by multiplying part-time enrolment by a factor of 0.3 while for the U.S. by a factor that varies by level of the institution and level of study provided by the Digest of Education Statistics. Total revenues which were originally reported in national currencies were recalculated into real (2005 = 100) euros. The total revenues were divided into prime sources (core funding, mainly from governments in the form of teaching or/and operating grants), student fees and third sources (e.g. from investments, donations etc.). The precise definition of respective revenue's shares is depended on the specific HE systems – see Table B3 in Supplementary material. The teaching output is measured by the total number of graduates; the research output is proxied by the number of publications indexed in the Web of Science of academics affiliated with a given institution.[11] Finally, some of the information (e.g. year of establishment, number of different departments, location) was obtained directly from the web pages of individual HEIs.

We are aware that due to the data limits and somewhat heterogeneous variable definitions across countries our analysis might not be free of a measurement error. We address this with a number of robustness checks. Additionally, we consider that the specific results of our study should be interpreted very cautiously, especially taking into account the limitations of the data on the one hand, and the importance and sensitiveness of the topic on the other.

Table 1 presents the key descriptive statistics on the institutions in our sample.

The first column shows the number of publications per academic staff member, which can be treated as a partial measure of scientific productivity. The highest value is achieved by Dutch HEIs, where in the period analysed one academic "produces" on average 1.5 publications per year. This is followed by the U.S. with a value of 1. Of course, it should be emphasized that these are average values and the variation within countries is considerable, e.g. in the Netherlands for the Rotterdam Erasmus University, the best university in terms of the number of publications per academic in 2012, the indicator equaled 4.5, and for the weakest unit – the University of Tilburg – below 1. When the indicator is expressed as the number of publications to institution's revenues the ranking changes. The Dutch institutions are still in first place, followed by German and Italian universities, with Polish HEIs in fourth place now. The improvement in the latter's position reflects the relatively low level of funding of HEIs in Poland, which is confirmed by the relationship of total revenues to students (column 5). The financial aspects of HEI functioning also differ between and within countries.

Overall, U.S. universities have many more resources than their European counterparts, except from Swiss universities which are characterised by the highest value of revenues per student: see column 5 of Table 1 and Fig. 1.

In the last two columns of Table 1, the sources of institutional revenue are presented. The lowest share of funding from primary sources (government) is recorded for universities in the U.K., where only 40% of total revenues come from core funding. In the U.S., almost 65% of revenues come from the total government appropriations (federal, state, and local), while state funds constitute around 30%. Again, Fig. 2 underlines the within and between countries variability in this context.

If we look at the share of the fees paid by students, the highest value is recorded for the U.S., where on average 25% of university income comes from fees paid by students. Over the analysed period of time there is a slight increase in revenues per student – shown in Table 2 – both for U.S. and European institutions. However, the trend is not common for all European countries; specifically, in Germany, Italy, the Netherlands and Sweden there was a drop in the ratio. Table 2 presents also changes in the share of revenue coming from government resources (column 4 and 5). In both groups there is a decline in the percentage (in the U.S. the drop is much more pronounced). The drop in the share of revenue from government sources is accompanied by an increase in revenue from tuitions fees.

In the second step of our analysis we will examine how this changed revenue structure influenced technical efficiency.

4. Assessment of higher education institution efficiency using DEA

The critical part of this stage is the definition of the inputs and outputs of university activity. The choice is guided by the state of the art (the inputs and outputs used in previous cross-country studies are reviewed in Table B1 in Supplementary material). However, it is also the result of the feasibility of collecting comparable data. The benchmark model (Model 1) considers three inputs: academic staff, total revenue[12] and total number of students; and two outputs: publications and graduates. Alternatively, we calculate Model 2: with two inputs and two outputs (without students as input) and a 4-input/3-output model (Model 3 inputs: academic staff, non-academic staff, total revenues, students; outputs: scientific articles, publications other than scientific articles, graduates) and Model 4: 1 input/2-output (input: total revenues). Furthermore, two different frontiers are distinguished: global frontiers (all HEIs pooled together) and European versus U.S. frontiers (European countries pooled together). Table B4 in the Appendix B in Supplementary material presents the basic descriptions of the DEA models together with justification of their different input/output mix and normalisation strategy.

We proceed with the analysis by evaluating output-orientated efficiency models with variable returns to scale (VRS) for every year between 2000 and 2012.

The first stage DEA results are presented in Table 3 – as country and period means and medians.[13] As can be seen from Table 3, the mean and median efficiency scores vary greatly between and within the analysed countries. The U.K., Poland, the Netherlands and Italy are the most efficient countries with the lowest mean efficiency scores (the lower the score, the higher the efficiency). The mean and median value for the whole European sample is 1.59 under the assumption of a common frontier and this drops to 1.50 for the European-US frontier. In both cases the values are greater than for U.S. universities. Since we are assuming an output-oriented approach, an inefficient university would have to increase its output by a factor of (DEA score -1) $\times$ 100% in order to reach the frontier. Therefore, the efficiency score of 1.57 (1.35) for the U.S. indicates that they could improve their output as much as by 57% (35%) keeping their inputs stable.

So far, we have employed either a global or a regional (European/US-specific) frontier. In both cases a single common frontier across

[11] In order to determine the number of publications of various universities, the total number of works in which at least one of the authors reported a working place her/his institution was counted for consecutive years during 2000–2012. The usual criticisms of bibliometric data on the results of research activity are also relevant to our study (overrepresentation of publications in English, different publication practices across fields etc.) – see e.g. Haustein (2016). We would like to thank an anonymous referee for pointing this out.

[12] Our use of total revenue as an indicator of the financial resources of HEIs is driven by the feasibility of collecting financial data for different countries. However, the correlation between revenue and expenditure or costs (for countries for which we possess such information, e.g. Sweden) is very high and equals 0.8.

[13] The detailed results of the DEA scores for each institution for each year and all the different DEA models are available from the author upon request.

