




Studying the heterogeneity of European higher education institutions

Renato Bruni¹ · Giuseppe Catalano¹ · Cinzia Daraio¹ · Martina Gregori^{1,2}  · Henk F. Moed¹

Received: 18 December 2019 / Accepted: 10 September 2020 / Published online: 29 October 2020
© The Author(s) 2020

Abstract

The heterogeneity of the Higher Education (HE) Institutions is one of the main critical issues in the assessment of their performance. This paper adopts a multi-level and multi-dimensional perspective, combining national (macro) and institution (micro) level data, and measuring both research and teaching activity, using performance indicators derived from the European Tertiary Education Register, CWTS Leiden Ranking, and PATSTAT patent database. Clustering and efficiency analysis are combined to characterize the heterogeneity of national HE systems in European countries, and reveal the potential of using micro level data to characterize national level performance. Large differences are observed between the European countries, partially due to the fact that they are in different phases of their scientific (and economic) development and of the re-structuring of their HE systems. Evidence is found that universities specializing either in teaching or in research tend to have a higher efficiency than those institutions balancing research and teaching. Tradeoffs are observed between undergraduate and post-graduate activities, and a “Matthew cumulative effect” seems in place on the European institutions analysed: high quality research is able to attract external funds that stimulate innovative and patenting activities that in turn are self-reinforcing to the scientific activities. The results reveal once more the limits and dangers of one-dimensional approaches to the performance of HEIs.

Keywords University · Europe · Heterogeneity · Clustering · Efficiency analysis

Introduction

The measurement of academic performance is a relevant issue at the intersection between political science and informetrics. Numerous international comparisons (rankings) of Higher Education Institutions (HEIs) are regularly published—such as Shanghai, Times

✉ Martina Gregori
martina.gregori@uniroma1.it

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

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

Higher Education and Leiden Ranking; nonetheless, HEIs performance analysis still represents a challenging task. Higher education systems are complex, characterized by multi-levels (course, institution, nation, etc.), multi-objectives (i.e. teaching, research, third mission activities) and heterogeneity.

Heterogeneity is one of the main critical issues to address in any benchmarking analysis. The comparability of the units of analysis is a necessary condition for any meaningful relative assessment or quantitative evaluation. Attention to the topic has been dedicated for long from both scholars and policy-makers; nevertheless, the diversity in higher education systems results to be difficult to tackle, a general conceptualization is still lacking (Huisman et al. 2015) and the empirical analysis of the related literature seems to lead to contradictory outcomes (Barbato and Turri 2019).

The choice of the most salient dimensions of heterogeneity is still controversial. Multiple sources are associated to heterogeneity, including the national context, the HEIs mission, the presence or absence of medical schools, the institutions' legal status and the adopted disciplinary orientation and degree of specialization (López-Illescas et al. 2011; Daraio et al. 2011). The dimension of internationalization has also been considered in recent studies, with nations that become increasingly interdependent and internationalization missions that are currently embodied in universities' strategies (Huisman et al. 2015). Differences in performance outcomes could also originate from different levels of autonomy and/or competitiveness experimented by universities (Aghion et al. 2010) and the economic development of their contexts (with more influence on research-related activities, rather than the teaching one; Agasisti and Bertolotti 2019).

One-dimensional approaches to the HEIs performance evaluation entangle the risk of potentially unbalanced or even invalid conclusions, forcing a homogeneous vision of success/failure, mission, characteristics. The literature is moving towards more complex methodological approaches, trying to include progressively a more *multi-dimensional* perspective; including investigations on how, and to which extent, elements of heterogeneity influence performance. It should be taken into account that, due to the heterogeneity and multi-dimensionality, a real overall valid HEIs "classification" is difficult to obtain.

Bonaccorsi and Daraio (2009) is one of the first attempts in analysing extensive data from different European countries with the aim of tackling their heterogeneity. Using a database from the AQUAMETH project,¹ they identified through cluster analysis different performance profiles across European countries, relating them to different strategic orientation adopted by the single institution (research oriented, teaching oriented, multi-purposes). Similar results were obtained by García-Aracil and Palomares-Montero (2012) and de la Torre et al. (2018), both with respect to the Spanish higher education system. The former, applying a cluster analysis, identified 3 groups: research-oriented universities, teaching oriented universities and Knowledge Transfer (KT)-oriented universities. The latter, applying a so-called DEA-MDS multidimensional analysis, identified 6 groups: universities oriented towards efficiency in the traditional missions (particularly teaching), universities oriented towards efficiency in research, universities oriented towards the efficiency in the traditional missions, universities oriented towards overall efficiency, universities oriented towards efficiency in KT, regional universities oriented towards efficiency in research and KT. The results obtained in the present work

¹ Aquameth stands for advanced quantitative methods for the evaluation of the performance of public sector research. The Aquameth project has been coordinated by Andrea Bonaccorsi and Cinzia Daraio. For more information on the dataset, see Bonaccorsi and Daraio (2007).

confirm the same line of categorization, working with extended database and more dimensions.

Daraio et al. (2011) provide an investigation on the identification of the heterogeneity, considering *horizontal* heterogeneity (i.e. decisions on subject mix, target audience, teaching methodologies, type of research, type of third mission's activities etc.) and *vertical* heterogeneity (i.e. *positioning* of the university in a hierarchy of quality of university service provision).

Catalano et al. (2017), focusing on the sources of heterogeneity induced by the *subject mix* of HEIs, and using the ETER database, propose to estimate “scale parameters” representing European students in different fields of education (namely, Engineering, Medicine, Natural and Physical Sciences, Social Sciences and Humanities), as tool to be able to compare educational production across different fields on a common ground.

Similarly, Zharova et al. (2017), using micro-level data on publications and citations (Scopus) over selected HEIs in Germany, identify differences across research fields over (i) the relationships between funding volume and research productivity and number of citations; (ii) the influence of past research results on likeliness to obtain external funding; (iii) the optimal answer to exogenous changes. The evaluation of research performance by disaggregating the disciplinary fields to low levels is also proposed by Bonaccorsi and Secondi (2017), that shows how research performance depends on variables at the level of university (e.g. size, teaching, governance) and the level of external regional environment (general effects—*level of development of the region, expenditure in R&D and technological intensity of the manufacturing sector*; specific effect—*variables used to sizing the health sector*).

Finally, Barbato and Turri (2019) compare two European countries, namely England and Italy, by considering different dimensions (core functions, subject mix, market size, structural information). Institutional positioning has been defined by Fumasoli and Huisman (2013) as the process through which HEIs locate themselves in specific niches within the HE system, reflecting the activities, resources (e.g. financial, human) and potential relations (competition, cooperation) that they assume to prosper in their system. Barbato and Turri (2019) identify two main approaches in positioning: more or less passive adaptation in the direction indicated by context external forces, and deliberate or emergent strategy. Institutional pressure (government regulation) and competition (students, researchers, funds, reputation) are the two main important external forces that impact on HEIs. The analysis results indicate a more differentiated system in England, while in general both Italian and English HEIs are becoming increasingly homogeneous in terms of research intensity, and increasingly more heterogeneous in terms of internationalisation.

In this context, it would be important for the research community and the policy maker, to be able to understand how heterogeneity would be tackled and how heterogeneity can influence the performance. The present work adopts a multi-level perspective by combining national (macro) level data and institution (micro) level data and analyses; also showing the potential in using micro-level data to characterize the national level performance. We consider a *systemic* perspective, integrating heterogeneous sources of available data, covering all the three dimensions of HEIs production process (namely, teaching, research and third mission) and including information on the national regulation measures introduced over time.

The current paper's objective is to characterize HEIs while accounting for the following aspects:

- (i) *Structural* heterogeneity (structure of the national system: systemic factors, e.g. number and types of HEIs that are involved, governance factors);
- (ii) *Internal* heterogeneity (linked to the type of the production process carried out within the HEIs);
- (iii) *Other* heterogeneity sources.

The analysis is focused on the *European* context. The European HEIs have been proved to be less performing if compared with their US counterparts (Aghion et al. 2010), making crucial to create tools useful for improvement. The US higher education system is characterized by significantly higher resources and a clear distinction between education-oriented institutions and doctoral universities, associated to overall higher volume of publications and citations with respect to revenues (Lepori et al. 2019).

