Characterizing the Heterogeneity of European Higher Education Institutions Combining Cluster and Efficiency Analyses

Renato Bruni[1], Giuseppe Catalano[1], Cinzia Daraio[1], Martina Gregori[2,1], and Henk F. Moed[1]

[1]bruni@diag.uniroma1.it, giuseppe.catalano@uniroma1.it, daraio@diag.uniroma1.it, martina.gregori@uniroma1.it, henk.moed@uniroma1.it

[1]Sapienza University of Rome, Department of Computer, Control and Management Engineering (DIAG), Via Ariosto 25, 00185, Rome (Italy)

[2]Sapienza university of Rome, Department of Mechanical and Aerospace Engineering (DIMA), Via Eudossiana 18, 00184, Rome (Italy)

Abstract

The heterogeneity of the Higher Education (HE) Institutions is one of the main critical issues to address properly the assessment of systemic performance. We adopt a multi-level perspective by combining national (macro) and institution (micro) level data and analyses. We combine clustering and efficiency analysis to characterize the heterogeneity of HE systems (at the national level) exploiting micro level data. We show also the potential of using micro level data to characterize national level performance. The obtained results may provide a quantitative support to identify the higher education institutions that need to be further investigated through qualitative case studies in political science analyses of HE systems.

Introduction

The measurement of academic performance is a relevant issue at the intersection between political science and informetrics. Although the analysis of the performance of Higher Education Institution (HEI) systems is a complex task, there are numerous international comparisons (rankings) of institutions such as Shanghai, Times Higher Education and Leiden Ranking that are published on a regular basis. The heterogeneity of the HEIs is one of the main critical issues to address properly the assessment of performance, in a multi-level (systemic) perspective. There are different sources of heterogeneity, including the mission, the national context, the presence or absence of medical schools, the legal status and the disciplinary orientation and degree of specialization (López-Illescas et al., 2011; Daraio et al., 2011).

We adopt a multi-level perspective by combining national (macro) level data and institution (micro) level data and analyses. We show also the potential of using micro level data to characterize national level performance. In a way we attempt to characterize HEIs accounting for their (i) *Structural* heterogeneity (structure of the national system: systemic factors, e.g. number and types of HEIs that are at place, governance factors), (ii) *Internal* heterogeneity (linked to the type of the production process carried out within the HEIs) and (iii) *Other* heterogeneity sources.

This work presents the first results from a larger project (see Acknowledgements), aimed to study the activities, the performances and the efficiencies of European HEIs. It focuses on a statistical exploration of a series of indicators linking Education, in a *systemic way*, with Research and Innovation. In terms of data analysis, it explores the combination of statistical data from ETER, the European Tertiary Education Register, with bibliometric data obtained from the Leiden Ranking, and with categorizations of national higher education policies obtained from more qualitative studies of national HEI systems. In the project, the existing problems of data availability, quantification and comparability go hand in hand with the need of conceptualization of the performance model before making the analysis (Daraio and Bonaccorsi, 2017). The notion of performance is characterized in a "progressive" way, starting from production ("volume" or extensive variables), going to productivity (intensive or

"size-independent" indicators of production), up to efficiency (combination of outputs/inputs) and more elaborated efficiency models, towards effectiveness and impact (Daraio, 2019).
The present work is organized in two parts. In the first part we tackle the heterogeneity of HEIs calculating country-level statistics based on micro data and analyzing them with qualitative and governance variables. We will call this section the Quali-quantitative analyses. In the second part of the work, we give an order to this heterogeneity calculating a teaching and research productivity score and providing a cluster analysis that allows us to identify some typologies of HEIs.
The main objective of this work is then to combine clustering and efficiency analysis to characterize the heterogeneity of HE systems (at country level) exploiting micro level data.

Methods

The methods used are K-means (MacQueen J.B., 1967) and DBSCAN (Density-Based Spatial Clustering of Applications with Noise; Ester, M. et al. 1996) clustering approaches, and nonparametric efficiency analysis (Free Disposal Hull, FDH estimation of efficiency scores and a more robust nonparametric estimation in progress, see Daraio and Simar (2007).
K-means is a well-established clustering technique. It aims at partitioning n observations into k clusters in which each observation belongs to the cluster with the nearest mean (which actually constitutes the centroid of the cluster). The application of this principle leads to a partition of the data space into Voronoi cells. Data are therefore iteratively clustered in n groups of equal variances, minimizing a criterion known as the inertia or within-cluster sum-of-squares. This algorithm requires the number of clusters to be specified.
DBSCAN, on the other hand, is a more recent clustering technique but is it one of the most used and cited approaches. The DBSCAN algorithm views clusters as areas of high density separated by areas of low density. Therefore, it groups together points that are closely packed together (points with many nearby neighbours), marking as outlier points that lie alone in low-density regions. This set of outliers can eventually be viewed as the last or residual cluster. Due to this density-based approach, the clusters obtained by DBSCAN can be of any shape, as opposed to K-means which assumes that clusters are convex shaped, and the number of clusters cannot be specified in advance.
We estimate the efficiency of universities in producing teaching and research and use the efficiency scores as an additional variable to characterize the groups of universities obtained from the cluster analyses. The DBSCAN cluster analysis lead us to identify three clusters. After that, we run the K-means clustering to characterize the three groups of universities. The combination of the two different approaches was useful to shed some lights on the robustness of the choice done in the K-means approach.