Table 1
Key statistics on HEIs – mean values by country, time period 2000–2012.[a]
Source: Own elaboration.

country	Publications per academic staff member	Publications per 1 m revenue	Graduates per academic staff member	Total number of students	Revenue per student per year	Revenue from government funding in% of total revenue	Revenue from tuition fees in% of total revenue
Austria	0.60	4.27	1.75	20386	9448	78	n.a.
N = 11	(0.27)	(1.74)	(0.80)	(19421)	(4846)	(8)	
Finland	0.63	4.83	1.61	12275	11028	65	n.a.
N = 13	(0.33)	(2.26)	(0.73)	(8639)	(2841)	(7)	
Germany	0.57	7.04	1.43	17911	9689	64	n.a.
N = 65	(0.30)	(4.85)	(0.73)	(10698)	(3871)	(12)	
ITALY	0.89	5.47	4.56	30076	5651	81	14
N = 54	(0.42)	(2.52)	(1.52)	(24841)	(2294)	(8)	(6)
Netherlands	1.54	7.09	2.23	18424	23971	60	7
N = 10	(0.95)	(2.35)	(1.33)	(6336)	(5554)	(8)	(2)
Poland	0.22	5.49	3.09	21262	2346	65	19
N = 30	(0.13)	(2.46)	(1.06)	(9974)	(797)	(7)	(8)
Spain	0.35	4.90	1.78	28620	4239	n.a.	n.a.
N = 47	(0.18)	(2.20)	(0.45)	(19560)	(1072)		
Sweden	0.68	3.19	2.77	11436	16028	72	n.a.
N = 24	(0.68)	(2.60)	(1.16)	(7885]	(16316)	(11)	
Switzerland	0.94	5.58	0.84	11849	31729	88	n.a.
N = 9	(0.32)	(1.54)	(0.41)	(5550)	(13253)	(5)	
UK	0.76	3.87	5.14	18368	12707	40	24
N = 85	(0.56)	(2.72)	(2.03)	(7254)	(7824)	(9)	(9)
Europe	**0.65**	**5.05**	**3.18**	**20768**	**10776**	**61**	**19**
N = 348	(0.51)	(3.32)	(2.09)	(15636)	(9272)	(19)	(10)
US	**1.04**	**2.53**	**3.90**	**21885**	**26101**	**64**	**25**
N = 152	(0.76)	(1.77)	(1.50)	(15755)	(16321)	(12)	(10)

[a] Data for the U.S. 2000–2010, in the case of missing values a regression imputation procedure was employed (e.g. data for Spanish universities were available only for every second year). Standard deviation in parenthesis. Revenues expressed in real euros, prices from 2005.

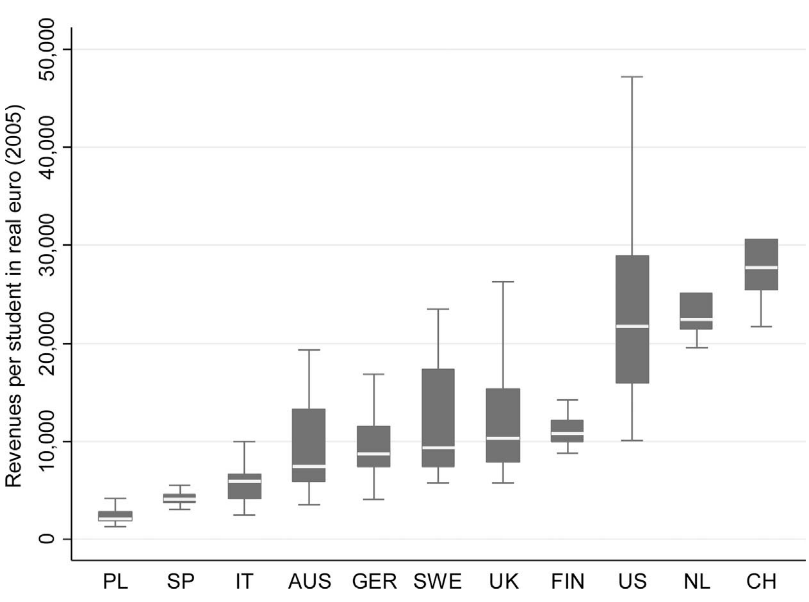

Fig. 1. Revenue per student (mean 2000–2012*), between and within country variability.
Note: * data for the US: 2000–2010. Revenues expressed in real euros, prices from 2005
Source: authors' elaboration.

European countries was assumed, which given the heterogeneity of European higher education systems may be a questionable choice. This can also result in relatively high level of efficiency scores obtained previously – the global or European benchmark for some of the countries can be too difficult to be reached. Therefore, we now test country-specific frontiers: we calculate separate DEA models for each country (in which every HEI is evaluated with respect to the units from the same country). However, such an exercise can only be performed for the DEA models with limited numbers of inputs and outputs (DEA model 2 and DEA model 4) because for some of the countries in our sample the number of units in not sufficient to estimate the frontier and ensure a reasonable level of discrimination.[14] The results of this exercise are presented in Table B6 in the Appendix B in Supplementary material. We see that the values of efficiency scores for European countries drop considerable: the most efficient HEIs are now Switzerland and Austria, the mean value for the whole Europe is 1.24. However mean and median scores for the U.S stay at the relatively high levels. This exercise showed that the frontier definition is important for the results of

[14] See e.g. Dyson et al. (2001) for a discussion of the low discriminatory power of DEA models.

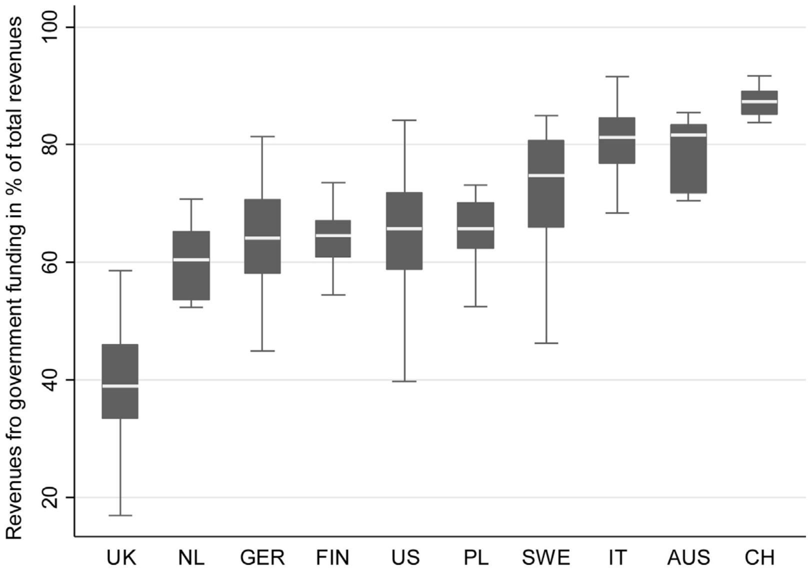

Fig. 2. The share of revenue from government funding (mean 2000–2012*), between and within country variability.
Note: *data for the US: 2000–2010. For Spain data not available.
Source: authors' elaboration.