Regulation settings, traditions, economic development contexts highly varying, substantially influence the level of heterogeneity between and within countries (Bonaccorsi 2014). The modernisation agenda for Higher Education in Europe (European Commission 2016) identifies the relevance of creating effective governance and funding mechanisms for higher education. Different models of governance (Agasisti and Catalano 2006; Capano et al. 2015) are applied by policy makers trying to improve the systemic performance of Higher Education, resulting at a European level in designs that represent each country's proper interpretation of a common template. After 30 years of adaptations, three systemic governance factors seem to have emerged (Capano and Pritoni. 2019): a performance-based mode, a re-regulated mode and a systemic goal-oriented mode.

Finally, reliable data recently started to be available, thanks to important advancement in data collection and data processing procedures and the activation of specific research projects with the aim of creating broad databases, with good coverage on different countries and different years (i.e. AQUAMETH—see Daraio et al. 2011, EUMIDA—see Bonaccorsi 2014, ETER²).

This work presents 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 third mission. 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,³ with information on innovation activities from PATSTAT database and with categorizations of national higher education policies obtained from more qualitative studies of national HEI systems. Notice that, in our analysis a series of variables associated with patents activities and funds composition are used as proxies for the third mission activities.

The third mission refers to the economic and social impacts generated by HEIs' activities through interactions with the external stakeholders, aiming to generate, apply and exploit knowledge (Secundo et al. 2017). Nevertheless, the concept still lacks a specific and unambiguous definition, mainly due to its dependence on contextual circumstances (Pineiro et al. 2015). This unclearness contributes to a critical data availability issue, along with the difficulties associated with output quantification and measurement, especially with

² The ETER data can be downloaded from the project website at: <https://eter-project.com/#/home> (last accessed on 17 October 2019).

³ The Leiden Ranking can be accessed at: <https://www.leidenranking.com/> (last accessed on 2 December 2019).

regard to the societal dimensions. According to the available data, it was then decided to represent the third mission only through its innovation side, following well-established approaches in the literature.

In the project the existing problems of data availability, quantification and comparability go hand in hand with the need for 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 structure of this paper is as follows. “Method” Section gives an overview of the methods that were applied in the study, and “Data” Section a detailed description of the data sources. The results are presented in three parts. As an introduction to the analyses, “Basic information on national higher education systems” Section outlines the set of higher educational institutions analysed in the paper and gives a characterisation of national HE systems in terms of governance structures. “Results from the Cluster and efficiency analyses” Section presents the outcomes of a cluster analysis of higher education institutions based on their similarity in terms of their bibliometric and governance characteristics, and focuses on the notion of efficiency of the institutions across clusters. Next, “Additional methodological approaches and case studies” Section illustrates two additional studies that represent lines of future research, aimed to further broaden the insights into the performance of higher education institutions and into the factors that influence this performance. The first relates to the methodology to identify clusters, and the second to case studies providing a detailed comparison of particular countries. Finally, “Discussion and conclusions” Section summarises the results and makes more suggestions for further research.

The main objective of this work is to characterize the heterogeneity of HE systems (at a country level and systemic level), exploiting micro-level data and making use of a multi-methodological approach (combining qualitative exploratory analysis, clustering analysis and efficiency evaluation). Beside this multi-methodology, another novelty of the paper is represented by the analyzed sample, in terms of covered countries and data quality.

Some of the results obtained by our analysis confirm previous literature by using a more complete database, adopting more recent and comparable data.

Methods

This work uses the multi-methodology approach summarized in Fig. 1. The main pillars are a descriptive analysis carried out at a country level (quali-quantitative analysis) and a cluster analysis, which includes as a variable the inefficiency scores calculated to assess the ability of higher education institutions to perform (both) teaching and research activities. By combining and interpreting the outcome of both analyses we characterize the heterogeneity of European HEIs identifying three main typologies.

The identification of possible HEIs typologies is usually carried out in the literature following either an expert base approach (subjective selection of classification criteria and threshold values), or mathematical approaches, among which the most common applied are Data Envelopment Analysis (DEA) and Cluster analysis. DEA analysis has been recently applied for example to the cases of Russia and Spain higher education systems (Abankina et al. 2016; de la Torre et al. 2018). Cluster analysis has already been used to tackle the

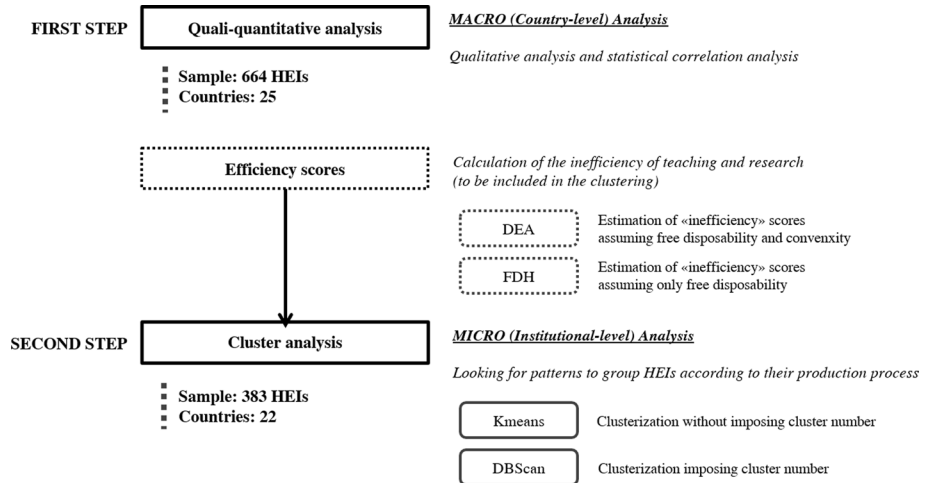


Fig. 1 Tackling the heterogeneity of European HEIs: an outline of our approach

higher education heterogeneity, for example, in China, Italy, Spain and EU country selections (Bonaccorsi and Daraio 2009; Rossi 2010; García-Aracil and Palomares-Montero 2012; Wang and Zha 2018; Barbato and Turri 2019).

DEA (Charnes, Cooper and Rhodes 1978) is a *nonparametric* approach to carry out efficiency analysis. It does not require many assumptions and particularly it does not need the specification of a functional form for the benchmarking frontier against which the performance of the units is assessed. DEA assumes the free disposability (that is the possibility to destroy goods without any cost) and the convexity of the feasible production set (the set of all possible values over which the benchmarking comparison is made). On the other hand, the Free Disposal Hull (FDH, Deprins et al. 1984), another popular nonparametric efficient estimator of the frontier, is based only on the free disposability and does not rely on the convexity assumption. DEA uses mathematical programming techniques to evaluate the performance of peer units in terms of multiple performance metrics/measures/features. These peer units are called Decision Making Units (DMUs). DEA and FDH are among the most applied nonparametric techniques for the measurement of the efficiency in production and service activities. These are the methods used in the second section of this work to calculate a nonparametric efficiency estimation of teaching and research. More robust nonparametric estimation (Daraio and Simar 2007) could be introduced in the next studies, focusing on this specific aspect and is left for future works.

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 in advance.

DBSCAN, on the other hand, is a more recent clustering technique, but is 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 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 leads 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 consider indicators related to all the different activities carried out in the academic operations, namely teaching, academic research and third mission activities (here approximated by innovation activities; e.g. collaboration with industries, patents, etc.).

With the purpose of gathering information about the three aforementioned areas, different sources were used. In particular, the following three databases were integrated for the analysis at the micro-level (single institution):

ETER database, *for the information at the micro level (single institution) regarding the teaching area;*

CWTS Leiden Ranking database, *for the information regarding the academic research (based on extensive cleaning of the data from the Web of Science);*

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

Table 1 reports the main considered variables. Besides, a database dedicated to the national regulatory characteristics of European countries was integrated, to account for a part of the heterogeneity among higher education macro-systems. The considered governance indicators (reported in Table 2) are based on Capano and Pritoni's (2019) data, covering the period 1988–2014 and considering 12 European countries. These governance indicators are in total 24, grouped in 4 dimensions (Regulation, Expenditure, Taxes, Information), and form a policy instrument repertoire. All the pieces of relevant legislation on Higher Education (e.g. laws, decrees, circulars and ministerial regulations) occurred over the observed period in each country, have been classified over the repertoire through an expert based process, that assigned each piece to a specific indicator according to its main issue nature.