Data

A HEIs performance evaluation analysis, to be as much as possible representative and complete, needs to take into account indicators related to all the different activities carried out in the academic operations, namely teaching, academic research and Third Mission activities (collaboration with industries, patents, etc.).
With the purpose of gather information about the three aforementioned areas, it was made use of different sources. In particular, the following three databases were integrated for the analysis at the micro-level (single institution):

- ETER database, *for the information at micro-level (single institution), regarding the TEACHING area*
- CWTS Leiden Ranking database, *for the information regarding the academic research*

PATSTAT PATENTS database, *for the information regarding the patents registered*

The list of the variables considered is reported in Tables 1, 2, 3. In addition, a database dedicated to the national regulatory characteristics of European countries was integrated in order to outline part of the heterogeneity present among higher education macro-systems. The considered governance indicators (reported in Table 3) are elaborated in Capano and Pritoni (2019), cover the period 1988–2014 and consider 12 European countries[1]. The governance indicators are in total 24, grouped in 4 dimensions (Regulation, Expenditure, Taxes, Information) and represent the number of government interventions on the observed period in each specific sub-areas. In order to include these indicators in our analysis, we applied few trasformation on data. For each country, the scores per dimension were summed together; next, per dimension a *percentage* score was calculated relative to the total score on this dimension over all countries. Moreover, 3 further indicators have been calculated based on the Capano and Pritoni (2019) data, to try to capture the trends of national government towards a more or less restrictive approach on HEIs system regulation and verify the grade of application of *control measure* on micro-level performance (see Table 3).

In the final dataset, it was necessary to structurally internalise temporal lags between inputs and outputs information. It is well known that a certain period of time has to pass in order to observe effects related to the interventions on academic staff, academic funds, and so on. As it is usually done in the empirical analyses, one year lag to observe effects on academic research publication and two years lag to observe effects on patents applications are acceptable average periods to be assumed. Hence, the data considered refer to the following time ranges: 2011–2014, ETER database (teaching and basic information on inputs); 2012–2015, CWTS data (academic research information)/INCITES database; 2013–2016, PATENTS database.

The teaching outputs (mainly, number of graduates for each degree class) are referred to the same time period of the inputs variables (e.g. academic staff, funds). The choice was driven both by the lack of data of high quality and completeness for years after 2014 and the difficulty in establishing an acceptable lag, due to the different degree classes considered in the analysis. Nevertheless, it has been verified that the annual values assumed by the teaching outputs variable do not vary significantly year by year, in a short range of time.

Furthermore, data on numbers of inhabitants (obtained from OECD and EUROSTAT for the year 2016) are used to draw the possible relation between countries dimensions in population and the HEIs produced output.

The final dataset contains the average variable values over the considered period of each included database; missing values had been excluded from the calculation.

Table 1. Research funding based indicators

Use of metrics in education *(0=NO; 0.5=LIMITED, 1.0=YES)*
Research performance based funding *(0=NO; 0.5=LIMITED, 1.0=YES)*
Use of quantitative formula in research funding *(0=NO; 0.5=LIMITED, 1.0=YES)*
Use of peer review in research funding *(0=NO; 0.5=LIMITED, 1.0=YES)*

We selected all the institutions categorized as universities in the ETER dataset, and for which data are available both on staff, students, graduates, and on publications and citations in the Leiden Ranking dataset. The total number of selected institutions for all ETER countries combined amounts to 664. Nevertheless, due to the presence of missing values on key

[1] The countries considered in Capano & Pritoni (2019) are: Austria, Denmark, England, Finland, France, Greece, Ireland, Italy, Netherlands, Norway, Portugal and Sweden.

variables (namely, academic staff and number of enrolled students) with respect to the cluster analysis procedure, the second quantitative analysis was performed only on a sub-selection of the database composed by 383 HEIs from 22 countries.

