DIPARTIMENTO DI INGEGNERIA INFORMATICA
AUTOMATICA E GESTIONALE ANTONIO RUBERTI

SAPIENZA
UNIVERSITÀ DI ROMA

**Studying the Heterogeneity of European Higher Education Institutions**

Renato Bruni
Giuseppe Catalano
Cinzia Daraio
Martina Gregori
Henk F. Moed

# Studying the Heterogeneity of European Higher Education Institutions

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

[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)*

**Corresponding author:** *martina.gregori@uniroma1.it*

**Abstract**
The heterogeneity of the Higher Education (HE) Institutions is one of the main critical issues in the assessment of their performance. A multi-level and multi-dimensional perspective is adopted, combining national (macro) and institution (micro) level data, and measuring both research and teaching activity integrated with performance indicators derived from the European Tertiary Education Register (ETER), CWTS Leiden Ranking, and PATSTAT patent database. Clustering and efficiency analysis are combined to characterize the heterogeneity of national HE systems in European countries, revealing the potential of using micro level data to characterize national level performance. We discover large differences 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. We find evidence that universities specializing either in teaching or in research tend to have a higher efficiency than those institutions balancing research and teaching. We observe tradeoffs 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, heterogeneity, clustering, efficiency analysis, Europe*

**Msc:** 97B10 Education research and planning, 97B40 Higher Education

**JEL:** I23 Higher Education - Research institutions

**Article Highlights**

- European universities are partitioned in three groups: research oriented, teaching oriented and balanced orientation.

- Deeper analysis on England and Italy show an opposite distribution of HEIs in the research and teaching two-dimensional space.

- Limits and dangers of one-dimensional approaches, that may lead to unbalanced or even invalid conclusions, are revealed.

# 1. 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 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 system 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, more than the teaching one; Agasisti and Bertoletti, 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 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 Garcia-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 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

---

[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).*

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 final objective is to characterize HEIs while accounting for their:

(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, ETER[2]).

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 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[3], 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 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 current 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.


## 2. Methods

This work uses the multi-methodology approach summarized in Figure 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 higher education heterogeneity, for example, in China, Italy, Spain and EU country selections (Bonaccorsi and Daraio, 2009; Rossi, 2010; Garcia-Aracil and Palomares-Montero, 2012; Wang and Zha, 2018; Abankina et al., 2018; Barbato and Turri, 2019).
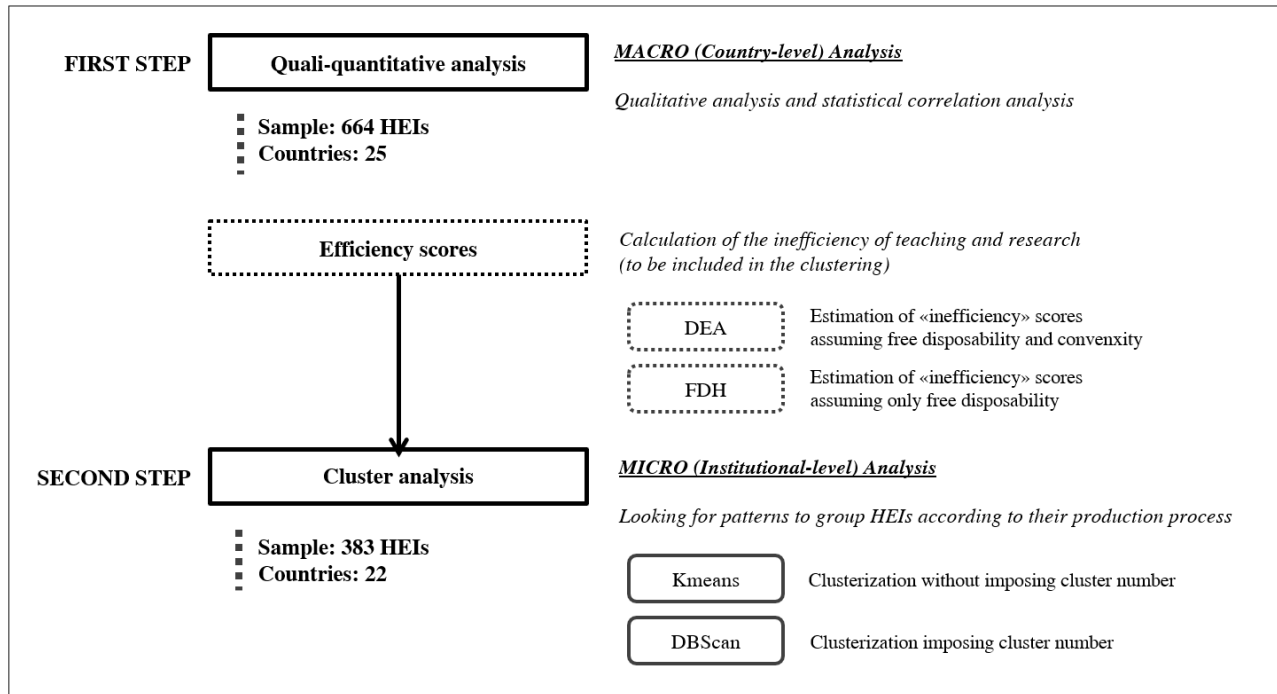


**Fig. 1 Tackling the heterogeneity of European HEIs: an outline of our approach**

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, Simar and Tulkens, 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 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.