Few data transformation were applied in order to include the governance data in our analysis. As first group of variables, *normalized governance structure indicators* were calculated, to give information on the national governance approach to HEIs legislation, normalized to the all sample observations. Each country has 4 variables; each variable represents the percentage of the total pieces of legislation applied on a dimension (Regulation, Expenditure, Taxes, Information) that occurred in the country. As second group of variables on *internal governance structure indicators* were calculated, retrieving them directly

Table 1 Definition and source of variables at micro-level

Category (Source)	Variables	Definition	
Cluster analysis variables (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	
Inefficiency scores	FDH	FDH Inefficiency scores. Are based on the free disposability assumption. The obtained values may be higher than, or equal to 1. It is 1 for efficient units, higher than 1 for units that can expand the production of their outputs given their inputs	
	DEA	DEA Inefficiency scores. Are based on the free disposability and on the convexity assumptions. The obtained values may be higher than, or equal to 1. It is 1 for efficient units, higher than 1 for units that can expand the production of their outputs given their inputs	
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	
	WomenProf_share	Percentage of women on the total number of professors	
	Admn/TOIstaff_FTE	Percentage of administrative staff on total staff (academic plus administrative)	
	Specialization (ETER)	Specialization	Expresses 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

Table 1 (continued)

Category (Source)	Variables	Definition
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)
	mnsr_(w-av)_av	Papers mean normalized citations (yearly average, weighted by the number of patent applications)
Third mission (Funds ETER)	Funds_external%	Percentage of funding from third parties on total funding
	Funds_third_part/ACADstaff_FTE	Third party funds (coming from public and private sources) per academic staff (expressed in FTE)
	Patent_application(av)/ACADstaff	Overall total number of patent applications (yearly average)
Third mission (PATENTS, ETER for Acad_staff)	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 2 Definition and sources of variables at macro-level

<i>Governance (Capano and Pritoni 2019)</i>	
GOV_Regulation	Percentage of policy intervention on Regulation, normalized and internal [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, normalized and internal [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, normalized and internal [tax exemption; tax reduction for particular categories of students; service-based student fees]
GOV_Information	Percentage of policy intervention on Information, normalized and internal [transparency; certification; monitoring and reporting]
GOV_Cons_trend	In each country, percentage of regulatory interventions aimed to add <i>more constraints</i> respect to the overall regulatory interventions in Regulation (full explanation in paragraph 4.3)
GOV_Opp_trend	In each country, percentage of regulatory interventions aimed to add <i>more opportunities</i> respect to the overall regulatory interventions in Regulation (full explanation in paragraph 4.3)
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 (full explanation in paragraph 4.3)
<i>System structure (ETER)</i>	
EU_fract_country	Total enrolled students in the country/Total enrolled student in ETER database (without Turkey)
NAT_UNI_fract (<i>number</i>)	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

from Capano and Pritoni (2019). Each country has 4 variables; each variable represents the percentage of pieces of legislation applied on a dimension, over all the pieces of legislation applied in the country. Governance variables implications should be considered carefully, taking into account the qualitative methodology used for the data collection, and the necessary approximation associated to the pieces of legislation categorization. It would not be robust to base quantitative analysis solely on this type of variables; nonetheless, they can be fruitfully used to tackle the general national approaches towards HEIs legislation and to outline possible relations with performance results.

With respect to the multi-dimensional activities data in the final dataset (Table 1), it was necessary to structurally internalise temporal lags between inputs and outputs information. It is well known that a certain time must pass to observe effects related to the interventions on academic staff, academic funds, and so on. As it is usually done in the empirical analyses, a 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, PATSTAT database (patent and international collaborations).

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

The final dataset contains the average variables values over the considered period of each included database; missing values had been excluded from the calculation. In order to better highlight the role played in heterogeneity by the considered factor, we decided to shrink the analysis perimeter only to HEIs categorized as “universities” in the ETER dataset, excluding universities of applied sciences and other institutions, highly differentiated among countries in the sample according to the national code for labelling.

Moreover, for data cleaning purposes, we included those universities whose data are available both on staff, students, graduates, on publications and citations in the Leiden Ranking dataset, on patents submissions and citations performance in the PATSTAT database. The total number of selected institutions for all ETER countries combined amounts to 664. Nevertheless, due to the presence of missing values on key variables (namely, academic staff and number of enrolled students) concerning the cluster analysis procedure, the quantitative analysis was performed only on a sub-selection of the database composed by 383 HEIs from 22 countries (Bulgaria, Croatia, Luxembourg and Malta are represented in the sample only by one higher education institution).

Basic information on national higher education systems

Number of HEIs per country

Table 3 presents per country the number of higher education institutions (HEIs) and the number of total enrolled students (ISCED 5-7) included, respectively, in the analyses presented below in “Results from the Cluster and efficiency analyses” Section, and in the ETER database. Table 3 is constructed making reference to the last year of currently available data which is 2016.

The last two columns of Table 3 show that the coverage of our sample considering the total number of enrolled students at all levels is higher than the coverage measured by the number of higher education institutions. Hence, our sample can reasonably be considered representative of the European university systems.

Basic data on governance orientation

The database presented by Capano and Pritoni (2019) allows important analyses with regard the regulation approaches adopted by European countries; in particular, it could be attempted to connect general tendency and time evolution, with specific results in performance (e.g. increases/decreases in efficiency). This type of study is beyond this paper objectives; nevertheless, a first analysis on the database is useful to give insights into the internal regulatory structure adopted by each country. As stated in “Data” Section, even if

Table 3 Number of HEIs classified as “university” per country, considered in this paper and in the ETER database

Country	Country code	Nr HEIs in our sample	Total Nr HEIs in ETER database	% HEIs included in our sample (%)	Total Nr Enrolled Students ISCED 5–7 available in ETER database	% Enrolled Students ISCED 5–7 included in our sample *(number of HEIs with missing data)
Belgium	BE	5	12	42	34,299	99% *(6)
Switzerland	CH	9	12	75	32,036	89%
Cyprus	CY	2	8	25	7572	29%
Czech Republic	CZ	15	29	52	68,409	75%
Germany	DE	61	103	59	261,110	74% *(6)
Denmark	DK	7	10	70	47,521	87% *(2)
Spain	ES	53	82	65	297,981	76%
Finland	FI	9	15	60	29,869	80%
Hungary	HU	6	25	24	52,658	41%
Ireland	IE	7	7	100	36,300	100%
Italy	IT	53	97	55	354,175	81% *(5)
Lithuania	LT	6	20	30	19,577	59% *(3)
Netherlands	NL	12	19	63	77,999	94% *(1)
Norway	NO	7	8	88	30,112	91%
Portugal	PT	8	38	21	51,059	49% *(1)
Sweden	SE	7	29	24	61,065	41%
Slovakia	SK	5	18	28	42,461	44%
United Kingdom	UK	107	128	84	702,755	91%

Bulgaria, Croatia, Luxembourg and Malta were not included in the table because only one observation was available, for each one of them

Table 4 National systems regulation structure and comparison with the typical behaviour. *Source:* values from Capano and Pritoni’s (2019) indicators and our elaborations

1989-2014 Country	Country level percentage				Differences from average			
	Regulation	Expenditure	Taxes	Information	Regulation	Expenditure	Taxes	Information
GR	0.697	0.224	0	0.079	↑ 0.081	↑ 0.002	↓ -0.049	↓ -0.009
SE	0.702	0.226	0	0.071	↑ 0.086	↑ 0.004	↓ -0.049	↓ -0.017
DK	0.662	0.230	0.014	0.095	↑ 0.046	↑ 0.008	↓ -0.035	↑ 0.007
NL	0.691	0.235	0.015	0.059	↑ 0.075	↑ 0.013	↓ -0.034	↓ -0.029
NO	0.631	0.252	0.019	0.097	↓ 0.015	↑ 0.030	↓ -0.030	↑ 0.009
IT	0.585	0.202	0.053	0.160	↓ -0.031	↓ -0.020	↑ 0.004	↑ 0.072
AT	0.600	0.160	0.053	0.187	↓ -0.016	↓ -0.062	↑ 0.004	↑ 0.099
FR	0.681	0.167	0.069	0.083	↑ 0.065	↓ -0.055	↑ 0.020	↓ -0.005
FI	0.535	0.349	0.070	0.047	↓ -0.081	↑ 0.127	↑ 0.021	↓ -0.041
IE	0.623	0.170	0.075	0.132	↑ 0.007	↓ -0.052	↑ 0.026	↑ 0.044
England	0.452	0.290	0.129	0.129	↓ -0.164	↑ 0.068	↑ 0.080	↑ 0.041
PT	0.585	0.226	0.151	0.038	↓ -0.031	↑ 0.004	↑ 0.102	↓ -0.050
Typical behaviour (Geometric Average)	0.616	0.222	0.049	0.088				

We highlighted (using bold and grey colour for the cells) the values for which the differences from average are the highest, in both direction (bigger, smaller)

the governance variables inherent nature does not allow the application of strictly quantitative methodologies, it is still possible to exploit information for comparing national patterns. Table 4 displays the within-country percentages for each regulatory macro-dimension of the observed 12 countries.