Table 2
Revenue per student in real euros and the share of revenue from government funding, 2000–2012.[a]
Source: Own elaboration.

Country	Revenue per student in real euros		The share of revenue from government funding	
	2000	2012	2000	2012
Austria	7495	8337	87	74
Finland	10320	13375	65	60
Germany	10672	9900	65	68
Italy	5884	5104	81	74
Netherlands	29357	21727	63	57
Poland	2007	2885	66	61
Spain	3308	5655	n.a.	n.a.
Sweden	16343	16170	65	72
Switzerland	30679	33374	86	93
UK	10878	14049	43	33
Europe	9570	10404	63	59
US	24436	28127	67	58

[a] Data for the U.S. 2010 (not 2012) in the case of missing values a regression imputation procedure was employed. Revenues expressed in real euros, prices from 2005.

Table 3
Summary statistics for efficiency measures using a common and European-US frontier.
Source: Own elaboration.

	Global frontier			European – US frontier		
	Mean DEA score	Median DEA score	Std. dev.	Mean DEA scores	Median DEA score	Std. dev.
Austria	2.25	2.27	0.41	1.95	1.94	0.42
Finland	2.15	2.18	0.48	1.90	1.91	0.38
Germany	1.83	1.81	0.53	1.71	1.67	0.46
Italy	1.50	1.40	0.40	1.36	1.30	0.31
Netherlands	1.46	1.32	0.43	1.35	1.25	0.31
Poland	1.29	1.19	0.30	1.26	1.17	0.28
Spain	1.84	1.84	0.41	1.79	1.80	0.36
Sweden	1.83	1.84	0.45	1.75	1.76	0.42
Switzerland	1.88	1.80	0.30	1.80	1.76	0.26
UK	1.28	1.27	0.19	1.23	1.22	0.17
Europe	1.59	1.47	0.48	1.50	1.39	0.41
US	1.57	1.53	0.37	1.35	1.30	0.27

efficiency scores estimation (their magnitudes), further we will check whether it affects the results of the second stage analysis.

The kernel distribution of efficiency scores (pooling all years) by country is shown in Fig. 3. Most of the countries are characterised by a leptokurtic and skewed distribution with a concentration of mass in the lower tail in the direction of more efficient units. The exceptions are: Austria with the distribution shifted to the less efficient units on the right; Finland and Germany with a flatter distribution; and Spain and Sweden with a rather central distribution. These density estimates appear to graphically support the previous findings of a high variability of efficiency measures within and between countries.[15]

5. Exploring the determinants of inefficiency

5.1. Empirical specification

In the previous section of this study, a relatively high level of technical inefficiency of HEIs in European countries and the U.S. has been shown with a substantial variability in efficiency scores both between and within countries. From the policy perspective, it is interesting to examine the determinants of university efficiency, which can be helpful to answer the question of what can be done to improve it. In order to check whether the impact of the potential external factors (describing: institution size, department composition, funding schemes, and country- and region-specific characteristics) is common for European and U.S. HEIs, the following regression is estimated separately for the two subgroups, elaborating the general Eq. (1):

$$\mathrm{DEA}_{i,t} = \alpha + \beta_1\mathrm{GDP}_{n,t} + \beta_2\mathrm{DEP}_{i,t} + \beta_3\mathrm{FOUND}_i + +\beta_4\mathrm{REVGOV}_{i,t}/\mathrm{REVFEE}_{i,t} + \beta X_{ij,t} + u_{ijt}, \quad (2)$$

where: i refers to a single HEI, and t denotes the time period. The dependent variables are bias-corrected DEA scores which are regressed on potential covariates. X – refers to control variables such as a dummy equalling one if the HEI has a medical or pharmacy department (MED) and dummy for technical universities (TECH) to take into account the specificity of faculty composition and the level of cost that these

[15] The bias-corrected efficiency scores based on the bootstrap algorithm were also calculated. They are on average higher than the previous estimates. However, the countries' rankings are sustained and the shape of the distributions follows the previous ones (see Fig. B1 in Appendix in Supplementary material). Additionally, DEA scores were calculated on different DEA models but the results are similar (the Pearson correlation matrix is offered in Table B5 in the Appendix B in Supplementary material).