## 3. 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 (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.*

The list of the variables considered is reported in Table 3. 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 represent the number of government interventions in the observed period for each specific sub-area. In order to include these indicators in our analysis, we applied a few data transformations. For each country, the scores per dimension were summed up. Next, per dimension, a *percentage* score was calculated relative to the total score on this dimension over all European countries considered.

A research funds dimension was also included in the analyses as a factor influencing HEIs positioning, mission and performance. The dimension can be represented by different elements. Table 1 lists research funds-based indicators that will be considered in the quali-quantitave analysis section, making reference to a useful classification of European countries according to whether they have research performance-based funding is proposed by Zacherewicz, Reale, Lepori and Jonkers (2018).

**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)* |

**Table 2. 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; tax 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. |

With respect to the multi-dimensional activities data in the final dataset (Table 3), 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, PATENTS database.

The teaching outputs (mainly, the number of graduates for each degree class) relate to the same horizon time of the input 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 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 variable values over the considered period of each included database; missing values had been excluded from the calculation.

We selected all the institutions categorized as "universities" in the ETER dataset, leaving outside the analysis the other two categories, namely "universities of applied sciences" and "others" including higher education institutions not very active in research.

**Table 3. 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. |
| *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. |
| | **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** | 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. |
| *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). |

Moreover, 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 second quantitative analysis was performed only on a sub-selection of the database composed by 383 HEIs from 22 countries.

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

## 4. Quali-quantitative analyses

*4.1 Characterizing the heterogeneity of HE systems by combining bibliometric indicators, higher education data and research performance-based funding (RPBF) information*

This section uses a classification of European countries according to whether they have research performance-based funding, proposed by Zacherewicz, Reale, Lepori and Jonkers (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* consider 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.
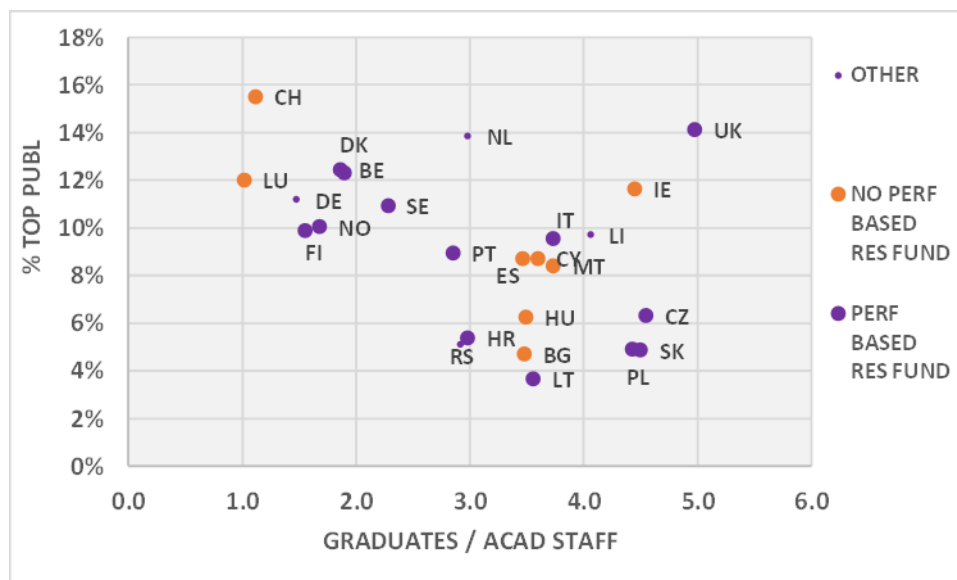


**Fig. 2 Scatterplot of % Top Publications against Graduates per Academic Staff**

The percentage of top publications (% TOP PUBL in Figure 2) 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 2 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 2 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 2 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.

*4.2 Statistical correlations between 6 key variables*

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

| | | |
|---|---|---|
| Total academic staff (ACADstaff_FTE); | Publications per academic staff (Pub_fract(av)/ACADstaff); | Percentage of Top Publications (Pub_in_top10%); |
| PhD intensity (PhD_intensity_2014); | Graduates per academic staff (Grads_ISCED.5-7/ACADstaff); | 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 are available both for governance indicators and for the first 5 key indicators, no governance indicators are included in the key set.