The data reveal an interesting similarity between the Greek regulatory scheme and that of a majority of the Scandinavian-northern countries (in particular Sweden; with less extent Denmark and Norway), with a large attention dedicated to general regulation and few interventions on taxes. On the other hand, an opposite regulatory scheme seems to be adopted by England, Portugal and, less strongly, Finland, which is also the country with the relatively most numerous interventions on the expenditure dimension. As reported already in Capano and Pritoni (2019), Austria and Italy show the largest interest in information interventions.

In general, relations can be observed between interventions on regulation on the one hand, and those on taxes (slightly negative), and on expenses (slightly positive; also in Capano and Pritoni 2019), on the other. By contrast, France, Finland and England seem to present opposite concordance between regulation and expenditure.

A second type of considerations can be carried out with respect to the more or less coercive policies applied at national level across European countries. Capano and Pritoni (2019) already include an analysis along this line, involving all the 4 governance dimensions. In our paper, we decided to repeat the analysis focusing exclusively on Regulation Indicators. We propose a new variable trying to capture the level of formal control applied over HEIs activities in each country. Regulation is the governance dimension interested by the higher number of pieces of legislation (around 50% for any considered country). It seems also to be the dimension interested by the highest heterogeneity (as is will be shown by cluster analysis results; in “Results from the cluster and efficiency analyses” Section). The tendency is assessed to either limit or encourage autonomous decisions and acts by the HEIs on four main areas of operative activities: assessment, academic career, courses,

Table 5 Governance tendency to adopt control measures on HEIs with respect to the tendency to either limit or encourage HEIs autonomy on Regulation Dimension

	%Constraints	%Opportunities	%Control measures
POR	0.21	0.47	0.40
ITA	0.13	0.58	0.35
SWE	0.37	0.32	0.35
NOR	0.33	0.45	0.34
AUS	0.27	0.40	0.33
IRE	0.40	0.42	0.33
DEN	0.36	0.32	0.31
FRA	0.38	0.31	0.30
ENG	0.33	0.48	0.28
FIN	0.18	0.50	0.24
NED	0.17	0.46	0.21
GRE	0.22	0.45	0.17
Total	0.27	0.44	0.30

administration. The government interventions on the subjects Content of curricula, Academic career, Regulation on students and Institutional and administrative governance have been registered by Capano and Pritoni (2019) in separate indicators according to the regulation aim, namely to generate more constraints or more opportunities for the HEIs; this separation is used to calculate the variables reported in Table 5. A proxy on “control measures” adopted by government is calculated, as reported in Table 2 in “Data” Section, based on the number of interventions registered on the areas of rules on goals in teaching, of assessment, and of monitoring and reporting. Table 5 seems to reveal that control measures are adopted especially in higher education systems that tend to be coercive in the regulatory dimension. In contrast, in Portugal, Italy and Norway the control measures could compensate for approaches towards a greater autonomy of the single institutions. This observations are strongly related to the regulatory level; considering all the 4 dimensions the country categorization could be different. For example, overall Italy it can be considered more oriented towards the apply constraints approach (Capano and Pritoni 2019).

Results from the cluster and efficiency analyses

The heterogeneity of HEIs exists both *across* and *within* countries. Hence, HEIs institutions are categorized regardless their national location, by considering, instead, a specific set of values representing characteristics and performances of each institution with respect to the dimensions of teaching, research and third mission. The result of such type of analysis can be also used to assess the internal coherence of the national education systems and to conduct in depth investigation on case studies.

The structure of this section is as follows. In “*K*-means cluster analysis (three clusters)” a cluster analysis is presented, based on *K*-Means clustering, fixing the number of clusters to three, thanks to a preliminary examination through DBSCAN application. The main characteristics of each cluster are being described. Next, “Efficiency analysis comparing the

identified clusters” Section focuses on one particular, most relevant aspect: efficiency. It compares the three clusters on the basis of two advanced inefficiency scores, DEA and FDH, outlined in “Methods” Section.

K-means cluster analysis (three clusters)

The variables used to compute the distances for the clusterization are: (i) average publications per academic staff ($\text{Pub_fract}(\text{av})/\text{ACADstaff}$; normalized to allow a balanced comparison with the other variables) and (ii) average graduates per academic staff ($\text{Grads_ISCED.5-7}/\text{ACADstaff}$).

In a first step, a DBSCAN analysis was conducted. Density-based spatial clustering of applications with noise (DBSCAN) is a density-based clustering non-parametric algorithm. It groups together data points that are closely packed together, identifying dense regions without the need to specify the number of clusters in advance. DBSCAN suggested the existence of three clusters in our dataset, which were then actually created by using *K*-means. *K*-means was preferred because the clusters produced by this algorithm are more internally homogeneous. However, *K*-means needs the number of clusters to be specified in advance. Since this is a crucial choice, algorithm DBSCAN was used to identify the number of underlying dense regions in our dataset and suggest it as the number of clusters in *K*-means.

The three groups of higher education institutions in these clusters are presented in Table 6. They are labelled as: research and teaching oriented (TEAC&RES), research oriented (RES_OR) and teaching oriented (TEAC_OR). It must be noted that the calculations include only institutions for which data are available for *all* the analysed variables. Therefore, if for some country no data are available on academic staff, all its institutions were discarded. Furthermore, it should be noted that, since the analysed HEIs all appear in the Leiden Ranking, they all have at least a minimum level of research activity. Figure 2 illustrates how well the three clusters are spread along the two clustering dimensions. In particular, it shows that the RES-TEACH cluster contains institutions that have both a relatively low publication productivity, compared to those in the RES_OR cluster, and a low graduation productivity, compared with the institutions in the TEACH_OR cluster.

Figure 3 clearly reveals large differences among national academic systems with respect to the distribution of institutions across the categories research oriented, teaching oriented and both research and teaching oriented institutions. The TEAC&RES cluster contains institutions from several Scandinavian countries (with a high scientific performance), Switzerland and—with a smaller percentage of national institutions, but always higher than 50%—Germany, Belgium, Spain, Italy, Hungary, Portugal and Slovakia. Around than the 50% of the national institutions of Lithuania, Czech Republic and UK are included in the TEAC_OR cluster; Slovakia and Hungary follows with smaller percentages. The RES_OR institutions percentage is, instead, particularly high for the Netherlands, Sweden, Portugal and Ireland.

Finland, Switzerland and the Netherlands seem to have uniform national systems, even though they are not in the same cluster: the first two are focused on joint teaching and research activities, whereas the latter one is more dedicated to research activities.

