ORIGINAL RESEARCH

Climbing university rankings under resources constraints: a combined approach integrating DEA and directed Louvain community detection

Simone Di Leo[1] · **Alessandro Avenali[1]** · **Cinzia Daraio[1]** ·
Joanna Wolszczak-Derlacz[2]

Abstract

Over recent years, scholarly interest in universities' allocation and effective utilisation of financial resources has been growing. When used efficiently, financial resources may improve universities' quality of research and teaching, and therefore their positions in world university rankings. However, despite the relevance of financial efficiency to university placement in academic rankings, universities' total available financial resources appear much more significant. In the present study, we propose an innovative methodology to determine realistic ranking targets for individual universities, based on their available financial resources. In particular, we combine data envelopment analysis, as developed by Banker et al. (Manag Sci 30(9):1078–1092, 1984), and a directed Louvain community detection algorithm to examine 318 universities across five countries, considering their ARWU scores alongside key financial indicators (i.e., long-term physical capital, total operating revenues). We identify clusters of universities with similar financial profiles and corresponding ARWU scores, as well as universities that have optimised their use of financial resources, representing benchmarks for similar universities to emulate. The approach is subsequently applied to Italian universities, as a specific national case. The findings may be useful for policy makers and university managers seeking reliable strategies for climbing academic rankings, particularly in countries with limited public investment in higher education.

Keywords DEA · Community detection · University rankings · ARWU ranking · Financial sustainability

---

✉ Simone Di Leo
   dileo@diag.uniroma1.it

   Alessandro Avenali
   avenali@diag.uniroma1.it

   Cinzia Daraio
   daraio@diag.uniroma1.it

   Joanna Wolszczak-Derlacz
   jwo@zie.pg.edu.pl

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

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

🐾 Springer

1 Introduction

As the financial resources available to academic institutions have decreased, the optimisation of these resources has become increasingly relevant. Universities, as complex institutions, require considerable financial support to carry out their educational, research and community service activities (i.e., "third mission" activities, as popularised by Etzkowitz & Leydesdorff, 1997). Thus, over the past decade, the financial sustainability of universities has become a critical issue. Only institutions that fully understand the costs of their activities can assess their financially viability, identify potential efficiency improvements in comparison to other universities, and define targeted strategies to prevent overspending (Di Carlo et al., 2019).

Universities' allocation of financial resources has a direct impact on their ability to attract high-calibre faculty, conduct quality research and offer competitive curricula. While university rankings aim at showcasing these strengths by classifying higher education institutions (HEIs) on the basis of various outputs (i.e., teaching, research, knowledge transfer, reputation), their methodologies often fail to account for the vast disparities in financial resources among universities, in terms of both total amounts and per capita (e.g., per student or staff member). University rankings were introduced as tools to inform university governance. Despite their limitations, they continue to be widely used and considered (see, e.g., the recent book by Hazelkorn & Mihut, 2021). However, debates around their underlying data, theory and methodology persist (Daraio et al., 2022), highlighting their significant impact on policies, strategies and competition dynamics within higher education. Therefore, it is necessary to deepen our understanding of the interaction between universities' financial resources and their positioning in prominent academic rankings. Indeed, efficient utilisation of financial resources enables universities to deliver high-quality research and teaching, thus bolstering their standings in the global rankings. However, while financial efficiency is certainly a relevant factor in determining universities' positioning in the rankings, the total amount of financial resources at their disposal appears much more significant.

To address this latter factor, the present research applied an innovative methodology to identify realistic ranking targets for universities based on their available financial resources. Our approach integrated data envelopment analysis (DEA) with variable returns to scale (VRS) and the directed Louvain community detection algorithm, leveraging data from the Academic Ranking of World Universities (ARWU) and financial metrics extracted from universities' 2019 financial statements (i.e., long-term physical capital, total amount of operating revenues). The analysis aimed at identifying clusters of universities with similar financial resources and comparable rankings in the ARWU, considering a set of 318 universities from five countries (i.e., USA, UK, Italy, Australia, Poland).[1]
Specifically, we utilised the DEA-VRS framework introduced by Banker et al. (1984) to assess universities' efficiency in maximising their positions in the ARWU ranking, relative to their available financial resources. The DEA considered long-term physical capital and total operating revenues as inputs, and ARWU score as the output. Subsequently, applying the Louvain network partition to the DEA results, we identified homogeneous communities of universities characterised by similar financial resources and ARWU scores, and we explored the correspondence between these groups and groups based solely on ARWU position.