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

| *Variable 1* | *Variable 2* | *Pearson corr.* | | *Partial corr.* | |
|---|---|---|---|---|---|
| | | R | Prob | R | Prob |
| Pub_in_top10% | Pub_fract(av)/ACADstaff | 0.74 | 0.00 | 0.82 | 0.00 |
| Pub_in_top10% | PhD_intensity_2014 | 0.53 | 0.01 | 0.44 | 0.07 |
| Pub_in_top10% | ACADstaff_FTE | 0.26 | 0.21 | 0.61 | 0.01 |
| Pub_in_top10% | Grads_ISCED.5-7/ACADstaff | -0.44 | 0.03 | -0.47 | 0.05 |
| Pub_in_top10% | RPBF | -0.13 | 0.58 | -0.57 | 0.01 |
| RPBF | Pub_fract(av)/ACADstaff | 0.12 | 0.60 | 0.49 | 0.04 |
| RPBF | ACADstaff_FTE | 0.23 | 0.30 | 0.48 | 0.04 |
| Grads_ISCED.5-7/ACADstaff | PhD_intensity_2014 | -0.61 | 0.00 | -0.22 | 0.39 |
| Grads_ISCED.5-7/ACADstaff | ACADstaff_FTE | 0.09 | 0.66 | 0.55 | 0.02 |

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.

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 counterintuitive. 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 3 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 4.
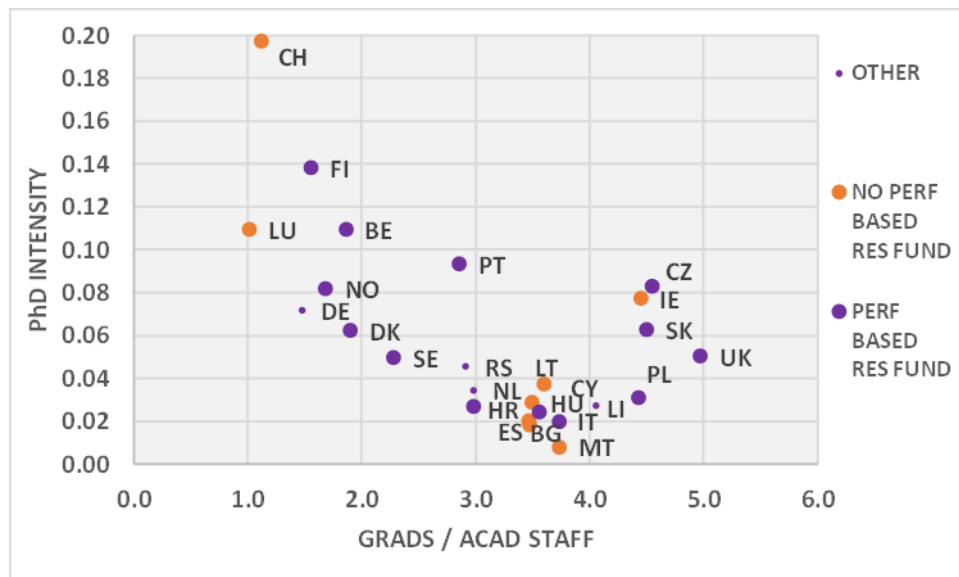


**Fig. 3 Scatterplot of PhD Intensity against Graduates per Academic Staff**

- 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 3.

- 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.

*4.3 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 can be attempted to connect general tendency and time evolution, with specific results in performance (e.g. increases/decreases in efficiency). A first analysis on the database gives insights into the internal regulatory structure adopted by each country. Table 5 displays the within-country percentages for each regulatory macro-dimension of the observed 12 countries.

**Table 5. National systems regulation structure and comparison with the typical behaviour**

| 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** | | | | |

Source: our elaborations on Capano and Pritoni's (2019) indicators.

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. 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), 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 position applied at national level across European countries. Focusing on the Regulation Indicators included in Capano and Pritoni (2019), 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, 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 6. A proxy on "control measures" adopted by government is calculated, as reported in Paragraph 3, based on the number of interventions registered on the areas of rules on goals in teaching, of assessment, and of monitoring and reporting. Table 6 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 (at least, on regulatory level).

**Table 6. Governance tendency to adopt control measures on HEIs with respect to the tendency to either limit or encourage HEIs autonomy on Regulation Dimension**

|  | Percentage constraints | Percentage 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 |

## 5. 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 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 variables) and (ii) average graduates per academic staff (Grads_ISCED.5-7/ACADstaff).

In particular, the results obtained by the K-means cluster analysis using 3 clusters, after the DBSCAN analysis that suggested the existence of three clusters, identified three groups of universities whose main characteristics are outlined in Table 6. We labelled the three groups as: research and teaching oriented (TEAC&RES), research oriented (RES_OR) and teaching oriented (TEAC_OR).