Nevertheless, the choice of analysing only universities (excluding universities of applied sciences and other institutions) is suitable for the analysis of research output and productivity, but might provide some biased results for the system-level analysis. The fact that countries, such as Switzerland and Netherland, appear very homogeneous could be due to

Table 6 Descriptive statistics on the main variables for the obtained three clusters

	TEAC&RES	RES_OR	TEAC_OR
<i>Cluster analysis variables</i>			
Grads_ISCED.5–7/ACADstaff	2.67	3.08	7.26
Pub_fract(av)/ACADstaff	4.61	9.57	2.07
<i>Inefficiency scores</i>			
DEA	3.72	2.14	2.44
FDH	2.43	1.67	1.64
<i>Basic data</i>			
Foundation_year	1847.84	1785.39	1924.75
Uni_Hospital	0.531	0.706	0.045
Enrolled_student_ISCED.5–7	19,368.25	21,196.18	20,143.51
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
WomenProf_share	0.1921	0.1943	0.2760
Admn/TOTstaff_FTE	0.4415	0.4797	0.5068
<i>Specialization</i>			
Specialization	0.269	0.261	0.244
<i>Research quantity and quality</i>			
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
mns_c_(w-av)_av	0.9894	1.1612	0.8673
<i>Third mission—Funds</i>			
Funds_external%	0.1809	0.2723	0.0971
Funds_third_part/ACADstaff_FTE	30,113.47	60,818.98	23,251.66
<i>Third mission—Patents</i>			
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
<i>National variables</i>			
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

the fact that the teaching-oriented HEIs are outside the university sector (unlike UK and Italy). Hence, the delineation of the perimeter might affect results of the comparisons and analyses in terms of efficiency presented below.

Czech Republic and UK show the lowest percentage of institutions balancing teaching and research (in TEACH&RES), and Italy, Spain and Germany the highest percentage of this type of institutions.

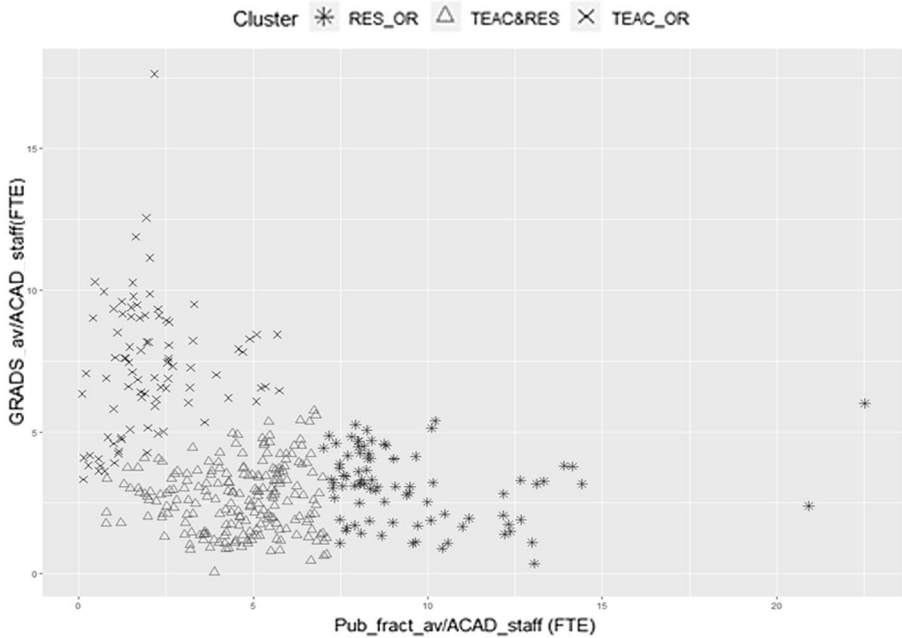


Fig. 2 Publications per Acad_staff vs graduates per Acad_staff for the three clusters

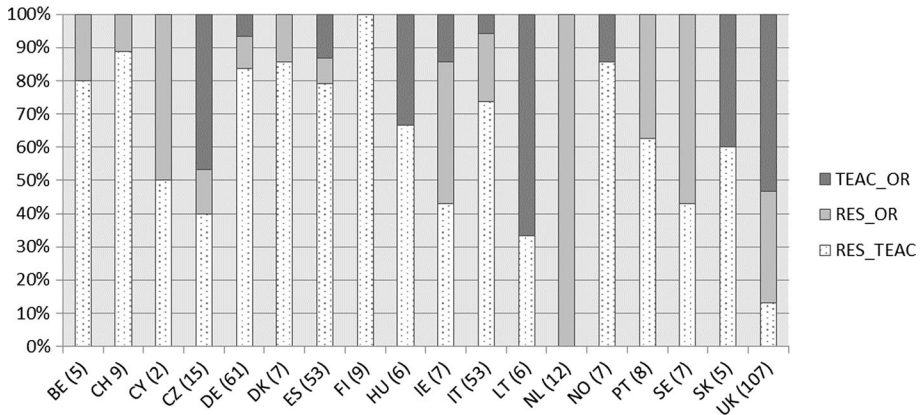


Fig. 3 Heterogeneity within countries according to the identified clusters, *Note On the X-axis, the number in brackets refers to the number of HEIs included in our sample, for each country. Notice that this number ranges from 107 for UK to 2 for Cyprus. Bulgaria and Malta were not included because only one observation was available. Vertical axis: Percentage of HE institutions

Focusing on large countries in terms of number of universities, a more detailed analysis reveals that a large fraction of UK institutions similar to those labelled in ETER as *universities of applied sciences* are categorized as *universities*, while in other countries many institutions with such label were not assigned to this category and consequently excluded from our analysis.

Table 7 Descriptive statistics on the governance variables for the obtained three clusters

	TEAC&RES	RES_OR	TEAC_OR	TEAC&RES	RES_OR	TEAC_OR
	Normalized governance structure			Internal governance structure		
GOV_Regulation	0.09	0.07	0.05	0.57	0.54	0.46
GOV_Expenditure	0.1	0.08	0.09	0.24	0.25	0.28
GOV_Taxes	0.12	0.13	0.18	0.07	0.09	0.12
GOV_Information	0.12	0.09	0.09	0.12	0.11	0.13
	Over regulation			Over all governance dimensions		
GOV_Constraints_trend	0.46	0.49	0.54	0.66	0.65	0.68
GOV_Opportunities_trend	0.54	0.51	0.46	0.34	0.35	0.32
GOV_Control_measures	0.33	0.29	0.29	0.33	0.29	0.29

According to Table 6 the RES_OR cluster is characterized by the largest 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 above the world average ($m_{nsc}(w-av)_{av} > 1.16$).

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”. This means that high quality research is able to attract external funds that are connected to innovative and patenting activities, which in turn are self-reinforcing to the scientific activities. Finally, the RES_OR cluster tends to contain relatively more institutions with academic hospitals than the other clusters do. On the other hand, we observe that the TEAC_OR cluster is characterized by the production of the largest 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 and highest policy interventions on Taxes. 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). TEAC_OR also tends to include institutions with a lower level of subjects mix specialization.

The TEAC&RES cluster shows, for most indicators, intermediary values among the two previously described groups, with a few significant exceptions. The HEIs in the cluster resulted to be the ones from countries with the highest attention on regulation, expenditure and information; consistent with the highest value registered for the percentage of control measures adopted.

Table 7 reports the values on governance variables, comparing the normalized structure measures and the internal structure measures. Considering the difference in proportion, the values of the two series of variables seem to be aligned (a slightly difference could be registered for the Expenditure dimension), showing a robustness in the variable that we proposed, that gives information on the normalized legislation activities in the considered countries. Regarding the values assumed by the three clusters (averages over the observations—institutions—contained in each group), even if it is possible to see differences over the values, the

distances are not pronounced. Taking into account the qualitative origin of governance data, the result can still be considered indicative about tendencies.

The connection between the country-level analysis and the institutional-level analysis is a difficult task. Although the inclusion of governance dimensions in the analysis may be a step forward, much more work is needed in articulating the two levels.

Efficiency analysis comparing the identified clusters

The indicators enrolled students, number of graduates over number of academic staff, percentage of administration personnel over the total staff, share of women professors, present the lowest values among the clusters, but only slightly. In general, the TEAC&RES institutions seem to show the lowest dimension (requiring less administrative staff for the management), but represent big players in their national context. We hypothesize that the lower average graduation productivity in the TEAC&RES cluster compared to that in the RES_OR cluster is due to the fact that the former shows a larger heterogeneity than the latter.

Finally, it may be speculated that the amount of efforts in producing a graduate student and the input needed to produce a scientific paper may not be equal over the cluster and, specifically, over different national HEIs system composition. The effect of these and other differences at the macro and the micro level upon the outcomes of the efficiency analysis could be the subject of a future secondary analysis. From the current results, it is interesting to note that the average FDH inefficiency score of the institutions in the group TEAC&RES (2.43) is higher (*i.e.*, they are less efficient) than the inefficiency scores of the RES_OR and those of the TEAC_OR groups (around 1.6). The same kind of result is obtained if the average DEA inefficiency score is considered (see Table 6).