Following our analysis of the international sample, we narrowed our focus to the Italian context (i.e., 46 Italian universities) to apply the proposed methodology within a more

---

[1] The year 2019 was selected to prevent the potential bias of the impact of COVID-19 on universities' investment choices in the subsequent period.

homogeneous context. In Italy, financial data are subject to a standardised national accounting system, facilitating a more nuanced policy assessment calibrated to the Italian landscape. Importantly, financial statement data from different countries present several criticalities due to variations in accounting standards, which may prevent a uniform data collection. Furthermore, the Italian case demonstrated the scalability of the proposed methodology and its adaptability to different analytical contexts.

The remainder of the paper is organised as follows. Section 2 describes the state of the art in the literature, introduces the Italian university system and details the main contribution of the present work. Section 3 describes the methodology and dataset. Finally, Sect. 4 reports the findings and Sect. 5 offers concluding remarks.

2 Literature review and contribution

The following section is divided into three subsections. The first subsection discusses how the literature has addressed the evaluation of university performance and the use of rankings in this regard. The second subsection provides insight into the Italian university system, while the third subsection introduces the identified gap in the literature and elucidates the contribution of the present work to the ongoing debate.

2.1 The performance evaluation of universities

The performance evaluation of universities has become a fundamental practice in higher education. Valmorbida and Ensslin (2017) described that the performance assessment process consists of four sequential steps. The first step involves identifying the evaluation context and the managers involved. Based on this identification, the second step involves identifying the relevant management objectives and their corresponding indicators. The third step involves measuring these indicators (i.e., the assignation of values). Finally, the fourth step involves leveraging this information for management (i.e., pinpointing areas for improvement, implementing corrective actions, learning from outcomes to enhance overall performance). This evaluative process encompasses various aspects, including research, teaching, technology transfer, social impact and governance, among others.

The evaluation of university performance is crucial for several reasons. First, it helps to guarantee the quality of higher education by ensuring that universities meet acceptable standards of teaching, research and student services. Second, it promotes accountability and transparency by providing an independent mechanism for assessing university performance and disclosing the results to stakeholders.

University rankings have garnered significant interest from both the academic community and policy makers, as tools for assessing university reputation and performance. Rankings classify universities based on predefined criteria, with the aim of highlighting real or perceived differences in performance or reputation (Merisotis & Sadlak, 2005). Notable rankings include the ARWU (also known as the Shanghai Ranking), the Times Higher Education World University Rankings (THE) and the Quacquarelli-Symonds World University Ranking (QS). Unlike the THE and QS, the ARWU does not incorporate financial indicators (the THE research and teaching scores include some income-based components). Additionally, it relies solely on publicly available and replicable data (in contrast, the QS and THE are based on data collected through surveys administered within universities). The ARWU evaluates universities based on indicators such as research quality, academic citations, awards and

international collaboration (Liu et al., 2016). Compared to other rankings, it is less prone to issues such as data manipulation and the overweighting of subjective indicators (Hazelkorn, 2017).

Initially conceived as information tools, academic rankings have evolved into performance assessment instruments, responding to the escalating competition within higher education and addressing the growing demand for transparency and accountability. However, this evolution has sparked debate regarding the limitations of rankings as comprehensive indicators of academic quality and their potential influence on university strategies and priorities. In more detail, several criticisms have been raised concerning the selection and weighting of indicators and the inherent limitations of the underlying data (for more detail, see Fauzi et al., 2020; Moed, 2017; Olcay & Bulu, 2017). In addition to these methodological and data criticisms, the impact of academic rankings on university finances has also been raised as a relevant concern. Benito et al. (2019) investigated the influence of funding on the positions obtained by the top 300 universities in the 2018 QS, finding that public funding significantly impacted 84% of these universities, explaining up to 51% of the variability in their ranking positions. Similarly, Berne (2020) demonstrated that the performance of HEIs, as measured by the ARWU, correlated closely with institutional finances (expressed in terms of annual budget per student) and student tuition fees. Furthermore, Avenali et al. (2024) developed a machine learning-based tool to predict university placement in the ARWU, primarily using financial statement data.