Note 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. Figure 4 illustrates how well the three clusters are spread along the two clustering dimensions.
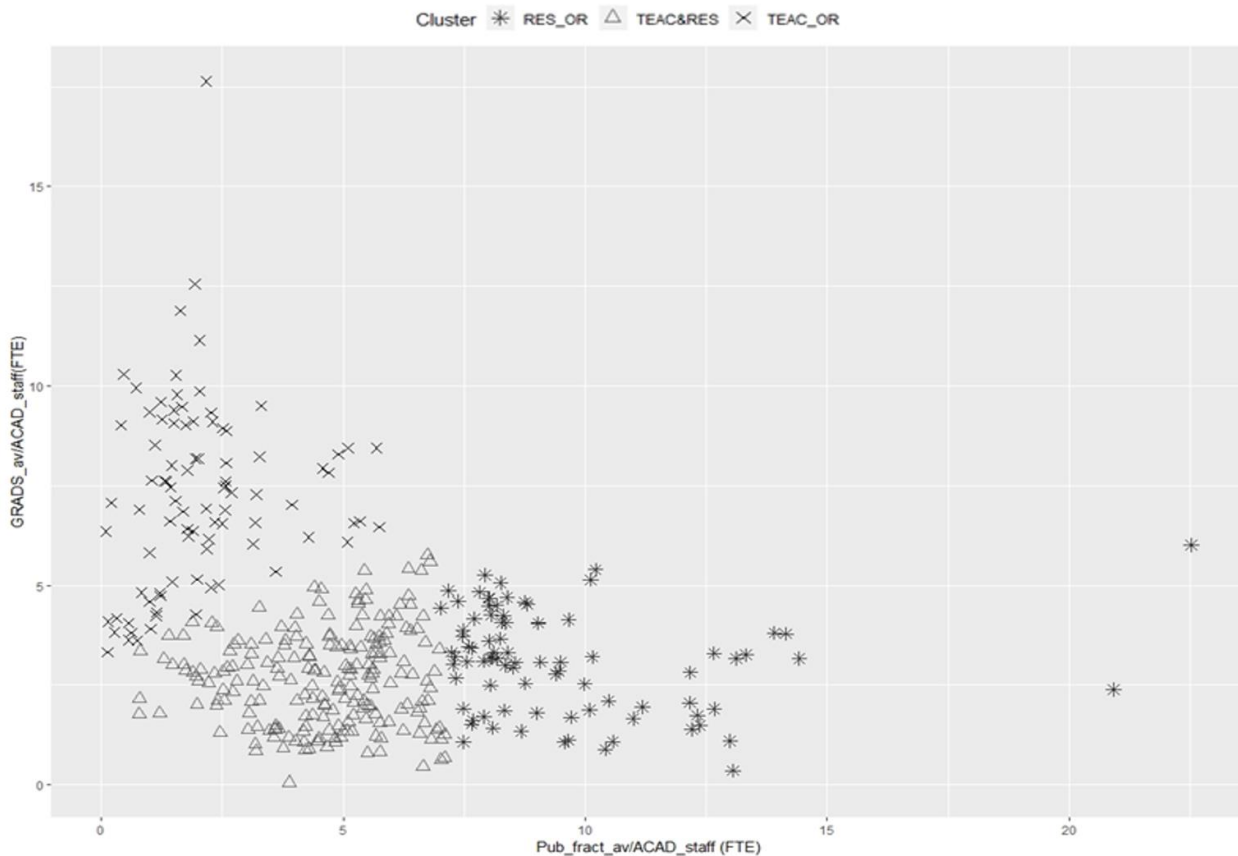


**Fig. 4 Publications per Acad_staff vs graduates per Acad_staff for the three clusters.**

Figure 5 shows the distribution of universities in the three clusters by country. It reveals large differences among countries. The TEAC&RES cluster contains institutions from several Scandinavian countries (with a high scientific performance)[4], Switzerland and – with a smaller percentage of national institutions, but always higher than 50% - Germany, Belgium, Spain and Italy. More than the 50% of the national institutions of Lithuania, Czech Republic, Slovakia, UK and Hungary are included in the TEAC_OR cluster. The RES_OR institutions percentage is, instead, particularly high for the Netherlands, Sweden, Portugal and Ireland.

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

---

[4] *Finland, Norway, Denmark and Switzerland have more than 70% or the national institutions included in TEAC&RES cluster.*