We remind to the reader that the inefficiency scores are calculated considering an *output orientation*, meaning that given the inputs or resources used (in this case the academic staff), the institutions look at the maximum expansion of their outputs (in this case teaching-graduates and research-publications). The “norm” against which the efficiency is assessed is represented by the estimated efficient (or benchmarking) frontier, the frontier that envelops the analysed institutions, given by those that are producing more outputs (graduates and publications) given their inputs level. 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 suggests that the specialization in teaching or in research pays also in terms of efficiency of the overall activities carried out: universities specializing in teaching or in research, tend to have a higher efficiency (are closer to the estimated efficient frontier or frontier of the best practice) than those universities that balance research and teaching activities and that are located further away from the efficient frontier.

Additional methodological approaches and case studies

This section sketches two lines of future research, aimed to further broaden the insights into the performance of higher education institutions and into the factors that influence this performance. The first relates to the methodology to identify clusters, and the second to case studies providing a detailed comparison of particular countries.

Additional results from K-means with two clusters

The robustness of the clustering presented above should be further investigated. One way to do this is to vary the specified number of clusters to be created in a *K*-means clustering routine. As an illustration, Table 8 presents the outcomes of a *K*-means clustering with *two* clusters. Two groups of higher education institutions are obtained, one research oriented and a second teaching oriented. The two groups have the same values of the two inefficiency measures. Figure 4 gives the values of the two indicators Publications per Acad_staff vs graduates per Acad_staff in the two clusters.

The institutions previously included in the TEAC&RES cluster were distributed among the two specialized clusters, creating changes in the average values assumed by the descriptive variables. It is still possible to clearly identify the general clusters orientation (towards teaching or research activities), but some of the differences in values seems “smoothened” or, conversely, sharpened up. The considerations outlined in the opening part of paragraph 5 are still valid.

Secondary analysis comparing pairwise particular countries: England and Italy

An interesting follow-up of the observations contained in this work, would be a closer examination of a smaller group of HEIs selected in case studies, in order to better see differences in their context (market, economic development, regulation), strategy and position (e.g. research oriented, teaching oriented, generalist). To this end, for selected countries, a series of analyses concentrated on the internal system composition has been run. If a deeper analysis is left out for further research development and beyond the scope of this paper, the comparison of the results obtained for the two cases of England and Italy is considered particularly interesting in relation with the objective of this work. England and Italy present higher education systems of notable sizes, both are well represented in our database (in terms of number of singular observations) and comparisons between the two countries has already been carried out by a series of publications (Barbato and Turri 2019), due to the interestingly different approaches adopted by policy maker and institutions in the two contexts. For these reasons, we considered interesting to show new evidence for such a comparison widely treated in literature. Notice that, only English universities are considered in the analyses to take into account the different regulation adopted across the UK members (England, Scotland, Wales and Northern Ireland).

Figures 5 and 6 represent all institutions included in the used database from, respectively, England and Italy. Each symbol inside the plots represents an institution and the symbol's format indicates the cluster to which it belongs (as established in the previous section: RES_OR=research oriented, RES_TEAC=research and teaching oriented, TEAC_OR=teaching oriented). Figure 5 depicts both the higher education systems, the colour shows the country in which the institution is active (red for Italy, light blue for UK). In Fig. 6 depicts the two higher education systems separately, including more information on the institutions. Each symbol represent an institution, its size is proportional to the number of enrolled students (with reference to ISCED, from grade 5 to 7) and its colour shows the performance the institution scored in the efficiency analysis. A white colour is associated to 1, which is the maximum efficiency score: the institution is acting efficiently on its efficient frontier; smaller values smaller point to institutions that could improve the realization of their teaching and research outputs—respectively graduates and publications.

Table 8 Descriptive statistics on the main variables for the obtained two clusters

	RES_OR	TEAC_OR
<i>Cluster analysis variable</i>		
Grads_ISCED.5–7/ACADstaff	2.71	5.81
Pub_fract(av)/ACADstaff	6.73	2.26
<i>Inefficiency scores</i>		
FDH	2.231	1.801
DEA	3.091	3.037
<i>Basic data</i>		
Foundation_year	1817.99	1911.97
Uni_Hospital	0.6367	0.1377
Enrolled_student_ISCED.5–7	20,736.51	18,564.98
ACADstaff_FTE	1856.69	856.47
PhD_intensity_2014	0.0817	0.0202
WomenProff_share	0.1939	0.2444
<i>Specialization</i>		
Specialization	0.2595	0.2646
<i>Research quantity and quality</i>		
Pub_top10(av)/ACADstaff	0.045	0.0113
Pub_in_top10%	0.1086	0.0724
Pub_international_coll	0.5399	0.4903
mns_c_(w-av)_av	1.693	0.8746
<i>Third mission—Funds</i>		
Funds_external%	63,952,377.1	17,940,308.5
Funds_third_part/ACADstaff_FTE	0.2199	0.117
<i>Third mission—Patents</i>		
Patent_application(av)/ACADstaff	0.0027	0.0008
Back_citations(av)/ACADstaff	0.0118	0.0037
NPL_av	37.37	2.36
NPL_av/SPA_av	6.26	2.63
<i>National variables</i>		
GOV_Regulation (normalized)	0.071	0.048
GOV_Expenditure (normalized)	0.086	0.086
GOV_Taxes (normalized)	0.11	0.15
GOV_Information (normalized)	0.011	0.1
GOV_Constraints_trend	0.48	0.52
GOV_Opportunities_trend	0.52	0.48
GOV_Control_measures	0.31	0.29
EU_fract_country	0.088	0.095
NAT_HEIs_fract	0.028	0.03
NAT_UNI_fract (number)	0.53	0.655
NAT_UNI_fract	0.041	0.034

By inspecting Figs. 5 and 6, it is clear that the two national systems (England and Italy) have quite an opposite distribution of the institutions in the two-dimensional space. In England, hosting a differentiated system, the institutions are well polarized

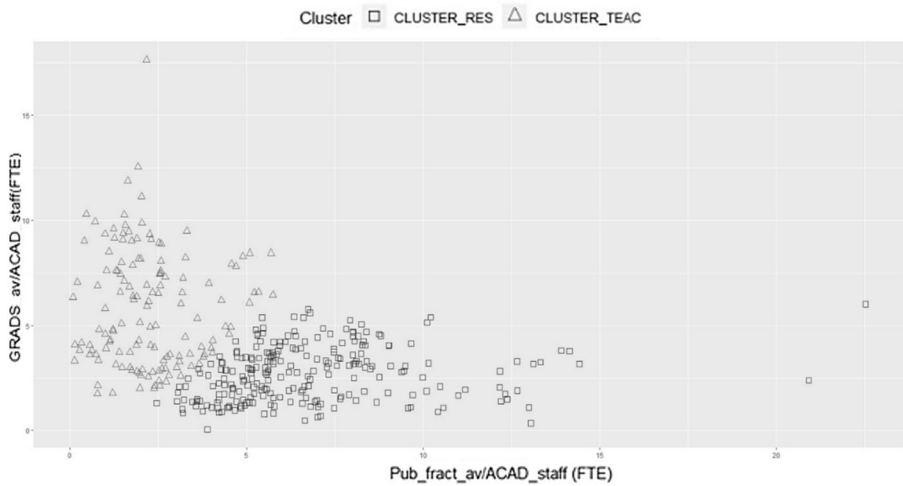


Fig. 4 Publications per Acad_staff versus graduates per Acad_staff for the two clusters