Table 2. Definition and source of variables at micro-level

Category (Source)	*Variables*	*Definition*
Cluster analysis variable (ETER, CWTS for Pub_fract)	**Grads_ISCED.5-7/ACADstaff**	Total graduates ISCED 5-7 divided by the Total academic staff (in Full Time Equivalent, FTE); both values represented by yearly averages.
	Pub_fract(av)/ACADstaff	Number of publications (fractional counting) divided by the Total academic staff (FTE); both values represented by yearly averages.
Efficiency analysis	**Mod.Teach.Res.X_ACADSTAFF.FDH**	FDH Inefficiency scores. It may be higher or equal to 1. It is 1 for efficient units, higher than 1 for units that can expand the production of their outputs.
Basic data (ETER)	**Foundation_year**	HEI foundation year.
	Uni_Hospital	Dummy; 1 = presence of an Hospital in the Institution.
	Enrolled_student_ISCED.5-7	Total student enrolled at ISCED 5-7.
	ACADstaff_FTE	Total academic staff (expressed in FTE).
	PhD_intensity_2014	PhD intensity (year of reference: 2014).
	FullProf/ACADstaff_Head	Percentage of full professor on the total academic staff.
	WomenProff_share	Percentage of women on the total number of professors.
	Admn/TOTstaff_FTE	Percentage of administrative staff on total staff (academic plus administrative).
Third mission (Funds ETER)	**Funds_external%**	Percentage of funding from third parties on total funding.
	Funds_third_part/ACADstaff_FTE	Third party funds per academic staff (expressed in FTE).
Specialization (ETER)	**Specialization**	Express the specialization with respect to the disciplinary areas; it is calculated making reference to the Herfindahl index on academic staff. The missing values are filled in with the Herfindahl index on PhD graduates and, in few cases, Herfindahl index on students ISCED 5-7. The values refer to the year 2014.
Research quantity and quality (CWTS, ETER for Acad_staff)	**Pub_top10(av)/ACADstaff**	Number of papers in top 10% (yearly average) divided by the Total academic staff (FTE).
	Pub_in_top10%	Percentage of papers in top 10% (yearly average).
	Pub_international_coll	Percentage of papers with international collaborations (yearly average).
	mnsc_(w-av)_av	Papers mean normalized citations (yearly average, weighted by the number of patent applications).
Third mission (PATENTS, ETER for Acad_staff)	**Patent_application(av)/ACADstaff**	Overall total number of patent applications (yearly average).
	Back_citations(av)/ACADstaff	Number of patents' backward citations (yearly average).
	NPL_av	Number of academic papers citations for patents (yearly average).
	NPL_av/SPA_av	Number of citations from academic papers for each patent (yearly average).

Table 3. Definition and sources of variables at macro-level

Governance (Capano and Pritoni, 2019)	**GOV_Regulation**	Percentage of policy intervention on Regulation [assessment, evaluation and accreditation; agency of assessment, evaluation and accreditation; content of curricula; academic career and recruitment; regulation on students (admission and taxation), institutional and administrative governance; contracts].
	GOV_Expenditure	Percentage of policy intervention on Expenditure [Grants; subsidies and lump-sum funding; targeted funding; loans; performance based institutional funding; standard cost per student].
	GOV_Taxes	Percentage of policy intervention on Taxes [tax exemption; tac reduction for particular categories of students; service-based student fees].
	GOV_Information	Percentage of policy intervention on Information [transparency; certification; monitoring and reporting].
	GOV_Cons_trend	In each country, percentage of regulatory interventions aimed to add *more constraints* respect to the overall regulatory interventions in Regulation.
	GOV_Opp_trend	In each country, percentage of regulatory interventions aimed to add *more constraints* respect to the overall regulatory interventions in Regulation.
	GOV_Control_measures	In each country, percentage of regulatory interventions in the monitoring and reporting, rules on goals in teaching, assessment subjects, respect to the overall regulatory interventions.
System structure (ETER)	**EU_fract_country**	Total enrolled students in the country / Total enrolled student in ETER database (without Turkey).
	NAT_UNI_fract (*number*)	Total number of HEIs of university type in the country / Total number of HEIs of any type in the country.
	NAT_UNI_fract	Total enrolled students in the university institutions in the country / Total enrolled student in HEIs of any type in the country.
	NAT_HEI_fract	Total enrolled students in an institution / Total enrolled students in the country.

Quali-quantitative analyses

Characterizing the heterogeneity of HE systems combining bibliometric indicators, higher education data and Research performance based funding (RPBF) information

This section uses a useful classification of European countries according to whether they have research performance-based funding, proposed by Zacherewicz, Reale, Lepori and Jonkers (Science & Public Policy, 2018, Table 1). This classification is available for 25 countries. Hence, the analyses presented in this section relate to institutions in these 25 countries.
The table by Zacherwicz et al. contains the following information on the research funding system. This system includes Bibliometrics (both Publications, Journal impact based measures, and Citations). As regards “Other formula, elements” it includes indicators on PhD graduates, Patents, Project funding, and Business funding. Finally, it takes into account information from Peer review and Performance Contracts.
The classification in their Table does *not* take into account the factor *time*, although the table’s legend gives some additional information about this factor. Funding systems change over time. If a system has been implemented, it takes several years before one can observe any effect at all. Hence, countries that have recently changed their funding system into a performance-based system may not show any effects in the data analysed in this report.
The percentage of top publications (% TOP PUBL in Figure 1) is one of the most frequently used indicators of citation impact. A top publication is a publication of which the citation rate of is among the top 10 percent most frequently cited papers in the subject filed covered by that publication. A country’s percentage of top publications is calculated relative to its total

publication output. The number of graduates per academic staff is an often used measure of the graduation productivity. The two indicators are probably among the best possible measures for citation impact and teaching performance.