Focusing on large countries in terms of number of universities, a more detailed analysis reveals that a large fraction of UK institutions 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. As outlined in the Data section, the analyses presented in this work concern to institutions categorized as universities in ETER.
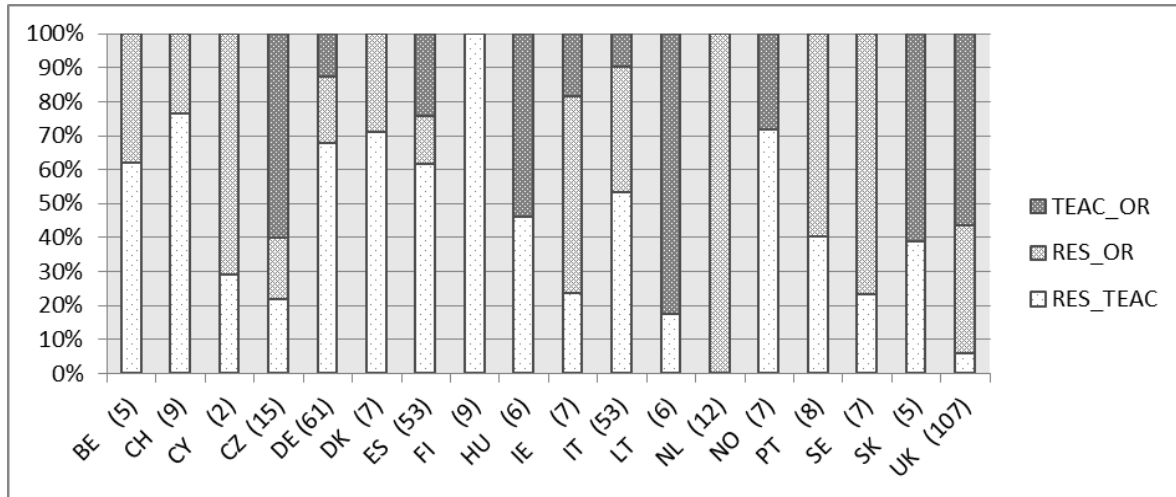


**Fig. 5 Heterogeneity within countries according to the identified clusters**
**\* Note: On the X-axe, the *number* in brackets refers to the number of HEIs analysed in each country. Notice that this number ranges from 107 for UK to 2 for Cyprus. Bulgaria and Malta were not included because only one observation was available.**

According to Table 7 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 (mnsc_(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. 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 (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 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. 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 |
| *Inefficiency scores* | DEA | 3.72 | 2.14 | 2.44 |
| | 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 |

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[5] (requiring less administrative staff for the management), but represent big players in their national context[6]. We hypothesize that the lower average graduation productivity in the TEAC&RES

---

[5] *See "EU_fract_country", Table 7.*
[6] *See "Nat_HEIs_fract, Table 7.*

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 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 7). 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.

*5.1 Additional results from K-means with two clusters*

In this section we propose some additional results of analyses carried out to identify (select) the HEIs that need to be further investigated through case studies. To check the robustness of the choice of the number of clusters, we run a K-means clustering with 2 groups. The results are described in Table 8 and in Figure 6, which highlights the presence of two clusters of universities: one teaching oriented and another research oriented.
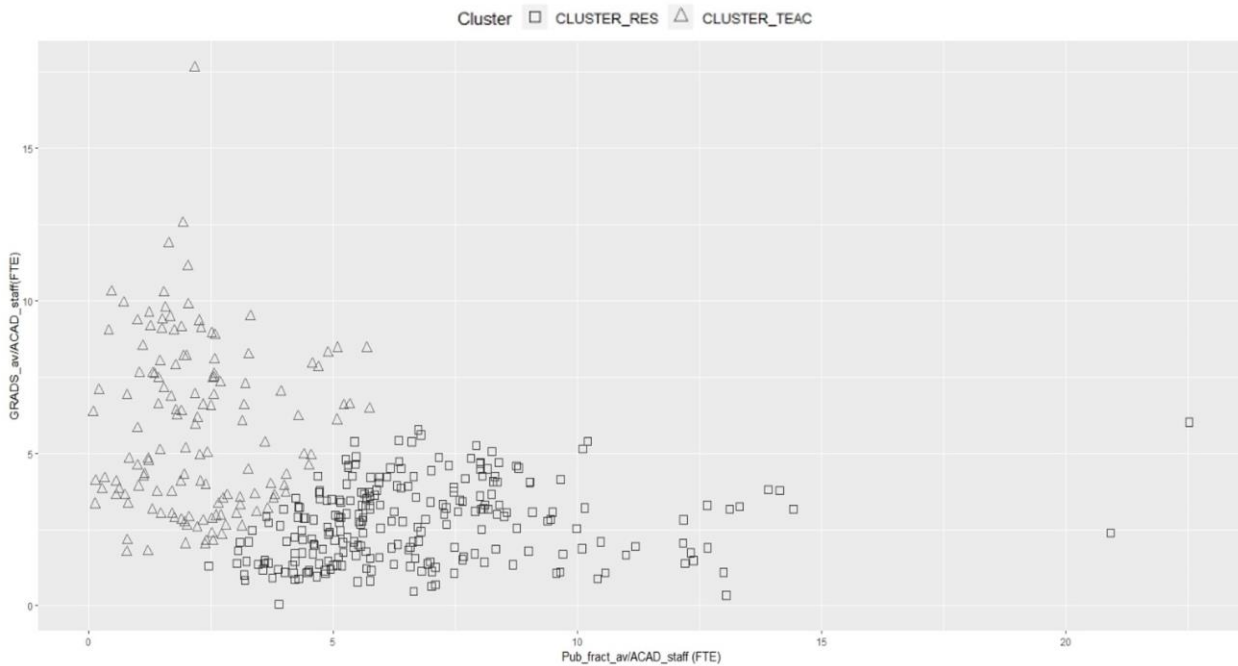


**Fig. 6 Publications per Acad_staff vs graduates per Acad_staff for the two clusters**

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.