Lepori et al. (2017) analysed the financial environment of HEIs as a possible selection mechanism. They identified a subset of research universities characterised by high research volume and significantly greater funding compared to the other HEIs in the sample. The authors suggested that the emergence of these HEIs was critically related to resource concentration. Later, Lepori et al. (2019) empirically demonstrated that global academic rankings can be interpreted as measures of wealth, as US universities, which typically enjoy significantly more resources, tend to produce more publications and citations than their European counterparts. This suggests that the incorporation of a resource metric may improve the accuracy of international rankings.

Regarding financial resources, Baltaru et al. (2022) studied resource disparities between 102 English universities over the period 2008–2017, in relation to their positions in the ranking provided by the Complete University Guide (CUG). The authors discovered that a university's position in the ranking affected all universities except those with historically established reputations (i.e., elite universities), even after controlling for previous levels of financial resources and institutional differences. They proposed that this relationship may be explained, in part, by universities' revenues from tuition fees. Similarly, Kim (2018) analysed U.S. News and World Report's Best Colleges ranking to explore the relationship between universities' positioning in the ranking and their available financial resources. The study revealed that the distinct ranking criteria encouraged increased spending on teaching and non-teaching activities, including academic and student services.

While criticisms of university rankings persist, the need for universities' effective performance evaluation has spurred alternative approaches to leveraging ranking data. Daraio et al. (2023) proposed the use of the Leiden Ranking to identify top-performing institutions ("outliers") in terms of research quality. Additionally, methodologies such as efficiency analysis (Bougnol & Dulá, 2006; Daraio et al., 2015) and multicriteria analysis (Ho et al., 2006; Sarrico & Dyson, 2000) have been proposed.

2.2 Italian university system

Italian universities have a rich legacy of academic and research excellence, and they are internationally recognised for their quality and commitment to innovation and advanced education. Nonetheless, the university system in Italy faces several challenges, including a persistent funding deficit.

Italian universities are primarily funded by the central government, via the Ministry of Universities and Research (MUR), which is meant to cover universities' operating expenses. However, in recent years, the Italian university system has suffered a series of budget cuts that have profoundly impacted its ability to uphold rigorous academic standards and maintain a stimulating academic environment. Insufficient funding has also led to a shortage of academic jobs, reduced resources for research and innovation, and diminished educational opportunities for students (Turri, 2016). Thus, the nexus between Italian universities' financial resources and their reputations and research performance represents a relevant area of research. The Italian university system comprises 97 universities, of which 67 are state universities, 19 are legally recognised non-state universities and 11 are legally recognised non-state telematics universities. Despite this expansive network, not all Italian universities are represented in international academic rankings. In fact, the 2019 edition of the ARWU included only 46 Italian universities. This underrepresentation extends to other rankings, as well. For instance, the latest Leiden Ranking included only 47 Italian universities.

2.3 Gap in the literature and contribution of the present work

Despite the significant contributions of previous research on university funding, academic performance and rankings, there remains a notable gap in the literature regarding the integration of different approaches to gain a more comprehensive and in-depth understanding of the relationship between universities' positioning in academic rankings and their available financial resources.

Most prior studies have explored the effect of funding on academic performance or university efficiency separately, considering academic rankings as independent indicators. Indeed, despite the acknowledged limitations of these rankings (as previously discussed), they continue to be widely used by policy makers. The present research aimed at partially filling this gap in the literature by developing a methodology combining DEA and the Louvain community detection technique, and applying it to a prominent academic ranking in order to deepen our understanding of the relationship between academic rankings and universities' available financial resources. Specifically, the research aimed at determining the financial resources required for universities to achieve international competitiveness and high positions in the ranking. Additionally, it aimed at identifying groups of comparable universities based on financial resources, to identify those with the greatest efficiency (i.e., those maximising their position in the ranking relative to their available financial resources).