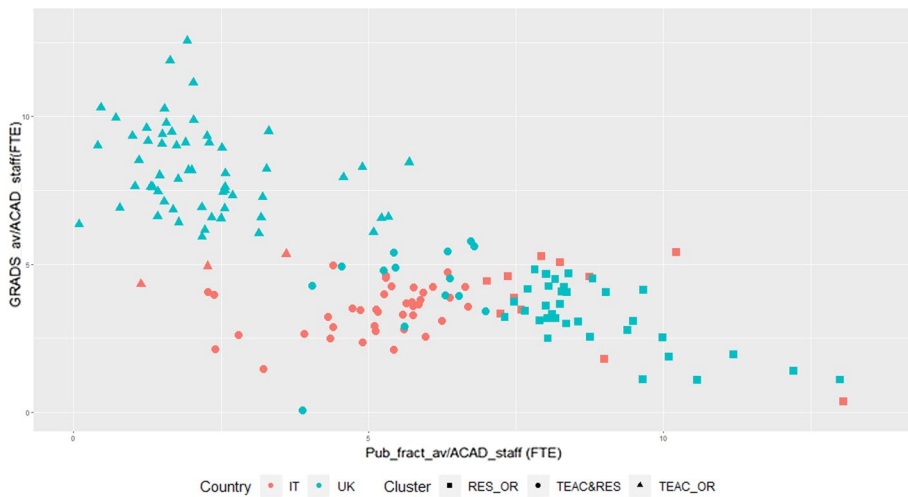


Fig. 5 Publications per Acad_staff versus graduates per Acad_staff, comparing England and Italy

in terms of a teaching versus research orientation; this result being in line with what it would be expected to observe in the case of countries in which academic funding is largely performance based already for several decades. Among the UK institutions labelled as ‘universities’ in ETER there are many *vocational universities*, also known as “*new*” *universities*, while in the Italian system this type of institution does not exist. The differences between the two sides of Fig. 6 (UK and Italy) clearly reflect this difference in the categorization of institutions. On the other hand, in Italy the opposite seems to stand, at the expense of the efficiency of the production of graduates and publications. Italy also seems to be characterized by institutions of bigger size and, thanks to

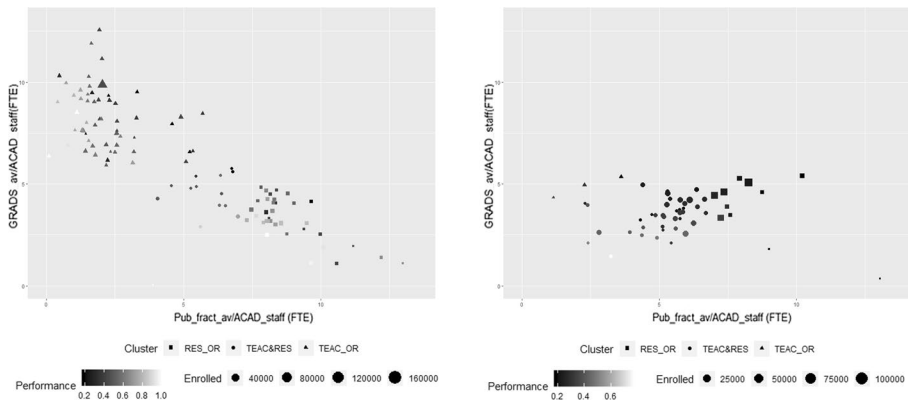


Fig. 6 Publications per Acad_staff versus graduates per Acad_staff (England and Italy) *Note Performance indicator is the inverse of the FDH efficiency score output oriented; performance = 1 corresponds to most efficient units, while performance = 0.2 corresponds to the less efficient unit in teaching and research efficiency. Size enrolled: number of enrolled students with reference to ISCED from 5 to 7

the lower degree of differentiation in terms of research versus teaching orientation, a higher value of publication per unit of academic staff seems to stand on average.

By inspecting Table 9 and Table 10, it appears that teaching oriented institutions in Italy are more efficient (research and teaching together), but have a much lower graduates per academic staff ratios. English research oriented institutions are more efficient and, accordingly, show a higher publication output. The English system seems to have a more distributed attention towards PhD programs, showing an interestingly high value of PhD intensity also in institutions included in the “generalist” cluster (24,8%; while Italy for the same variable presents a value lower than 3,5%); this observation being confirmed also by the qualitative analysis in paragraph 4. On the other hand, the Italian universities included in the teaching oriented cluster presents high specialization, compared to both the English counterparts and the Italian institutions included in the other clusters.

The English system shows a higher participation of women as full professor; nevertheless, in both the countries it still can be noticed a tendency to possible segregation of women in institution concentrated more on the teaching activities.

English institutions receive more funds from third parts, both in absolute values and in percentage with respect to the total amount of funds, especially for the institutions included in the research oriented clusters. Also, the English system seems to perform better on research quality, especially regarding the international collaboration in the published papers, that could be positively influenced by the country’s mother-language and the general higher presence of international students and international professors in the system. On the other hand, the Italian system seems to perform better on third mission activities, represented by quantitative and qualitative indicators on patents production.

The Italian system is more regulated and presents more requirement in transparency and information sharing. English system, instead, highly concentrates intervention on taxes and, less strongly, on expenditures. Both systems seem to be equilibrated with respect to the tendencies to apply constraints or create opportunities of enhancement on regulation dimension; same tendency can be registered also on all 4 governance dimensions overall, as confirmed in Capano and Pritoni (2019).

Table 9 Descriptive statistics on the main variables for the obtained three clusters- Italy and England

	Italy			England		
	T&R	RES	TEAC	T&R	RES	TEAC
<i>Cluster analysis variables</i>						
Grads_ISCED.5–7/ACADstaff	3.436	3.834	4.869	4.289	3.212	8.209
Pub_fract(av)/ACADstaff	5.081	8.533	2.340	5.636	8.985	2.237
<i>Inefficiency scores</i>						
FDH	1.715	1.259	1.530	1.743	1.604	1.375
DEA	3.044	2.020	3.092	2.705	2.164	2.001
<i>Basic data</i>						
Foundation_year	1771	1633	1941	1922	1830	1904
Uni_Hospital	0.641	0.636	0	0.25	0.793	0.064
Enrolled_student_ISCED.5–7	25,177	41,363	11,551	11,665	17,257	22,298
ACADstaff_FTE	1.389.455	1.866.682	458.083	946.406	2.213.276	872.553
PhD_intensity_2014	0.034	0.094	0.017	0.248	0.092	0.013
FullProf/ACADstaff_Head	0.167	0.160	0.164	0.130	0.162	0.089
WomenProff_share	0.195	0.175	0.23	0.241	0.22	0.314
Admn/TOTstaff_FTE	0.388	0.441	0.381	0.533	0.531	0.539
<i>Specialization</i>						
Specialization	0.265	0.241	0.372	0.237	0.234	0.178
<i>Research quantity and quality</i>						
Pub_top10(av)/ACADstaff	0.027	0.052	0.015	0.044	0.077	0.012
Pub_in_top10%	0.09	0.101	0.1	0.134	0.143	0.082
Pub_international_coll	0.446	0.472	0.403	0.608	0.603	0.499
mnsc_(w-av)_av	0.945	1.015	1.022	1.231	1.278	0.927
<i>Third mission – Funds</i>						
Funds_external%	0.087	0.183	0.072	0.186	0.306	0.069
Funds_third_part/ACADstaff_FTE	15,912	56,982	38,842	37,596	69,877	14,663
<i>Third mission – Patents</i>						
Patent_application(av)/ACAD-staff	0.001	0.003	0.001	0.002	0.002	0.001
Back_citations(av)/ACADstaff	0.004	0.013	0.003	0.007	0.006	0.002
NPL_av	10.036	17.891	1.000	11.375	21.290	1.834
NPL_av/SPA_av	6.085	6.940	1.000	7.658	3.805	2.566
<i>National variables</i>						
EU_fract_country	0.1			0.13		
NAT_HEIs_fract	0.015	0.017	0.011	0.005	0.008	0.01
NAT_UNI_fract (number)	0.449			0.801		
NAT_UNI_fract	0.016	0.018	0.012	0.005	0.008	0.01

Table 10 Descriptive statistics on the governance variables for Italy and England

	Italy	England
	Normalized governance structure indicators	
GOV_Regulation	0.091	0.04
GOV_Expenditure	0.073	0.09
GOV_Taxes	0.11	0.16
GOV_Information	0.17	0.1
	Over regulation	
GOV_Constraints_trend	0.55	0.54
GOV_Opportunities_trend	0.45	0.46
GOV_Control_measures	0.35	0.28

Discussion and conclusions

The aim to tackle the heterogeneity of European Higher Education Institutions is ambitious and requires an in-depth analysis from a multi-level and multi-methods perspective. This work presents an integrated set of approaches to the problem, developed to capture the main elements underlying the differentiation of European HEIs in performance, production strategies and missions. Further research is needed to capture the systemic implications of the subject.