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

|  |  | RES_RES | TEAC_OR |
|---|---|---|---|
| *Cluster analysis variable* | Grads_ISCED.5-7/ACADstaff | 2.71 | **5.81** |
|  | Pub_fract(av)/ACADstaff | **6.73** | 2.26 |
| *Inefficiency scores* | FDH | 2.231 | 1.801 |
|  | DEA | 3.091 | 3.037 |
| *Basic data* | Foundation_year | 1817.99 | 1911.97 |
|  | Uni_Hospital | 0.6367 | 0.1377 |
|  | Enrolled_student_ISCED.5-7 | 20736.51 | 18564.98 |
|  | ACADstaff_FTE | 1856.69 | 856.47 |
|  | PhD_intensity_2014 | 0.0817 | 0.0202 |
|  | WomenProff_share | 0.1939 | 0.2444 |
|  | Funds_external% | 63952377.1 | 17940308.5 |
|  | Funds_third_part/ACADstaff_FTE | 0.2199 | 0.117 |
| *Third mission - Funds* | Specialization | 0.2595 | 0.2646 |
|  | Pub_top10(av)/ACADstaff | 0.045 | 0.0113 |
| *Specialization* | Pub_in_top10% | **0.1086** | 0.0724 |
| *Research quantity and quality* | Pub_international_coll | 0.5399 | 0.4903 |
|  | mnsc_(w-av)_av | **1.693** | 0.8746 |
|  | Patent_application(av)/ACADstaff | 0.0027 | 0.0008 |
|  | Back_citations(av)/ACADstaff | **0.0118** | 0.0037 |
| *Third mission - Patents* | NPL_av | **37.37** | 2.36 |
|  | NPL_av/SPA_av | **6.26** | 2.63 |
|  | GOV_Regulation | **7.071** | 4.820 |
|  | GOV_Expenditure | 8.550 | 8.633 |
| *National variables* | GOV_Taxes | 11.430 | 14.809 |
|  | GOV_Information | 10.450 | 10.127 |
|  | 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 |

## 5.2 Secondary analysis comparing England and Italy

An interesting follow-up of the observations contained in this work, would be a closer examination of a small selection 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. The comparison of the results obtained for the two cases of England and Italy are considered particularly interesting in relation with the objective of this work. 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. 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).