The use of DEA enabled us to identify universities achieving higher levels of efficiency. Subsequently, the Louvain algorithm allowed us to establish groups of universities with similar levels of funding and academic performance, facilitating a deeper analysis of higher education dynamics. Such communities of universities were identified both internationally and nationally, with respect to the Italian context. By integrating both qualitative (i.e., university positioning in the academic ranking) and quantitative (i.e., university efficiency and heterogeneity) factors, our approach represents a novel contribution to the field. As we will describe in more detail in the "Methodology and data" section, to the best of our knowledge,

this study represents the first application of a combination of DEA and the directed Louvain community detection algorithm to identify communities of comparable universities based on their available financial resources and performance in an academic ranking.

3 Methodology and data

In this section, we will introduce DEA and briefly explain our process of constructing a DEA network and applying the Louvain community detection method in the present study. Subsequently, we will present the dataset that was utilised for the university analysis.

3.1 DEA

DEA is a non-parametric method that is used to assess the productivity or cost efficiency of a decision making unit (DMU) in comparison with other similar units. DMUs are the units of analysis, which can range from companies to departments or, in the present case, universities. First introduced by Charnes et al. (1978), DEA aims at evaluating the operational efficiency of a given sample of DMUs near the efficient boundary of their production set. It is a non-parametric approach that is capable of managing multiple inputs and outputs, assuming free disposability (i.e., the ability to destroy assets without cost) and convexity.

In input-oriented DEA, the objective is to evaluate the potential reduction in inputs a DMU can achieve while maintaining its current level of outputs and production technology. Essentially, this assesses how close a DMU can move towards an estimated efficient frontier, representing best practices observed across all DMUs under the same technological constraints.

Conversely, output-oriented DEA focuses on maximising the output levels a DMU can achieve while keeping the number of inputs and amount of technology constant. In the present analysis, we adopted an output-oriented DEA because we were interested in assessing the potential expansion of a university output (i.e., ARWU score) while holding the inputs (i.e., financial resources) and technology constant.

In the DEA formulation of Charnes et al. (1978), the frontier is modelled assuming constant returns to scale (CRS) and DMUs operating at an optimal scale. Later, Banker et al. (1984) introduced a DEA model allowing for variable returns to scale (VRS)-a development that has been seminal for the success of the DEA approach, as described in Nepomuceno et al. (2023). Formally, the DEA VRS output-oriented linear programming problem can be expressed as follows:

$$max\ \phi$$
$$s.t.\ \sum_{j=1}^{n} \lambda_j x_{ij} - x_{i0} \leq 0\ \forall i, i = 1, ..., p$$
$$\sum_{j=1}^{n} \lambda_j y_{rj} - \phi y_{r0} \geq 0\ \forall r, r = 1, ..., q$$
$$\sum_{j=1}^{n} \lambda_j = 1$$
$$\lambda_j \geq 0$$

Where ϕ represents efficiency, $x \in \Re_+^p$ is the vector of p inputs and $y \in \Re_+^q$ is the vector of q outputs, with λ_j denoting the weight of peers and $\lambda = (\lambda_1, ..., \lambda_n)$ is the vector of the weights and n is the number of DMUs. For the sack of clarity, we indicate the optimal values of λ_j after the optimization as λ_j^*.

Consistent with Kneip et al. (2016), we applied tests to assess the convexity hypothesis required by the DEA model and the returns to scale, with the aim of empirically determining whether the hypotheses were empirically supported by the data. These tests, implemented in FEAR (Wilson, 2008), were conducted on our empirical datasets, confirming the coherence of a VRS-DEA model (see Supplementary Material A for details).

While much of the literature focuses on ϕ (i.e., the efficiency score), our analysis focused on the weights λ_j^* which reflect the relative importance of different peers in establishing an efficiency benchmark. The literature on the linear programming (LP) problem in DEA is extensive, with considerable discussion of the constraints and weights restriction in DEA models. Notable contributions to this discussion include works by Pedraja-Chaparro et al. (1997), Podinovski and Athanassopoulos (1998), Joro and Viitala (2004) and Podinovski (2016), as well as various chapters in handbooks on DEA (e.g., Cooper et al. 2011; Podinovski, 2015; Thanassoulis et al., 2004). While the assessment of LP values is important, also in the DEA benefit of doubt (BOD) model (see, e.g., Lavigne et al., 2019), use of λ values for the representation and visualisation of the network among DMUs has been less explored (see Sect. 3.2 for details). Leveraging the representation and visualisation of λ_j^* values enables not only an evaluation of their "numerical" values, but also the relation between inefficient and efficient DMUs, thereby representing all DMUs in a network.