Figure 1 shows a scatterplot of these two variables. Moreover, it indicates whether or not a country has a research performance-based research funding (RPBF) system. The category 'Other' in Figure 1 contains three countries for which PBRF-classifications are unavailable: Germany (DE), Liechtenstein (LI) and Serbia (RS). In Germany, institutional funding of universities is mainly provided at the regional level. As allocation procedures differ from state to state, the authors have not assigned a score to the country as a whole. For a fourth country, The Netherlands, the PBRF table indicates a 'limited PBRF', because in this country 'performance contracts' constitute a determinant for institutional funding.

Figure 1 reveals a rather scattered pattern, showing substantial differences among countries, but there is no sign of a statistical correlation between graduation performance or research impact on the one hand, and RPBF on the other. The next section further quantifies this degree of correlation.

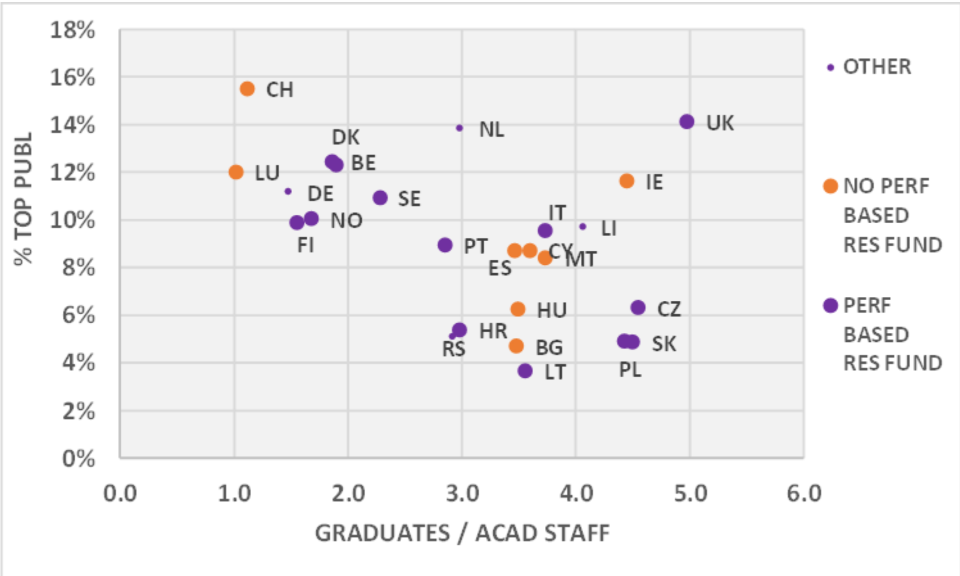

Figure 1. Scatterplot of % Top Publications against Graduates per Academic Staff.

Statistical correlations between 6 key variables

Statistical correlations were calculated pair-wise between the following six indicators:

Total academic staff;	Publications per academic staff;	Percentage of Top Publications;
PhD intensity;	Graduates per academic staff;	Degree of research performance based funding (RPBF);

Total academic staff is size dependent and a good measure of 'size'. The next four indicators are size independent, and measure research intensity, publication and graduation productivity, and citation impact, respectively. The Research performance based funding (RPBF) indicator is derived from Table 1 in Zacherewicz et al. (2018). If this table indicates RPBF, a value of one is assigned; no RPBF corresponds to the value zero. Since there are only nine countries for which data is available both for governance indicators and for the first 5 key indicators, no governance indicators are included in the key set.

Pearson correlations were calculated between each pair of variables. In addition, partial correlations between each pair were calculated, partially out the other four indicators. The number of countries for which data is available for each of the 6 indicators amounts to 25. Table 4 gives results for pairs for which the significance level in at least one of the two computations is above 95 per cent.

Table 4. Statistically significant Pearson and partial correlation coefficients (6 key variables)

Variable 1	*Variable 2*	*Pearson corr.*		*Partial corr.*	
		R	Prob	R	Prob
%Top Publications	Publications per Acad_staff	0.74	0.00	0.82	0.00
%Top Publications	PhD Intensity	0.53	0.01	0.44	0.07
% Top Publications	Total Acad_staff	0.26	0.21	0.61	0.01
%Top Publications	Total grads per Acad_staff	-0.44	0.03	-0.47	0.05
% Top Publications	Research perf. based funds	-0.13	0.58	-0.57	0.01
Research perf. based funds	Publications per Acad_staff	0.12	0.60	0.49	0.04
Research perf. based funds	Total Acad_staff	0.23	0.30	0.48	0.04
Total grads per Acad_staff	PhD Intensity	-0.61	0.00	-0.22	0.39
Total grads. per Acad_staff	Total Acad_staff	0.09	0.66	0.55	0.02

The following observations can be made.

- At the level of countries, citation impact (% Top publications) positively correlates with publication productivity (strongly) and PhD intensity (moderately). It correlates significantly with 'size' (Total academic staff) only if the other factors are partially out. It should be noted that a country's total academic staff is largely determined by demographical factors, for instance, the number of inhabitants.

- Citation impact correlates negatively with graduation productivity. This outcome reveals that, at least at the level of countries, a strong focus on research tends to go hand in hand with a lower graduation performance. It also shows a negative correlation with the degree of research performance based funding – statistically significant only when controlling for the other variables. This outcome is perhaps counter-intuitive. One should keep in mind that the effect of recently implemented RPBF systems may still be invisible in the indicators analysed.