Figure 7 and Figure 8 represent all institutions from, England and Italy, respectively. Each symbol inside the plot 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). The size of the symbol 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 green 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.
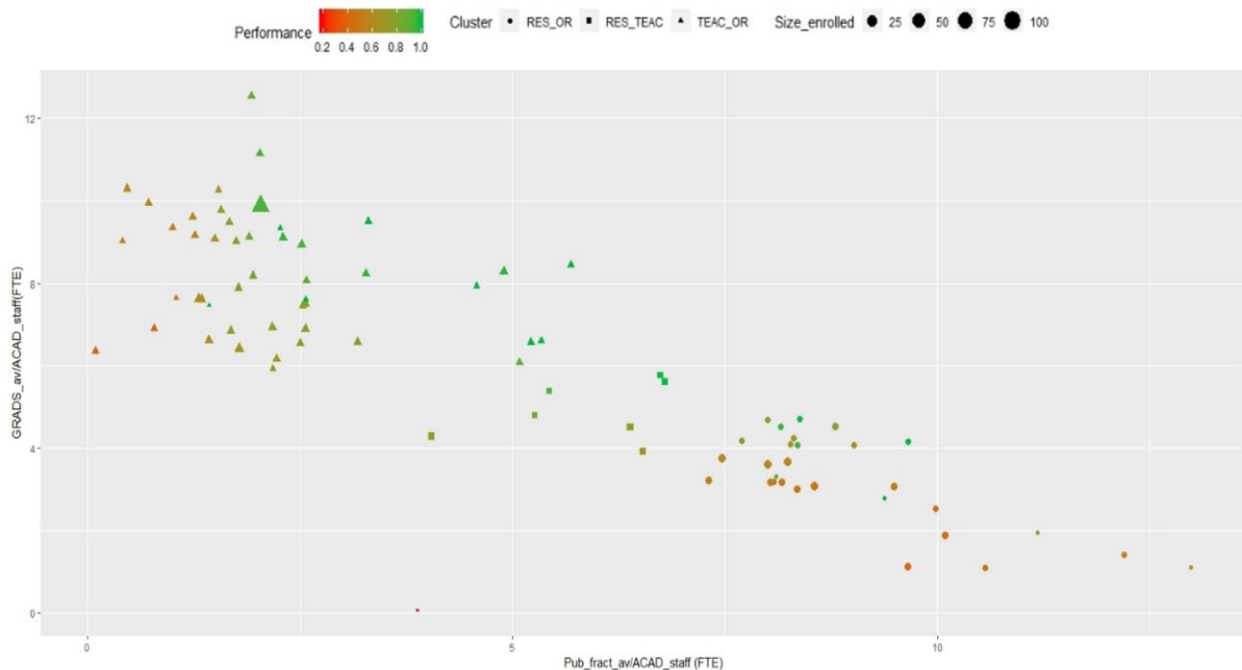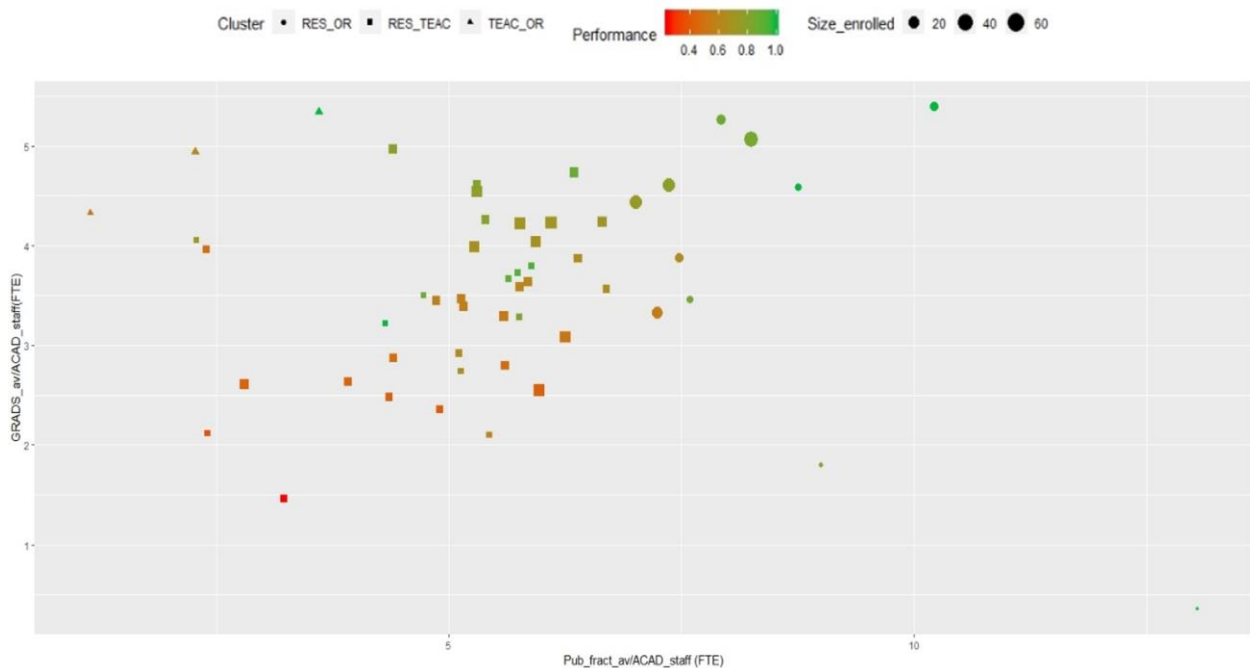


**Fig. 7 Publications per Acad_staff vs graduates per Acad_staff (England)**

*\* 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.*

By inspecting Figure 7 and 8, 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 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. 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 the lower degree of differentiation in terms of research versus teaching orientation, a higher value of publication per unit of academic staff seem to stand on average.



**Fig. 8 Publications per Acad_staff vs graduates per Acad_staff (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.**

By inspecting Table 9, 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%). 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.

**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** |
| 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 |
| Mod.Teach.Res.X_ACADSTAFF.FDH | 1.715 | **1.259** | **1.530** | 1.743 | **1.604** | **1.375** |
| Mod.Teach.Res.X_ACADSTAFF.DEA.VRS | 3.044 | 2.020 | 3.092 | 2.705 | 2.164 | 2.001 |
| 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 | 25177 | 41363 | 11551 | 11665 | 17257 | 22298 |
| 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 |
| Funds_external% | 0.087 | **0.183** | 0.072 | 0.186 | **0.306** | 0.069 |
| Funds_third_part/ACADstaff_FTE | 15912 | 56982 | 38842 | 37596 | 69877 | 14663 |
| Specialization | 0.265 | 0.241 | **0.372** | 0.237 | 0.234 | **0.178** |
| 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 |
| Patent_application(av)/ACADstaff | 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 |
| GOV_Regulation | **9.146** | | | 3.862 | | |
| GOV_Expenditure | 7.292 | | | **8.854** | | |
| GOV_Taxes | 10.811 | | | **16.216** | | |
| GOV_Information | **16.667** | | | 9.524 | | |
| GOV_Constraints_trend | 0.55 | | | 0.54 | | |
| GOV_Opportunities_trend | 0.45 | | | 0.46 | | |
| GOV_Control_measures | 0.35 | | | 0.28 | | |
| EU_fract_country | 0.103 | | | 0.131 | | |
| 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 |

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.