Of note, the λ_j^* values obtained from a DEA formulation are generally not unique, as highlighted in Banker et al. (1984). The problem of multiple optimal solutions is well-discussed in the literature, with Banker and Thrall (1992) and Seiford and Zhu (1998) noting its significance in DEA CRS cases but not in DEA VRS models, as applied in our analysis. Nevertheless, we conducted tests of robustness on our empirical applications to assess λ_j^* variation as the order of the analysed DMUs changed. These tests aimed at "forcing" the solver of the linear programming problem to provide different optimal solutions based on the DMU order. In more detail, they involved randomly altering the DMU order in our two datasets 100 times, performing DEA for each of the 100 random sets and extracting the calculated λ_j^* values. The resulting λ_j^* values were consistent up to the 12th decimal place, primarily due to numerical approximation issues, indicating unique solutions. Although the results of these tests are not reported in this paper (for brevity), they are available from the authors upon request. For DEA CRS models, the issue of multiple λ_j^* values is more relevant, and it may be addressed using the recently proposed bootstrap approach by Kang et al. (2024).

3.2 From DEA to peers communities

The DEA analysis allowed us to construct a directed weighted graph G, relating each DMU to its peers and reference weights (i.e., a relational network). Formally, given two DMUs a and b, an edge is created if a is the peer of b, with the weight of this edge determined by the λ_j^* estimated in the DEA linear programming problem. Figure 1 illustrates an example of graph creation from the DEA linear programming results. The example includes five DMUs (i.e., the five nodes) and five weighted arcs.

The graphical representation enabled us to apply network analysis techniques to gain deeper insight into the relationships between nodes. Our research objective was to identify similar peers in terms of funding, while also considering positionality in the ranking. Community identification appeared to be the best approach for this, because universities function as communities of academics, non-academic staff and students. Communities are highly inter-connected sets of nodes (Fortunato & Castellano, 2007), and their identification can reveal similar structures-in the present case, in terms of the resources used in the production process
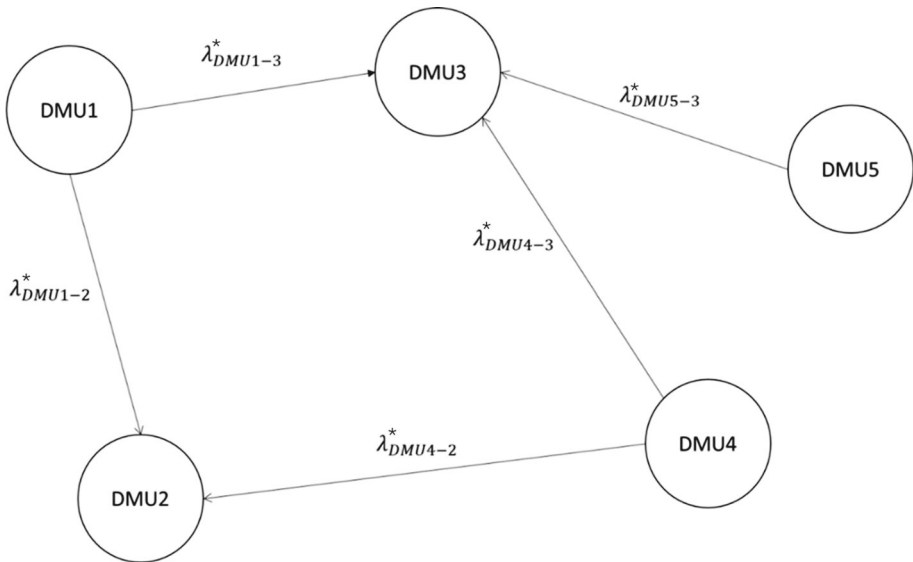
Fig. 1 Example of a directed network of weights and peers. Source: own elaboration

(inputs: long-term physical capital and total operating revenues) to generate an ARWU score (output), as modelled in the DEA. Through community detection, we aimed at identifying the network of universities. We considered densely connected nodes indicative of communities, and weakly connected nodes indicative of separate communities.