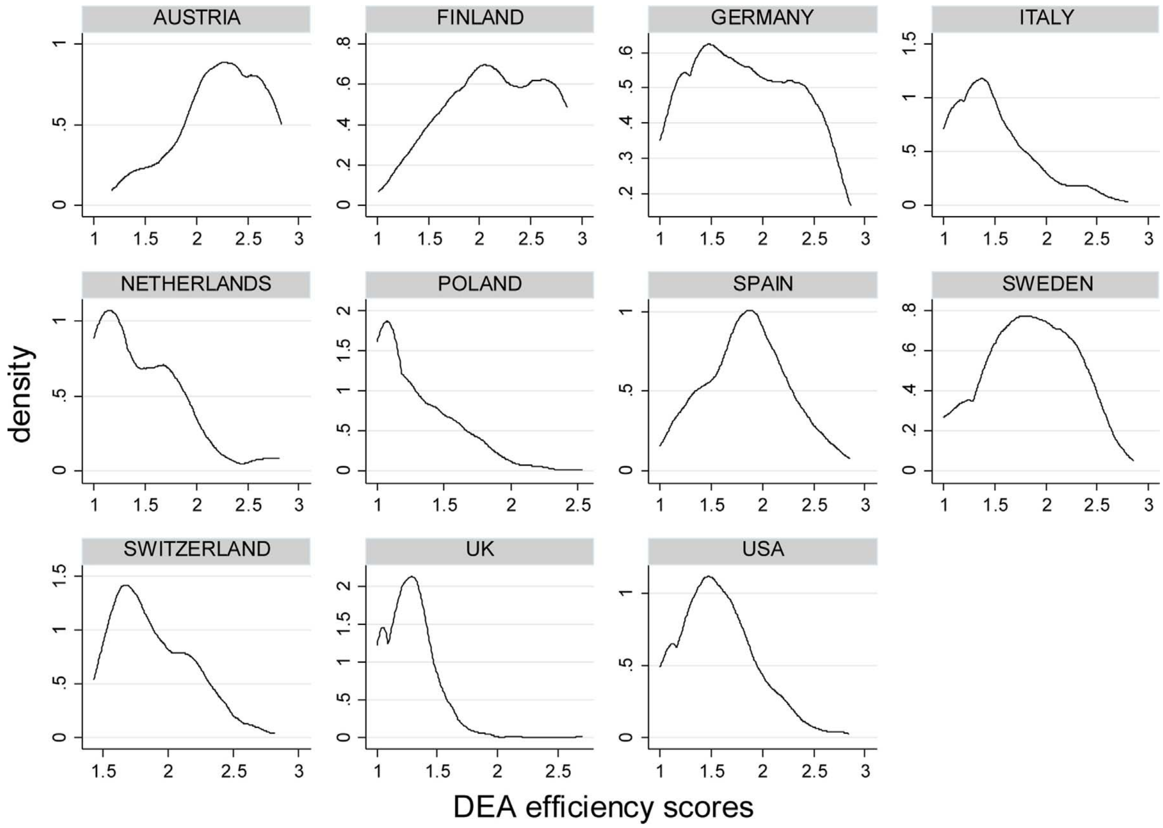

Fig. 3. The distribution of efficiency scores by country (all years pooled), common frontier.
Source: Own elaboration.

departments can impose. Additionally we incorporate time effects and country dummies (j-in the case of the European sample) in order to gauge country-specific effects of HE systems.

Among the environmental variables we include a proxy for location expressed as GDP per capita of the region n (NUTS2) where the institution is located (GDP). For the U.S. sample GDP refers to the state. University location can have an ambiguous impact on performance: if institutions take advantage of a wealthy region (e.g. through cooperation with local business) then there should be a positive correlation between GDP and efficiency; however, it is also possible that universities revitalise poorer regions and an inverse relationship is plausible.

Next, a variable representing the number of different departments (DEP) is included. This can represent either an economy of scale (larger institutions have more departments) and/or an economy of scope (different departments representing various disciplines). The problem of the potential existence of economies of scale in higher education has been much debated (for a review of relevant studies see Bonaccorsi et al., 2007a). The general conclusion is that larger institutions are more efficient. Some studies have confirmed economies of scale up to a certain level after which diseconomies can materialise e.g. through excessive bureaucracy (see Daraio et al., 2015a).

Next, an association between the year when a given institution was established and its efficiency is also tested by employing a variable representing the year of foundation (FOUND). We may expect older institutions to be more efficient (for reasons of tradition and reputation); on the other hand younger units might be more flexible.

Finally, two variables representing the structure of funding are introduced: REV_GOV, representing the share of government funding in total revenues; and REV_FEE for the share of tuition fees. Due to a high correlation between these two variables (Pearson coefficient = −0.67) they are introduced in separate regressions. Although the relationship between a university's revenue structure and efficiency is of great importance from a political perspective, in previous studies the issue has been addressed in the limited way. For Europe, Wolszczak-Derlacz and Parteka (2011) show that the greater the share of core funding the lower the efficiency. In the U.S. context, the association can be different: Robst (2001) finds signs of an inverse relationship between the share of state funds and inefficiency, but without statistical significance when other variables are controlled for. Similar results (the more state funding the higher the efficiency, but again without statistical significance) are obtained by Sav (2012, 2013). He concludes that greater tuition-fee dependency promotes inefficiency in the case of American public universities. In the recent paper Bolli et al. (2016), although utilising parametric methodology, show that the competition for international public funds disciplines European universities as evidenced by a positive impact on efficiency, but at the same time it decrease the productivity of the best performing HEIs. Conversely, tuition fees enhance the productivity of the best performing universities but increase the spread of universities with lower productivity. Their analysis is performed for the HEIs from eight European countries for the period 1994–2003.

Our estimation strategy involves truncated regression and a bootstrap simulation methodology is employed to account for a potential serial correlation of the DEA scores and a possible correlation of the error term with the covariates, as discussed in Section 2.

5.2. Results of the benchmark model

The results of the benchmark regressions corresponding to the DEA scores for the Model 1: 3-input/2-output model with a common frontier are presented in Table 4 for the European sample, and in Table 5 for the U.S. Since the dependent variables are equal to or greater than one, a positive/negative sign on the estimated regression parameter indicates lower/higher efficiency. For each of the subsamples three specifications are reported: the first one not controlling for funding structure, next,

Table 4
The determinants of inefficiency scores for the European sample – DEA 3-input/2-output model with common frontier.
Source: own calculations.

	(1) Bias-adjusted coefficients	95% bootstrap confidence intervals		(2) Bias-adjusted coefficients	95% bootstrap confidence intervals		(3) Bias-adjusted coefficients	95% bootstrap confidence intervals	
		low	high		low	high		low	high
GDP	−0.309[c]	−0.429	−0.198	−0.323[c]	−0.454	−0.205	−0.227[c]	−0.321	−0.133
DEP	−0.029[c]	−0.036	−0.022	−0.025[c]	−0.033	−0.016	−0.005[a]	−0.013	0.002
FOUND	0.080[c]	0.066	0.100	0.085[c]	0.067	0.107	0.030[c]	0.015	0.043
REV_GOV				0.376[b]	0.088	0.686			
REV_FEE							−1.193[c]	−1.63	−0.794
Obs.	4174			3355			1725		

Note: Constants are not reported. Dummies for medical departments and technical institutions as well as year and country individual effects included in all models.
[a] Indicates that the value zero does not fall within the 90% confidence interval.
[b] Indicates that the value zero does not fall within the 95% confidence interval.
[c] Indicates that the value zero does not fall within the 99% confidence interval. Confidence intervals obtained from 1000 bootstrapping interactions.

REV_GOV is added, and in the third specification we substitute it with REV_FEE. In the first columns, the bias-adjusted coefficients from a basic regression are presented. The next two columns show the lower and upper bounds of the 95% bootstrap confidence interval, which is used to check the statistical significance of the estimation.