- Apart from its positive correlation with citation impact, PhD intensity correlates negatively with graduation productivity as well (when controlling for the other 4 variables not significant at P=0.05). Figure 2 presents a scatterplot of these two measures. It is hypothesized that this is due to the fact that when a HEI is shifting its orientation towards research, its academic staff puts more efforts in the training of PhD students at the expense of the production of graduate students.

- Interestingly, PhD intensity correlates positively (but weakly) with publication productivity (R=0.35, p=0.09) but their partial correlation is negative (R=-0.25; p=0.34). Since these two correlations are not significant at p=0.05, they are not included in Table 7.

- Apart from the negative correlation with citation impact mentioned above, the degree of research performance based funding (RPBF) correlates positively with publication productivity and total academic staff. But these correlations are only significant if they control for the other four variables in the analysis. The first correlation is in agreement with one would expect to find as effect of RPBF, for the second the current authors do not have an explanation.

- It must be noted that the absolute number of students or graduates is a component in both indicators: it constitutes the denominator in the PhD intensity indicator, and a numerator in the graduation productivity measure. Hence, the indicators are statistically dependent, and a negative correlation between the two is not surprising. This dependence explains the hyperbolic (“f(x)=1/x”-like) left part of the curve in Figure 2.
- Despite the above limitations, and focusing on PhD Intensity, Figures 1 and 2 suggest that substantial differences exist in PhD policies among European countries. The relatively low PhD intensity for Italy and Spain compared to Northern European nations suggests that institutions in these two countries have –at least until recently – given a rather low priority to the foundation of a policy towards the training of PhD students.

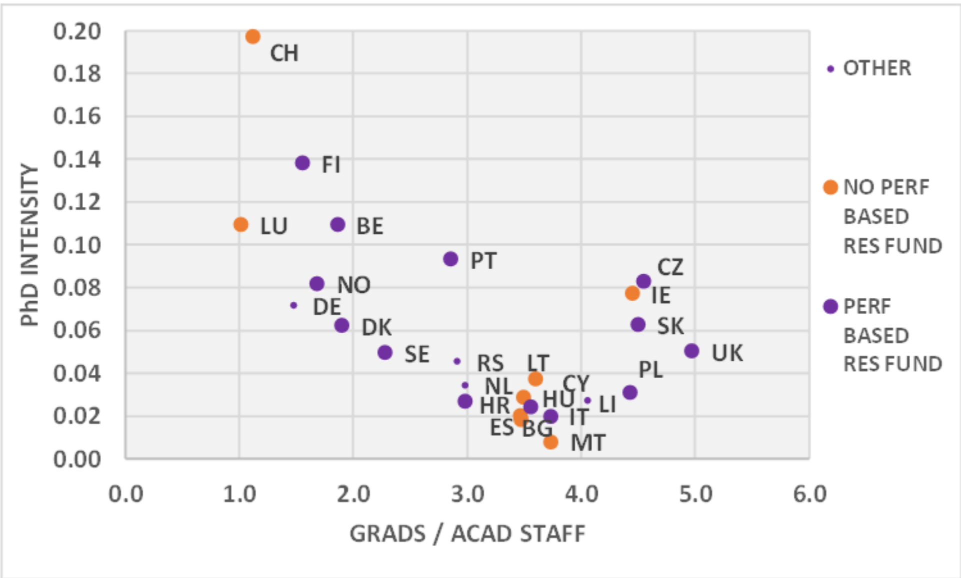

Figure 2. Scatterplot of PhD Intensity against Graduates per Academic Staff

Results from the Cluster and efficiency analyses

The heterogeneity of HEIs exists both across country and within country. Hence, it could be interesting attempt to categorize the HEIs institution regardless their national localization and considering, instead, a specific set of values representing characteristics and performance of each institution with respect to the dimensions of: teaching, research and third mission. The result of such type of analysis could be also used to assess the internal coherence of the national education systems. It could be very helpful to identify institutions to further investigate through case studies.

The variables used to compute the distances for the clusterization are: (i) average publications per academic staff (Pub_fract(av)/ACADstaff; normalized to allow a balanced comparison with the other variable) and (ii) average graduates per academic staff (Grads_ISCED.5-7/ACADstaff). In particular, the results obtained by the K-means (3 clusters) cluster analysis, after the DBSCAN analysis that suggested the existence of three clusters, identify three groups of universities whose main characteristics are outlined in Table 5. We labelled the three groups as: research and teaching oriented (TEAC&RES), research oriented (RES_OR) and teaching oriented (TEAC_OR).

See Figure 3 for an illustration that shows how well the three clusters are spread along the two clustering dimensions. Figure 4 reports the distribution of universities in the three clusters by country.

It appears (see Table 5) that the RES_OR cluster is characterized by the highest number of publications per academic staff (9.57), the highest PhD intensity and the highest proportion of publications in the highly cited journals (0.124), with an average mean normalized citation score (mnsc_(w-av)_av) above the world average (1.16).