While the combination of network analysis and DEA has been previously proposed in the literature, existing methods do not aim at identifying communities in directed weighted graphs derived from DEA. Instead, the main approaches consider the network as a social network and apply social network analysis (SNA) techniques. Liu et al. (2009) were among the first to combine DEA with SNA to differentiate efficient DMUs. They employed a two-step method, constructing a directed and weighted network based on the DEA results and subsequently calculating eigenvector centrality to rank the analysed DMUs. To resolve the convergence issues associated with eigenvector centrality measures, de Blas et al. (2018) proposed a new perspective on the DEA-SNA network. They considered the DEA result as a bipartite directional graph and applied the hyperlink-induced topic search (HITS) algorithm (Kleinberg, 1999) to identify peers as efficient units and hubs as inefficient units, thereby defining a "new" rank. However, one limitation of Blas et al.'s method is that only efficient DMUs are considered candidates for benchmarking units.

While these approaches attempt to address the challenge of creating partitions or communities, we did not consider them suitable for our analytical context, as we aimed at modelling not only the weight of relationships, but also their direction (i.e., from inefficient to efficient units). Furthermore, we sought to identify communities of similar DMUs without modelling new "virtual" DMUs, as proposed by the previous authors. Thus, we opted to instead utilise one of the primary methods available for identifying communities within a network: the Louvain community detection algorithm (Blondel et al., 2008), which optimises modularity as it progresses. Modularity is a scaling value between -1 (non-modular clustering) and 1 (fully modular clustering), measuring the relative density of edges within communities compared to edges outside communities. For an undirected weighted graph, modularity is defined as:

$$Q = \frac{1}{2m} \sum_{s,v} [A_{sv} - \frac{k_s k_v}{2m}] \delta(c_s, c_v)$$

where A_{sv} represents the weight of the edge between s and v, $k_s = \sum_{s,v} A_{sv}$ is the sum of the weights of the edges attached to vertex s, c_s is the community to which vertex s is assigned, the δ function $\delta(c_s, c_v)$ is 1 if c_s and c_v are the same community and 0 otherwise and $m = \frac{1}{2} \sum_{s,v} A_{sv}$.

The Louvain method proceeds iteratively in two steps, in order to efficiently increase modularity. First, each node is assigned to a different community, resulting in as many communities as there are nodes. Second, the neighbours s of node v are evaluated to determine the gain in modularity that would occur by removing s from its community and placing it in the community of v. If the gain is positive and maximum, s is moved accordingly; otherwise, s remains in its current community. Thus, a node is only moved if the modularity gain is positive (i.e., if moving the node improves overall modularity) and the modularity gain is the highest among all possible options (i.e., if moving it elsewhere would not lead to greater improvement).

This process is continually applied for each node until no further improvement can be achieved. At this point, the first phase is complete. Of note, the algorithm output, in terms of found communities and calculation time, may depend on the order in which the nodes are considered. However, the authors (Blondel et al., 2008) assert that the order of the nodes does not significantly influence the obtained modularity. To ensure the robustness of the output concerning node order, we randomly generated several node orderings and ran the algorithm for each one. In all cases, we consistently found the same communities, ensuring no uncertainty in terms of the community allocation. The second phase of the algorithm consisted of building a new network, with nodes representing the communities found in the first phase. The weights of the links between the new nodes were determined by summing the weight of the links between nodes in the corresponding two communities. Links between nodes in the same community resulted in self-loops for that community in the new network. Once this second phase was complete, it became possible to reapply the first phase of the algorithm to the resulting weighted network and to iterate. However, in our case, the version proposed by the authors presented a fundamental problem, as the concept of modularity was only formulated for weighted undirected graphs.

In our scenario, as demonstrated above, we were analysing a directed weighted graph. If we were to have performed a naive directed-to-undirected graph transformation, disregarding edge directionality and treating the graph as undirected, two main biases could have arisen (Malliaros & Vazirgiannis, 2013): data ambiguity (i.e., deviations in the clustering results) and biases in the clustering results (by discarding edge directionality, valuable information would not be used in the clustering process, generating deviations in the results).