Two separate regressions are run: for European institutions and for U.S. ones only. In fact, a number of similarities are found for two subgroups, but also a couple of noteworthy differences. In all the specifications, the results reveal a negative and statistically significant coefficient for GDP for both the European and U.S. samples, indicating a greater efficiency of universities located in richer regions. Similarly, the statistical significance of the number of different departments is confirmed. The negative parameter in front of the DEP variable shows that HEIs with a greater number of different departments have lower DEA scores (more efficient), which can be a sign of economies of scope. However, it could also be a sign of economies of scale, as larger units usually have a greater number of different departments. The year of foundation is statistically significant only in the case of European institutions and its sign indicates that younger units are less efficient − in the case of the U.S. the coefficient is not statistically significant.

Turning to the potential impact of funding structure on the technical efficiency of universities, there are some interesting results. For the European sample (specification 2 in Table 4), the results indicate a positive relationship between the share of funds from government resources and inefficiency. However, this is not confirmed for U.S. institutions, for which the relationship is not statistically significant. In contrast, tuitions fees are negatively associated with the technical inefficiency of HEIs in Europe (specification 3 in Table 4) and a positively with the inefficiency of units in the U.S. (specification 3 in Table 5). Most of these results are confirmed when the regional frontier is imposed (see Tables B7 and B8 in the Appendix B in Supplementary material). However, for the American universities none of the revenue's shares are now statistically significant. Additionally, the control variables as well as time dummies (not reported due to the space constraints) are highly significant for all specifications.

5.3. Sensitivity analysis

A number of robustness checks are carried out in which we alter the input/output mix, the frontier definition, the set of independent variables for the second-stage regression and finally also go beyond the semi-parametric approach and employ a wholly nonparametric model based on conditional frontier analysis (Daraio and Simar, 2005; Bădin et al., 2012, 2014). The results of this are briefly described in this section and the details are presented in the Appendix C in Coelli et al., 2005; Cooper et al., 2004; Wilson, 2008.

First, to check the sensitivity of the results, the same exercise is repeated for DEA models with alternative sets of input and output variables (see Tables C1 and C2 in Supplementary material). Some noteworthy differences need to be acknowledged. For both the European and U.S samples, in all the specifications the negative correlation between regional development and inefficiency of the institutions is confirmed, as is the relationship between the number of different departments and inefficiency scores. In the case of European HEIs, their age is always associated with greater efficiency, but for the U.S institutions the picture is less clear. In most cases the coefficient is not statistically significant, but when it is, it goes in the opposite direction to that of the European institutions – the younger HEIs are more efficient (Table C2, columns 4 and 5 in Supplementary material). The negative correlation of core revenue with efficiency is confirmed for European universities (the coefficients are positive and statistically significant for three of the four models) as well the opposite direction

Table 5
The determinants of inefficiency scores for the U.S. sample – DEA 3-input/2-output model with common frontier.
Source: own calculations.

	(1) Bias-adjusted coefficients	95% bootstrap confidence intervals		(2) Bias-adjusted coefficients	95% bootstrap confidence intervals		(3) Bias-adjusted coefficients	95% bootstrap confidence intervals	
		low	high		low	high		low	high
GDP	−0.815[c]	−0.985	−0.643	−0.858[c]	−1.028	−0.684	−0.843[c]	−1.018	−0.667
DEP	−0.036[c]	−0.044	−0.029	−0.035[c]	−0.042	−0.028	−0.032[c]	−0.040	−0.025
FOUND	−0.021	−0.071	0.033	−0.019	−0.071	0.036	−0.007	−0.058	0.045
REV_GOV				0.156	−0.066	0.378			
REV_FEE							0.281[a]	0.017	0.583
Obs.	1976			1669			1670		

Note: Constants are not reported. Dummies for medical departments and technical institutions as well as year and country individual effects included in all models.
[a] Indicates that the value zero does not fall within the 90% confidence interval.
[c] Indicates that the value zero does not fall within the 99% confidence interval. Confidence intervals obtained from 1000 bootstrapping interactions.

for the share of fee revenue. As far as the U.S. is concerned, we cannot draw strong conclusions: neither REV_GOV nor REV_FEE are statistically significant. Further, we check the specification utilising now the country-specific frontiers – relevant for European sample. The results for REV_GOV and REV_FEE are confirmed, but for some specifications the coefficient for FOUND loses its statistical significance (Table C3 in Supplementary material). In additional robustness checks, a limited European frontier is created based on nine countries (as opposed to the original ten). We exclude each individual country from the sample one by one in order to check whether any country does not influence either the shape of the frontier or the results of the second-stage analysis.[16] Limitation of the country coverage yields qualitatively similar results (Tables C4 and C5 in Supplementary material).

The association between funding structure and efficiency are further investigated. First, we check country-specific models (restricting the second-stage analysis to a single country). As can be seen from Table C6 in Supplementary material, the positive correlation between REV_GOV and inefficiency is confirmed for six out of nine cases (we have no data for REV_GOV for Spanish HEIs), while the parameter is not statistically significant in the cases of Germany, Poland and Switzerland. As Fig. 2 shows, a high share of Swiss universities' budgets generally comes from government resources and the institutions are relatively homogenous in this respect. However, this is not the case for German and Polish HEIs, which are more heterogeneous. Unfortunately, in our setting we are not able to classify the different donors of external resources which can have the impact on the specific results (see e.g. Bolli and Somogyi, 2011 study of private and public third-party funds on the efficiency of Swiss universities).

The same exercise is performed for the share of revenue coming from tuition fees (Table C7 in Supplementary material). We only have data for four countries – Italy, the Netherlands, Poland and the UK – and apart from the UK the share of tuition fees is confirmed as being negatively and statistically significantly correlated with inefficiency scores. Interestingly, for British HEIs the coefficient of REV_FEE is not statistically significant. This result may be related to what was previously found for REV_GOV for Swiss universities: both countries have relatively high shares of one type of financial resource.

In addition, to take into account possible time delays in the impact of the shares of government/tuition-fee revenue on efficiency, and to account for reverse causality, the regression analysis is altered by including time-lagged values of the funding structure. The estimation results are similar to those obtained previously, but in the European model the coefficient on the tuition fee turns insignificant (Table C8 in Supplementary material).