## Discussion and conclusions

This work proposes a multi-level and multi-dimensional study of the heterogeneity of European Higher Education institutions.

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.

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 present better quantitative and qualitative performance on research activities, associated to higher funds from third parties; it is having also a higher the 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.

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, 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.

## References

Abankina, I., Aleskerov, F., Belousova, V., Gokhberg, L., Kiselgof, S., Petrushchenko, V., ... and Zinkovsky, K. (2016). From equality to diversity: Classifying Russian universities in a performance oriented system. *Technological Forecasting and Social Change*, vol. 103, pp. 228 – 239.

Agasisti, T., and Bertoletti, A. (2019). Analysing the determinants of higher education systems' performance–a structural equation modelling approach. *Science and Public Policy*.

Agasisti, T. and Catalano, G. (2006). Governance models of university systems—towards quasi‐markets? Tendencies and perspectives: A European comparison. *Journal of Higher Education Policy and Management*, vol. 28, n. 3, pp. 245 – 262.

Aghion, P., Dewatripont, M., Hoxby, C., Mas-Colell, A., and Sapir, A. (2010). The governance and performance of universities: evidence from Europe and the US. *Economic policy*, vol. 25, n. 61, pp. 7 – 59.

Barbato, G., and 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. (Ed.). (2014). *Knowledge, Diversity and Performance in European Higher Education: A Changing Landscape*. Edward Elgar Publishing.

Bonaccorsi, A. and 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., and Daraio, C. (2009). Characterizing the European university system: a preliminary classification using census microdata. *Science and Public Policy*, vol. 36, n. 10, pp. 763 – 775.

Bruni R., Catalano G., Daraio C., Gregori M. and 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 & 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. and Ramesh, M. (Eds.). (2015). *Varieties of governance: Dynamics, strategies, capacities*, Palgrave MacMillan, UK.

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

Catalano, G., Daraio, C., Gregori, M., Moed, H. and 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. and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, vol. 2, n. 6, pp. 429 – 444.

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., pp. 633 – 666.

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

Daraio, C., Bonaccorsi, A., Geuna, A., Lepori, B., Bach, L., Bogetoft, P., ... and Fried, H. (2011). The European university landscape: A micro characterization based on evidence from the Aquameth project. *Research policy*, vol. 40, n. 1, pp. 148 – 164.

Daraio, C. and 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. and 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.

de la Torre, E. M., Casani, F., and Sagarra, M. (2018). Defining typologies of universities through a DEA-MDS analysis: An institutional characterization for formative evaluation purposes. *Research Evaluation*, vol. 27, n. 4, pp. 388 – 403.

Deprins, D., Simar, L. and Tulkens, H. (1984), Measuring labor inefficiency in post offices. In: *The Performance of Public Enterprises: Concepts and measurements*. M. Marchand, P. Pestieau and H. Tulkens (eds.), Amsterdam, North-Holland, pp. 243 – 267.

Ester, M., Kriegel, H.P. Sander, J. and 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.

European Commission (2016), *Education and Training 2020. Highlights from the Working Groups 2014 – 2015* (http://ec.europa.eu/education/policy/strategic-framework/expertgroups_en.htm).

Fumasoli, T., and Huisman, J. (2013), Strategic Agency and System Diversity: Conceptualizing Institutional Positioning in Higher Education, *Minerva*, vol. 51, n. 2, pp. 155 – 69.

Huisman, J., Lepori, B., Seeber, M., Frølich, N., and Scordato, L. (2015). Measuring institutional diversity across higher education systems. *Research Evaluation*, vol. 24, n. 4, pp. 369 – 379.

Lepori, B., Geuna, A., Mira, A. (2019). Scientific output scales with resources. A comparison of US and European universities. *PloS one*, vol. 14, n. 10.

López-Illescas, C., de Moya-Anegón, F. and 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.

Rossi, F. (2010). Massification, Competition and Organizational Diversity in Higher Education: Evidence from Italy, *Studies in Higher Education*, vol. 35, n. 3, pp. 277 – 300.

Wang, C., and Zha, Q. (2018). Measuring Systemic Diversity of Chinese Universities: A Clustering-Method Approach, *Quality and Quantity*, vol. 52, n. 3, pp. 1331 – 47.

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

Zharova, A., Härdle, W. K., and Lessmann, S. (2017). Is Scientific Performance a Function of Funds?, *SFB 649 Discussion Paper - Economic Risk*, Berlin.