To explain these biases, consider a network in which the connections represent universities. The relationship between efficient and inefficient universities would not be reciprocal; efficient universities would serve as "peers" (or best-performing units), while inefficient universities would look to efficient ones to improve their performance. This would necessitate the use of directed graphs, since we could not disregard the possibility of one-sided relationships among universities. Failure to account for such one-sided relationships and naively using undirected graphs could lead to the following biases:

1. *Data ambiguity*: When considering two universities in close proximity within the network, one efficient and the other inefficient, it would not be specified which must learn from the other. They would have a reciprocal link, creating ambiguity that may distort the results.

2. *Biases in clustering results*: Without a clear distinction between efficient and inefficient universities through directed graphs, we may misidentify groups of universities, underestimate the influence of certain universities, overestimate the influence of others and draw incorrect conclusions about network dynamics.

To avoid these biases and successfully identify communities, we adopted the directional version of Louvain's algorithm (Dugúe & Perez, 2015). This version, which we call the directed Louvain community detection algorithm, differs from the classical algorithm presented above (Blondel et al., 2008) through the modularity gain applied (with all other steps of the algorithm remaining unchanged). The modularity gain considered in the classical Louvain algorithm does not account for direction, whereas the directional version does, according to the following formula:

$$Q = \sum_c [\frac{L_c}{m} - \gamma \frac{k_c^{in} k_c^{out}}{2m}^2]$$

where c represents a community, L_c is the number of intra-community links for community c, k_c^{in} and k_c^{out} are the sum of in-going and out-going degrees of the nodes in community c, and γ is the resolution parameter (typically equal to 1, as in our case). For more detail on the γ selection see Newman (2016). From the results of the community detection, we were able to identify "similar" communities of inefficient and efficient units (i.e., peers). Furthermore, by maintaining the orientation of the graph, we were able to understand which efficient unit(s) exerted the most influence on their respective communities.

3.3 Dataset

All available information in the 2019 ARWU (https://www.shanghairanking.com/rankings/arwu/2019) was collected for the 318 universities in our sample (from five countries). The ARWU is based on six indicators:

- Alumni: Number of alumni who have won Nobel Prizes and Fields Medals.
- Award: Number of staff who have won a Nobel Prize in Physics, Chemistry, Medicine or Economics, or a Fields Medal in Mathematics.
- HiCi: Number of highly cited researchers selected by Clarivate Analytics.
- N&S: Number of papers published in *Nature* and *Science* (considering only articles).
- PUB: Number of papers indexed in the *Science Citation Index-Expanded* and *Social Science Citation Index*.
- PCP: Weighted scores of the above five indicators, divided by the number of full-time equivalent academic staff.

In the ARWU, the indicators N&S, HiCi and PUB represent a university's research performance, while the variables Alumni and Award can be seen as proxies for a university's reputation. For the final position of a university ***uni*** in the ranking, ARWU calculates the total score as a weighted sum of these indicators (https://www.shanghairanking.com/methodology/arwu/2019):

$$ARWU\,Score_{uni} =$$
$$0.1 Alumni_{uni} + 0.2 Award_{uni} + 0.2 HiCi_{uni} + 0.2 N\&S_{uni} + 0.2 PUB_{uni} + 0.1 PCP_{uni}$$

ARWU scores range from 0 to 100, with 100 denoting the top position in the ranking. In addition to the ARWU data, we also gathered information from universities' financial statements for the fiscal year 2019. Specifically, we collected the net value of long-term physical capital

Table 1 International sample of universities (318 observations): Correlation matrix between the collected variables

	ARWU score	Physical capital	Revenues
ARWU score	1		
Physical capital	0.79	1	
Revenues	0.85	0.84	1

Source: own elaboration based on data from ARWU and balance sheets

from university balance sheets and total operating revenues from university income statements, while harmonising the different accounting systems to eliminate external revenues (e.g., hospital activities) and extra activities that might influence and distort the comparison between universities. For additional discussion of the problems created by accounting system heterogeneity and the harmonisation that is required prior to comparative analysis, see Avenali et al. (2024).