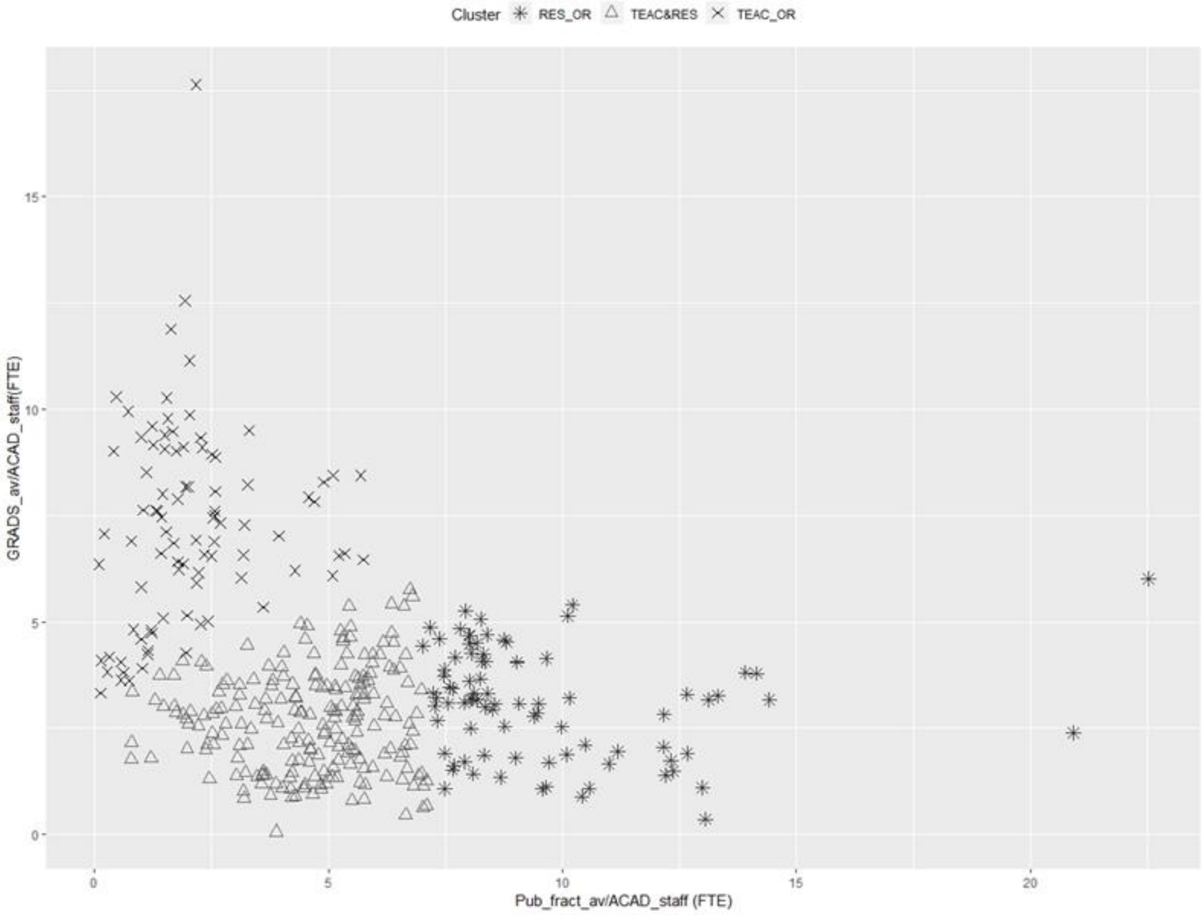

Figure 3. Publications per Acad_staff vs graduates per Acad_staff for the three clusters.