Finally, we examine possible nonlinearities by augmenting our specification with the quadratic term of REV_GOV and REV_FEE (Table C9 in Supplementary material). Regarding share of government funding, nonlinearities can be excluded: the coefficients on both the linear and quadratic terms remain insignificant. It is confirmed that the share of tuition fees have different impact on the efficiency of European and American institutions. For the European sample linear term is negative while the quadratic term significantly positive (columns 3 and 4) which provides evidence on the u-shaped effect on inefficiency. However, the estimation turning point is relatively high (26–28% of budget share) and most of the observations (around 80%) lie on the decreasing side of the hump-shaped curve confirming that tuition shares are negatively correlated with European university inefficiency. For American institutions the quadratic term is either statistically non-significant or significantly negative which suggest an inverse u-shaped relationship, but most of the observations lie on the increasing side of the hump-shaped curve providing evidence that increases in the share of their budget financed by the tuition fees raises university inefficiency.

We further explore the robustness of the findings by altering the estimation method, the normalisation of the variables and the specification *per se.*[17] In the case of the American sample, regional dummies (to distinguish 8 geographical regions) are added to the specifications, and alternatively state dummies. In the latter case, the GDP variable loses its statistical significance, which is quite reasonable, but the remainder of the results are maintained.

The final sensitivity analysis regards the conditional efficiency frontier, which is said to be less sensitive to outliers, is not under the influence of the curse of dimensionality, and in which a separability condition is not required to provide meaningful results (Bădin et al., 2014). We compare the conditional and unconditional efficiency scores and regress them nonparametrically on our independent variables (Zs) from the second-stage analysis in order to check their marginal effects. In the case of our model, which is output-oriented, if a given variable is favourable it means that it operates as a freely available extra input and the ratio of conditional to unconditional efficiency will increase with the value of Z. When the ratio is decreasing, Z acts as an unavoidable output and has a negative effect on efficiency (Daraio and Simar, 2005, 2007). As we have not only continuous variables in the set of independent Zs but also ordered and unordered discrete ones, we apply De Witte and Kortelainen's (2013) methodology.[18] Analogously, as in the benchmark specification, we performed an analysis on the yearly basis (calculation the ratio of conditional to unconditional scores for each year separately), but this time we do not take into account the variable REV_FEE. This is because the conditional efficiency frontier (as a wholly parametric method) requires a complete set of non-missing observations (not only of inputs and outputs but also of Zs) and inclusion of REV_FEE would significantly decrease the number of HEIs in our European sample.

The results of nonparametric significance tests are presented in Table C10 in Supplementary material. The average effects are revealed by partial plots, where the ratio of conditional to unconditional efficiency (y axis) are shown for a given Z assuming that all other exogenous variables are at their median (Figs. C1 and C2 in Supplementary material). For Europe, the results indicate that the number of departments has a favourable impact on efficiency, as does the level of development of the region where the university is located. However, the statistical significance of the number of departments is not confirmed − in fact some nonlinearities are detected. The unfavourable impact of REV_GOV on efficiency and of year of foundation: older universities being more efficient are confirmed. For the U.S. sample (Fig. C2 in Supplementary material) we obtain similar results to those from the benchmark analysis but statistically significant are only DEP and GDP.

From this intensive sensitivity analysis, we can draw some conclusions as to which of our results can be generalised. There are positive associations between both regional GDP per capita and number of departments and an institution's efficiency (for both the European and U.S. samples). On average, older European HEIs are more efficient, but this is not confirmed for American ones. Finally, government funding seems to have a negative effect on the efficiency of universities in Europe, which again is not confirmed for the U.S. However, we should

[16] Bolli et al. (2016) argue that such an exercise can also check the character of the measurement error (random versus systematic error) due to potential data heterogeneity across countries.

[17] The detailed results for this part are available from the author on request. For example, a single bootstrap procedure is utilized regressing the 'original' DEA scores on the basis of the truncated bootstrap regression, altering the truncation points (from 0.999 to 1) and changing the number of bootstrap replications. Additionally, we perform the analysis with different normalisations of the variables and on the original data (excluding the imputed values). Finally, as alternatives to time and country dummies we also incorporate time-country-specific dummies in order to allow for differential time effects across countries. Generally, results are sustained, one noticeable difference is the drop in the statistical significance of REV_GOV (it is still positive and statistically significant but at a lower level).

[18] We would like to thank Kristof De Witte for providing us with the R codes as well as suggestions on ways to adjust them to our setting.

underline that our results show some country heterogeneity, and this should be analysed more thoroughly in future studies.

6. Conclusions and discussion of findings

In this study, DEA has been employed in order to evaluate the relative efficiency of a sample of 500 higher education institutions (from ten European countries and the U.S.) for the period between 2000 and 2012. This is the most comprehensive (as far as country, time period and input/output measures) dataset at the level of individual institutions to be employed for this purpose, and one of the first one composed of countries from different continents (to the best of the author's knowledge).

The results reveal a relatively high level of technical inefficiency of HEIs and a substantial variability in the efficiency scores both between and within countries. For European universities, the highest average technical efficiency score is obtained under the assumption of a common frontier suggesting that to become fully efficient at the global arena, much more outputs should have been generated. However, when the country-specific frontier is employed for European countries, the magnitude of their efficiency scores drops considerable indicating that they are more successful at the local markets.

In a second step of the analysis, the previously-estimated DEA scores were related to their potential determinants. Some major differences between the two groups regarding funding structure are evident. As far as the European sample is concerned, a shift to government funding as a revenue source decreases a university's technical efficiency while this relationship is statistically insignificant for the U.S. Furthermore, the results indicate that the technical efficiency of European universities is positively associated with tuition fees as a source of revenue. For both groups: European and American HEIs, economies of scale (larger units have higher efficiency) have been confirmed as has a positive impact of location (units located within wealthier regions are more efficient).

It would be interesting to compare our findings with previous studies, although a direct comparison is problematic due to the sample composition, the years covered, and the model specification (e.g. different input-output sets). The results are in line with Wolszczak-Derlacz and Parteka (2011)'s analysis, in which they assess the impact of the share of revenue from core funding (mainly governmental) on the technical efficiency of public universities in seven European countries. For the public U.S. sector, Sav (2012) finds high values of mean inefficiency, which decrease with the share of government funding (although the result is statistically insignificant).

It may seem quite surprising that U.S. institutions are quite inefficient, with efficiency scores slightly below the mean European value, but clearly above the levels that characterise the most efficient countries. A possible explanation may lie in the sample compositions and the exclusion of private institutions from the analysis, which in the case of the American higher education sector are quite successful and have relatively high efficiency (Sav, 2012). However, inclusion of the U.S. private sector would surely have distorted the present analysis and comparison between private and public institutions is beyond the scope of our paper.