We applied clustering techniques, including a teaching and research efficiency variable to analyse national (macro) and institution (micro) level data, revealing the potential of using micro level data to characterize national level performance.

The cluster analysis shows the existence of three groups of European universities clearly characterized by their orientation towards teaching activities, research activities or balancing among the two activities. Interestingly, the universities specialized in teaching or research show on average a higher efficiency then those oriented to the production of both teaching and research activities. The analysis reveals a possible trade-off for the academic staff between activities dedicated to PhDs’ training and undergraduates’ programs. Also, a “Matthew cumulative effect” can be observed: 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. We assessed the internal homogeneity of national systems, revealing differences in the ways countries carry out research and teaching activities. Focusing our analysis on the category of “universities” we left outside those institutions from countries with a binary system that are not active in research to offer a fair comparison on their overall activities including both research and teaching.

A more detailed analysis of the results obtained for two national systems, namely England and Italy, shows quite an opposite distribution of the institutions of the two countries in the research and teaching (two-dimensional) space. In England, that has a differentiated system, the institutions are well polarized between teaching and research oriented universities, while in Italy the opposite seems to stand, at the expense of the efficiency of the production of graduates and publications. Italy seems also to be characterized by institutions of bigger size, better quantitative and qualitative performance on third mission and higher specialization for institutions concentrating on teaching activities. England seems to have better quantitative and qualitative performance on research activities, associated to higher funds from third parties; it is having also a higher percentage of women full professors.

The results reveal once more the limits and dangers of *one-dimensional approaches* to the performance of HEIs. Analyses dealing merely with one single 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. However, 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.

Aside from the 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, and 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.

Acknowledgements This work is a substantially extended version of the paper “Characterizing the Heterogeneity of European Higher Education Institutions Combining Cluster and Efficiency Analyses” by the same authors (Bruni et al. 2019), published in the Proceedings of the ISSI 2019 Conference. 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 No 824091) is gratefully acknowledged. We warmly thank Ed Noyons for providing us the CWTS bibliometric data analysed in this work.

Funding Open access funding provided by Università degli Studi di Roma La Sapienza within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abankina, I., Aleskerov, F., Belousova, V., Gokhberg, L., Kiselgof, S., Petrushchenko, V., et al. (2016). From equality to diversity: Classifying Russian universities in a performance oriented system. *Technological Forecasting and Social Change*, 103, 228–239.
- Agasisti, T., & Bertoletti, A. (2019). Analysing the determinants of higher education systems’ performance—a structural equation modelling approach. *Science and Public Policy*, 46, 834–852.
- Agasisti, T., & Catalano, G. (2006). Governance models of university systems—towards quasi-markets? Tendencies and perspectives: A European comparison. *Journal of Higher Education Policy and Management*, 28(3), 245–262.
- Aghion, P., Dewatripont, M., Hoxby, C., Mas-Colell, A., & Sapir, A. (2010). The governance and performance of universities: evidence from Europe and the US. *Economic Policy*, 25(61), 7–59.
- Barbato, G., & Turri, M. (2019). What do positioning paths of universities tell about the diversity of higher education systems? An exploratory study. *Studies in Higher Education*, pp. 1–14.

- Bonaccorsi, A. (2014). *Knowledge Diversity and Performance in European Higher Education Changing Landscape*. Cheltenham: Edward Elgar Publishing.
- Bonaccorsi, A., & Daraio, C. (Eds.). (2007). *Universities and Strategic Knowledge Creation: Specialization and Performance in Europe. PRIME Series on Research and Innovation Policy in Europe*. Cheltenham: Edward Elgar.
- Bonaccorsi, A., & Daraio, C. (2009). Characterizing the European university system: a preliminary classification using census microdata. *Science and Public Policy*, 36(10), 763–775.
- Bonaccorsi, A., & Secondi, L. (2017). The determinants of research performance in European universities: A large scale multilevel analysis. *Scientometrics*, 112(3), 1147–1178.
- Bruni R., Catalano G., Daraio C., Gregori M., & Moed H.F. (2019). Characterizing the Heterogeneity of European Higher Education Institutions Combining Cluster and Efficiency Analyses, in Catalano G., Daraio C., Gregori M., Moed H.F. and Ruocco G. (2019), *Proceedings of the 17th International Conference on Scientometrics and Informetrics*, 2–5 September 2019, Sapienza University of Rome, Italy. Edizioni Efesto, ISBN: 978-88-3381-118-5, August 2019, (pp. 2094–2105).
- Capano, G., Howlett, M., & Ramesh, M. (Eds.). (2015). *Varieties of governance: Dynamics, strategies, capacities*. UK: Palgrave MacMillan.
- Capano, G., & Pritoni, A. (2019). Varieties of hybrid systemic governance in European higher education. *Higher Education Quarterly*, 73(1), 10–28.
- Catalano, G., Daraio, C., Gregori, M., Moed, H. & Ruocco, G. (2017). A scaling approach to tackle the heterogeneity of HEIs, *ISSI 2017-16th International Conference on Scientometrics and Informetrics, Conference Proceedings*, (pp. 1805–1806).
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429–444.
- Daraio, C. (2019). Econometric approaches to the measurement of research productivity. In W. Glänzel, H. F., Moed, H. Schmoch and M. Thelwall (Eds) *Springer handbook of science and technology indicators* (pp. 633–666).
- Daraio, C., & Simar, L. (2007). *Advanced robust and nonparametric methods in efficiency analysis, methodology and Applications*. New York (USA): Springer.
- 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. *Research Policy*, 40(1), 148–164.
- 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*, 68(2), 508–529.
- de la Torre, E. M., Casani, F., & Sagarra, M. (2018). Defining typologies of universities through a DEA-MDS analysis: An institutional characterization for formative evaluation purposes. *Research Evaluation*, 27(4), 388–403.
- Deprins, D., Simar, L., & Tulkens, H. (1984). Measuring labor inefficiency in post offices. In M. Marchand, P. Pestieau, & H. Tulkens (Eds.), *The Performance of Public Enterprises: Concepts and measurements* (pp. 243–267). Amsterdam: North-Holland.
- European Commission (2016). *Education and Training 2020. Highlights from the Working Groups 2014–2015* (https://ec.europa.eu/education/policy/strategic-framework/expertgroups_en.htm).
- Fumasoli, T., & Huisman, J. (2013). Strategic agency and system diversity: Conceptualizing institutional positioning in higher education. *Minerva*, 51(2), 155–169.
- García-Aracil, A., & Palomares-Montero, D. (2012). Agrupación alternativa para la evaluación de las universidades públicas españolas. *Regional and Sectoral Economic Studies*, 12(3), 177–192.
- Huisman, J., Lepori, B., Seeber, M., Frölich, N., & Scordato, L. (2015). Measuring institutional diversity across higher education systems. *Research Evaluation*, 24(4), 369–379.
- Lepori, B., Geuna, A., & Mira, A. (2019). Scientific output scales with resources. A comparison of US and European universities. *PLoS ONE*, 14(10), e00223415.
- López-Illescas, C., de Moya-Anegón, F., & Moed, H. F. (2011). A ranking of universities should account for differences in their disciplinary specialization. *Scientometrics*, 88(2), 563–574.
- Pinheiro, R., Langa, P. V., & Pausits, A. (2015). One and two equals three? The third mission of higher education institutions. *European Journal of Higher Education*, 5(3), 233–249.
- Rossi, F. (2010). Massification, Competition and Organizational Diversity in Higher Education: Evidence from Italy. *Studies in Higher Education*, 35(3), 277–300.
- Secundo, G., Perez, S. E., Martinaitis, Ž., & Leitner, K. H. (2017). An Intellectual Capital framework to measure universities' third mission activities. *Technological Forecasting and Social Change*, 123, 229–239.

- Wang, C., & Zha, Q. (2018). Measuring Systemic Diversity of Chinese Universities: A Clustering-Method Approach. *Quality and Quantity*, 52(3), 1331–1347.
- Zharova, A., Härdle, W. K., & Lessmann, S. (2017). Is scientific performance a function of funds?, SFB 649 Discussion Paper-Economic Risk, Berlin.