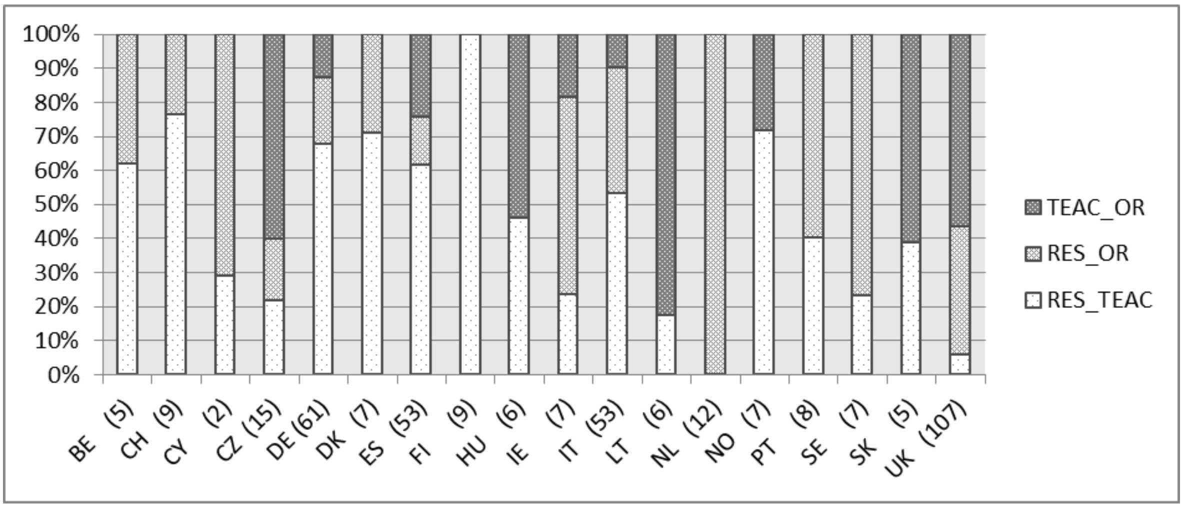

Figure 4. Heterogeneity within countries according to the identified clusters

*** Note: On the X-axe, the *number* in brackets refers to the number of HEIs analysed in each country. Notice that this number ranges from 107 for UK to 2 for Cyprus. Bulgaria, Hungary, Lithuania and Malta were not included because only one observation was available.**

Table 5. Descriptive statistics on the main variables for the obtained three clusters

		TEAC&RES	*RES_OR*	*TEAC_OR*
Cluster analysis variable	Grads_ISCED.5-7/ACADstaff	2.67	3.08	**7.26**
	Pub_fract(av)/ACADstaff	4.61	**9.57**	2.07
Efficiency analysis	Mod.Teach.Res.X_ACADSTAFF.FDH	2.43	**1.67**	**1.64**
Basic data	Foundation_year	1847.84	1785.39	1924.75
	Uni_Hospital	0.531	0.706	0.045
	Enrolled_student_ISCED.5-7	19368.25	21196.18	20143.51
	ACADstaff_FTE	1645.03	1931.45	731.41
	PhD_intensity_2014	0.0652	0.0933	0.0140
	FullProf/ACADstaff_Head	0.1166	0.1491	0.0998
	WomenProff_share	0.1921	0.1943	**0.2760**
	Admn/TOTstaff_FTE	0.4415	0.4797	0.5068
Third mission - Funds	Funds_external%	0.1809	**0.2723**	0.0971
	Funds_third_part/ACADstaff_FTE	30113.47	**60818.98**	23251.66
Specialization	Specialization	0.269	0.261	0.244
Research quantity and quality	Pub_top10(av)/ACADstaff	0.0270	0.0705	0.0105
	Pub_in_top10%	0.0949	**0.1240**	0.0700
	Pub_international_coll	0.5147	**0.5731**	0.4904
	mnsc_(w-av)_av	0.9894	**1.1612**	0.8673
Third mission - Patents	Patent_application(av)/ACADstaff	0.0022	0.0030	0.0008
	Back_citations(av)/ACADstaff	0.0094	**0.0133**	0.0034
	NPL_av	26.76	**43.87**	1.80
	NPL_av/SPA_av	5.63	6.32	2.08
National variables	GOV_Regulation	**8.00**	**6.08**	**4.26**
	GOV_Expenditure	8.81	8.12	8.79
	GOV_Taxes	11.03	11.70	**15.65**
	GOV_Information	**11.62**	9.22	9.85
	GOV_Constraints_trend	0.46	0.49	0.54
	GOV_Opportunities_trend	0.54	0.51	0.46
	GOV_Control_measures	0.33	0.29	0.29
	EU_fract_country	0.0846	0.0917	0.1035
	NAT_HEIs_fract	0.0350	0.0242	0.0187
	NAT_UNI_fract (*number*)	0.5139	**0.5878**	**0.7066**
	NAT_UNI_fract	0.7884	0.8122	0.9198

Interestingly, the RES_OR cluster shows also the highest percentage of funds from third parties (an average of 60,819 euro per academic staff) and the highest intensity of patents per academic staff and patents backward citations, pointing out to the existence of a “Matthew cumulative effect” in place. This means that high quality research is able to attract external funds that are connected to innovative and patenting activities that in turn are self-reinforcing to the scientific activities. On the other hand, we observe that the TEAC_OR cluster is characterized by the production of the highest number of graduates per academic staff (7.26) and presents the highest share of women (0.28) confirming a kind of segregation of women in teaching oriented universities. The TEAC_OR cluster is made, by and large, by institutions

belonging to countries with less regulation policies (GOV_regulation is 4.26 against 6.08 of the RES_OR cluster and 8 of the TEAC&RES cluster) and highest policy interventions on Taxes (GOV_Taxes =15.65, against 11 for the other two clusters). Finally, the TEAC_OR cluster is composed mostly by institutions coming from the biggest countries in Europe (EU_fract_country 0.10) and with the highest proportion of universities on the overall number of HEIs (NAT_UNI_fract (number) =0.71, higher than that of the other two clusters).
The TEAC&RES cluster shows instead intermediary values among the two previously described groups.
Finally, it is interesting to note that the average FDH inefficiency score of the group TEAC&RES (2.43) is higher (*i.e.*, they are less efficient) of the inefficiency scores of the RES_OR and of the TEAC_OR groups (around 1.6). We remind to the reader that an inefficiency score equal to 1 means that the institution is fully efficient, so it is producing its outputs (teaching-graduates and research-publications) being on the efficient frontier of its possibilities. On the other hand, an inefficiency score higher than 1 points out to the possibility of improving the production of its outputs given the available resources (or inputs). This result seems to show that the specialization in teaching and in research pays also in terms of efficiency of the overall activities carried out, that is specialized universities, in teaching or in research, tend to have a higher efficiency than those universities that balance research and teaching activities.