Some shortcomings of this study need to be admitted, mainly regarding the specification and the limited number of inputs and outputs, which were measured purely quantitatively. Moreover, many of the outputs of HEIs are not measurable at all. For example, it is difficult to measure the so-called third mission – a university's contribution to the surrounding community. Consequently, the lack of adequate quality controls and omitted variables can bias the estimation e.g. greater expenditure on quality may have been attributed to inefficiency. However, in the cross-sectional time series analysis this problem should be less severe (Robst, 2001). Additionally, we want to underline that any strict causality between efficiency scores and their potential determinants could be problematic. For example, universities located in better-off regions can take advantage of wealthier surroundings, while it may be the case that efficient HEIs are more successful in revitalising the surrounding area (e.g. by providing a well-educated labour force). Moreover, efficient universities can attract more third-party funding; on the other hand, universities with a greater share of external funding may benefit from more financial resources and improve their efficiency.

Nevertheless, this analysis is one of the first attempts to compare the technical efficiency of European and U.S. HEIs. Importantly, it has shown distinct differences related to the potential impact of funding schemes on the efficiency of the institutions in these two groups. The question arises of why the share of government funding seems to bring disadvantages in terms of technical efficiency only in the case of European institutions, while it does not hurt the efficiency of their American counterparts. There are some possible explanations. The first of these is connected with the different procedures for obtaining these funds in Europe and in the U.S. For example, most U.S. federal grants are awarded through a competitive process (e.g. through the National Science Foundation, which distributes funds using merit-based research competitions). Additionally, more and more states have introduced a performance-based procedure to allocate funds among universities. Furthermore, as shown by Aghion et al. (2010), when universities receive a positive funding shock, they become more productive if they are more autonomous and face more competition, e.g. from private research universities, as in the U.S.

We argue that DEA techniques can be used as an additional tool to help strategic planning and/or evaluation. The efficiency scores obtained through DEA can serve as extra information guiding the institution management on what it should focus on and where changes are needed, e.g. a higher pressure on research or the teaching output or on any specific output. The positioning of a given institution against international benchmarks can help understanding of its relative strengths and weaknesses. Furthermore, DEA can provide exact information about which output should be increased and by what amount in order to reach efficiency.[19] Moreover, the results of the second step of our analysis can again be informative both to management and policymakers. In our opinion the most important result concerns public resources, which turn out to be detrimental to university efficiency (in the case of European institutions). These results indicate that third-party funding does not have to be a threat to the quality of teaching and basic research as it was advocated by some scholars (Salter and Martin, 2001).

Both this study and a recent initiative at the European level (ETER) show the feasibility of creating a comprehensive cross-country database at the level of individual institutions containing detailed descriptions of the resources of units together with the results of their activities. Nevertheless, efforts need to be strengthened. Daraio and Glänzel (2016) list a number of challenges identified in research and innovation data integration, among which are data quality, comparability, standardization, interoperability, extensibility, updating and data availability. Without addressing these challenges, further studies of the efficiency of HEIs and its determinants will be extremely difficult, despite the importance of such studies from the perspectives of university administration, students, and whole economies.

Acknowledgements

I am grateful to three anonymous referees for useful comments and suggestions. Part of this paper was written when the author was a visiting scholar at University of California, Berkeley. This work was supported by the Polish Ministry of Science and Higher Education under the programme 'Mobility Plus'. The research assistance provided by Sabina Szymczak is also acknowledged.

[19] Since DEA is non-parametric method based largely on quantitative techniques, for the final evaluation of HEI performance it should be accompanied by quality controls (e.g. based on expert or peer review). Additionally, applications cannot be made automatically without full knowledge of the methodology utilized (e.g. its limitations). We thank an anonymous refer for pointing this out.

Appendix A-C Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.respol.2017.07.010.

References

Agasisti, T., Johnes, G., 2009. Beyond frontiers: comparing the efficiency of higher education decision making units across countries. Educ. Econ. 17 (1), 59–79.
Agasisti, T., Pérez-Esparrells, C., 2010. Comparing efficiency in a cross-country perspective: the case of Italian and Spanish state universities. High. Educ. 59 (1), 85–103.
Agasisti, T., Pohl, C., 2012. Comparing German and Italian public universities: convergence or divergence in the higher education landscape? Manage. Decis. Econ. 33 (2), 71–85.
Agasisti, T., Wolszczak-Derlacz, J., 2016. Exploring efficiency differentials between Italian and Polish universities, 2001–11. Sci. Public Policy 431, 128–142.
Agasisti, T., 2011.. Performances and spending efficiency in higher education: a European comparison through non-parametric approaches. Educ. Econ. 19 (2), 199–224.
Aghion, P., Dewatripont, M., Hoxby, C., Mas-Colell, A., Sapir, A., 2010. The governance and performance of universities: evidence from Europe and the US. Econ. Policy 25 (61), 7–59.
Altbach, P.G., Gumport, P.J., Berdahl, R.O., 2011. American Higher Education in the Twenty-First Century: Social, Political, and Economic Challenges. The Johns Hopkins University Press, Baltimore.
Aubyn, M., Pina, A., Garcia, F., Pais, J., 2009. Study on the efficiency and effectiveness of public spending on tertiary education. European Economy Economic Paper 390. European Commission, Directorate-General for Economic and Financial Affairs, Brussels.
Bădin, L., Daraio, C., Simar, L., 2012. How to measure the impact of environmental factors in a nonparametric production model. Eur. J. Oper. Res. 223 (3), 818–833.
Bădin, L., Daraio, C., Simar, L., 2014. Explaining inefficiency in nonparametric production models: the state of the art. Ann. Oper. Res. 214 (1), 5–30.
Bolli, T., Somogyi, F., 2011. Do competitively acquired funds induce universities to increase productivity? Res. Policy 40 (1), 136–147.
Bolli, T., Olivares, M., Bonaccorsi, A., Daraio, C., Aracil, A.G., Lepori, B., 2016. The differential effects of competitive funding on the production frontier and the efficiency of universities. Econ. Educ. Rev. 52, 91–104.
Bonaccorsi, A., Daraio, C., Simar, L., 2007a. Efficiency and productivity in European Universities. Exploring trade-offs in the strategic profile. In: Bonaccorsi, A., Daraio, C. (Eds.), Universities and Strategic Knowledge Creation. Specialization and Performance in Europe Edward Elgar PRIME Collection.
Bonaccorsi, A., Daraio, C., Raty, T., Simar, L., 2007b. Efficiency and University Size: Discipline-wise Evidence from European Universities. MPRA Paper No. 10265. Munich Personal RePEc Archi.
Bonaccorsi, A., Brandt, T., De Filippo, D., Lepori, B., Molinari, F., Niederl, A., Schmoch, U., Schubert, T., Slipersaeter, S., 2010. Final Study Report. Feasibility Study for Creating a European University Data Collection. The European Communities.
Bonaccorsi, A., Haddawy, P., Cicero, T., Saeed, H., 2017. Explaining the transatlantic gap in scientific excellence. Scientometrics 110 (1), 217–241.
Bougnol, M.L., Dula, J.H., 2006. Validating DEA as a ranking tool: an application of DEA to assess performance in higher education. Ann. Oper. Res. 145, 339–365.
American Universities in a Global Market. In: Clotfelter, Ch.T. (Ed.), NBER, The University of Chicago Press, LTD, London.
Coelli, T.J., Rao, D.S.P., ODonnell, C.J., Battese, G.E., 2005. An Introduction to Efficiency and Productivity Analysis, 2nd ed. Springer, New York.
Colbert, A., Levary, R.R., Shaner, M.C., 2000. Determining the relative efficiency of MBA programs using DEA. Eur. J. Oper. Res. 125, 656–669.
Cooper, W.W., Seiford, L.M., Zhu, J., 2004. Handbook on Data Envelopment Analysis. Kluwer Academic Publishers, Hingham, MA.
Daraio, C., Glänzel, W., 2016. Grand challenges in data integration state of the art and future perspectives: an introduction. Scientometrics 108, 391–400.
Daraio, C., Simar, L., 2005. Introducing environmental variables in nonparametric frontier models: a probabilistic approach. J. Product. Anal. 24 (1), 93–121.
Daraio, C., Simar, L., 2007. Advanced Robust and Nonparametric Methods in Efficiency Analysis. Methodology and Applications. Springer, New York.
Daraio, C., Bonaccorsi, A., Geuna, A., Lepori, B., Bach, L., Bogetoft, P., et al., 2011. The European university landscape: a micro characterization based on evidence from the aquameth project. Res. Policy 40 (1), 148–164.
Daraio, C., Bonaccorsi, A., Simar, L., 2015a. Efficiency and economies of scale and specialization in European universities: a directional distance approach. J. Informetr. 9, 430–448 D.
Daraio, C., Bonaccorsi, A., Simar, L., 2015b. Rankings and university performance: a conditional multidimensional approach. Eur. J. Oper. Res. 244, 918–930.
Daraio, C., Scannapieco, M., Catarci, T., Simar, L., 2016. ETER Quality Report on the First Wave of Data Collection. European Tertiary Education Register (ETER), Rome.
De Witte, K., Kortelainen, M., 2013. What explains the performance of students in a heterogeneous environment? Conditional efficiency estimation with continuous and discrete environmental variables. Appl. Econ. 45 (17), 2401–2412.
Dyson, R.G., Allen, R., Camanho, A.S., Podinovski, V.V., Sarrico, C.S., Shale, E.A., 2001. Pitfalls and protocols in DEA. Eur. J. Oper. Res. 132 (2), 245–259.
Emrouznejad, A., Yang, G., 2017. A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. Socioecon. Plann. Sci. 61 (1), 1–5.
Haustein, S., 2016. Grand challenges in altmetrics: heterogeneity, data quality and dependencies. Scientometrics 108 (1), 413–423.
Joumady, O., Ris, C., 2005. Performance in European higher education: a nonparametric production frontier approach. Educ. Econ. 13 (2), 189–205.
Kwiek, M., 2015. Competing for public resources: higher education and academic research in europe. a cross-sectoral perspective. In: Brada, J.C., Bienkowski, W., Kuboniwa, W. (Eds.), International Perspectives on Financing Higher Education. Palgrave Macmillan, pp. 6–24.
OECD, 2015. Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities. OECD Publishing, Paris.
Parteka, A., Wolszczak-Derlacz, J., 2013. Dynamics of productivity in higher education: cross-European evidence based on bootstrapped Malmquist indices. J. Product. Anal. 40 (1), 67–82.
Reichmann, G., Sommersguter-Reichmann, M., 2006. University library benchmarking: an international comparison using DEA. Int J. Product. Econ. 100, 131–147.
Robst, J., 2001. Cost efficiency in public higher education institutions. J. High. Educ. 72 (6), 730–750.
Salter, A.J., Martin, B.R., 2001. The economic benefits of publicly funded basic research: a critical review. Res. Policy 30, 509–532.
Sav, G.T., 2012. Stochastic cost frontier and inefficiency estimates of public and private universities: does government matter? Int. Adv. Econ. Res. 18 (2), 187–198.
Sav, G.T., 2013. Effects of financial source dependency on public university operating efficiencies: data envelopment single-stage and tobit two-stage evaluations. Rev. Econ. Finance 3, 63–73.
Simar, L., Wilson, P.W., 2000. A general methodology for bootstrapping in non-parametric frontier models. J. Appl. Stat. 27 (6), 779–802.
Simar, L., Wilson, P.W., 2007. Estimation and inference in two-stage, semi-parametric models of production processes. J. Econ. 136, 31–64.
UNESCO-UIS/OECD/EUROSTAT, 2010. UOE Data Collection on Education Systems. UNESCO, Montreal, Paris, Luxembourg.
Veiderpass, A., McKelvey, M., 2016. Evaluating the performance of higher education institutions in Europe: a nonparametric efficiency analysis of 944 institutions. Appl. Econ. 48 (16), 1504–1514.
Wilson, P.W., 2008. FEAR: a software package for frontier efficiency analysis with R. Socioecon. Plann. Sci. 42 (4), 247–254.
Wolszczak-Derlacz, J., Parteka, A., 2011. Efficiency of European public higher education institutions: a two-stage multi-country approach. Scientometrics 89 (3), 887–917.
Wolszczak-Derlacz, J., 2016. Assessment of TFP in European and American higher education institutions-application of Malmquist indices. Technol. Econ. Dev. Econ. 1–22. http://dx.doi.org/10.3846/20294913.2016.1213197.