Discussion and conclusions

From the analyses carried out in the present work, a rather heterogeneous picture emerges, that does not allow for 'simple' interpretations and conclusions. The statistical findings seem to be broadly consistent with the following observations.
The outcomes most of all reflect the heterogeneity of the European higher education and research system. Large differences exist between countries. The countries are in different phases of their scientific (and economic) development. During the past decade, in several countries, major changes took place in the funding structure and management of HEI, the effects of which are not yet visible in the analyses presented above. A longer term perspective is certainly needed. Therefore, correlations or concordances between quantitative measures on the one hand and more qualitative indicators (such as governance indicators or degree of research performance based funding) on the other hand are difficult to interpret, as they may relate to different time periods.
The results reveal once more the limits and dangers of *one-dimensional approaches* to the performance of HEIs. Analyses dealing merely with one singe dimension, *e.g.*, either research performance or teaching performance, may easily result in unbalanced or even invalid conclusions. As an example, for the teaching-oriented universities, a key part of their performance remains invisible in a purely bibliometric approach. This is perhaps common knowledge. But universities in the process of expanding their research funding and activities may easily show a declining graduation productivity (graduates per academic staff) if an increase in the size of their academic staff is deployed in research, while research output will increase with a delay of several years.
Apart from funding formula, another important aspect of a national HE system is the degree and the modus of quality assessment of research and education. For instance, in the Netherlands, assessment exercises by research discipline (e.g, Physics, Chemistry, Biology) have been conducted every 4–5 years for at least 25 years. Even though the outcomes do not play a formal role in the allocation of government funding of HEI, they do play a role in internal assessment and management processes within HEIs. The prominent position of The Netherlands in several analyses presented above may be at least partly a result of these long lasting and intensive assessment practices.

The combined efficiency analysis and cluster exercises showed the existence of three groups of European universities clearly characterized in their orientation towards teaching activities (TEAC_OR), research activities (RES_OR) or balancing among the two activities (TEAC&RES). Interestingly, the universities specialized in teaching or research show on average a higher efficiency in their main purpose then those oriented to the production of both teaching and research activities.
The obtained results may be useful to identify (select) the HEIs that need to be further investigated through case studies. In this way, our results may provide an evidence-based support to further investigate the heterogeneity of HE systems through qualitative case studies in political science studies of HE.

Acknowledgments

The financial support of the Italian Ministry of Education and Research (through the PRIN Project N. 2015RJARX7), of Sapienza University of Rome (through the Sapienza Awards no. PH11715C8239C105) and of the EU Horizon 2020 RISIS2 Project (grant agreement N° 824091) is gratefully acknowledged. We warmly thank Ed Noyons for providing us the CWTS bibliometric data analysed in this work.

References

Capano, G. & Pritoni, A. (2019). Varieties of hybrid systemic governance in European Higher Education, *Higher Education Quarterly*, vol. 73, n.1, pp. 10–28.

Daraio C. (2018). *Nonparametric Methods and Higher Education*, in: Teixeira P., Shin J. (eds) Encyclopedia of International Higher Education Systems and Institutions. Springer, Dordrecht.

Daraio C. (2019). *Econometric approaches to the measurement of research productivity*, in Springer Handbook of Science and Technology Indicators edited by Glänzel W., Moed H.F., Schmoch H. and Thelwall M., forthcoming.

Daraio C. et al. (2011). The European University landscape: A micro characterization based on evidence from the Aquameth project, *Research Policy*, vol. 40, pp. 148–164.

Daraio C. & Simar L. (2007). *Advanced Robust and Nonparametric Methods in Efficiency Analysis, Methodology and Applications*, Springer, New York (USA).

Daraio, C. & Bonaccorsi, A. (2017). Beyond university rankings? Generating new indicators on universities by linking data in open platforms, *Journal of the Association for Information Science and Technology*, vol. 68, n.2, pp. 508–529.

Daraio, C., Heitor, M., Meoli, M. & Paleari, S. (2019). Policy turnaround: Towards a new deal for research and higher education. Governance, evaluation and rankings in the big data era, *Higher Education Quarterly*, vol. 73, n. 1, pp. 3–9.

Ester, M., Kriegel, H.P., Sander, J. & Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, In *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, Portland, OR, AAAI Press, pp. 226–231.

López-Illescas, C., de Moya-Anegón, F. & Moed, H. F. (2011). A ranking of universities should account for differences in their disciplinary specialization, *Scientometrics*, vol. 88 n 2, pp. 563-574.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations, in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, n. 14, pp. 281–297.

Zacharewicz, T., Lepori, B., Reale, E., & Jonkers, K. (2018). Performance-based research funding in EU Member States—a comparative assessment, in *Science and Public Policy*, pp. 1–11 (doi: 10.1093/scipol/scy041).