The net value of long-term physical capital (henceforth referred to as "physical capital") represented the long-term assets used to support universities' production processes while fulfilling their institutional missions. This encompassed the net value of property (e.g., land, buildings, real estate), plants, equipment, machinery, vehicles, intangible assets (e.g., software licenses, databases) and goodwill. It was calculated as the value of intangible and tangible fixed assets, net accumulated depreciation and amortisation, as extracted from university balance sheets. All values were converted into euro using the exchange rate provided by the Central Bank of Europe on 31 December 2019.

The total operating revenues (henceforth referred to as "revenues") represented the ongoing overall returns universities received in a year to support their institutional mission-related operational activities. This information was extracted from university income statements. In cases where universities engaged in hospital activities or extra activities (e.g., external laboratories), returns from these activities were identified from the income statement and subtracted from the revenues, to ensure data homogeneity. Once again, revenues were expressed in euro.

Of note, while the ARWU includes more than 1,000 universities, the financial statements of individual universities are not always public or accessible. Thus, it was not possible to collect financial data for all of the ranked universities. However, the dataset included universities from various countries, such as the US (171 universities), UK (61 universities), Italy (46 universities), Australia (32 universities) and Poland (8 universities). Additionally, we extended the dataset for the Italian case by collecting information from the open data platform of the Italian Ministry of Universities and Research, USTAT (http://ustat.miur.it/). Specifically, for each Italian university, we collected data on the number of academic staff (headcount), number of enrolled students (at all educational levels), teaching load (i.e., ratio of students to academic staff) and student fees. We noted high correlations among the variables considered, for both the international (Table 1) and the Italian universities (Table 2). In both cases, there was a correlation greater than 0.8 between revenues and ARWU score. Regarding the relationship between physical capital and ARWU score, a very high correlation was observed in the international case (0.79), which dropped to 0.57 in the Italian case For the Italian universities, we also examined the correlation between ARWU score and the other collected variables. We found a high correlation between ARWU score and the number of academic staff, a medium correlation between student fees (a subset of revenues) and ARWU score and a very low correlation between teaching load and ARWU score (0.12). Table 3 presents the descriptive statistics of the variables collected for the international analysis, while Table 4 reports the descriptive statistics of the variables collected for the Italian case.

Table 2 Italian sample (46 universities): Correlation matrix between the collected variables

	ARWU score	Physical capital	Revenues	Student fees	Academic personnel	Teaching load
ARWU score	1					
Physical capital	0.57	1				
Revenues	0.81	0.65	1			
Student fees	0.58	0.43	0.75	1		
Academic personnel	0.77	0.6	0.95	0.85	1	
Teaching load	0.12	0.4	0.4	0.12	0.29	1

Source: own elaboration based on data from ARWU and USTAT

Table 3 International sample of universities (318 observations): Descriptive statistics on the inputs (physical capital (x_1) and revenues (x_2)) and the output (ARWU score (y)). Physical capital and revenues expressed in euro

Variable	25th perc.	Median	Average	75th perc.	Std.
ARWU score (y)	8.34	12.41	16.48	18.47	12.62
Physical capital (x_1)	352,280,621	713,865,415	1,054,286,909	1,466,095,518	1,107,559,709
Revenues (x_2)	258,641,092	480,024,267	803,330,956	902,927,035	1,018,710,964

Source: own elaboration based on data from ARWU and balance sheets

Table 4 Italian sample (46 universities): Descriptive statistics on the inputs (physical capital (x_1) and revenues(x_2)), the output (ARWU score (y)) and other variables considered in the analysis

Variable	25th perc.	Median	Average	75th perc.	Std.
ARWU score (y)	7.57	9.31	10.60	11.74	3.89
Physical capital (x_1)	89,782,452.40	152,039,606.00	199,735,091.00	269,927,318.30	165,066,014.00
Revenues (x_2)	130,994,945.00	208,248,230.00	248,997,317.00	300,926,415.50	172,641,476.00
Student fees	17,359,470.20	30,528,771.30	40,928,994.00	39,021,096.01	42,927,912.50
Academic personnel	878.25	1159.50	1476.15	1863.75	987.20
Teaching load	16.33	20.97	20.47	24.81	6.51

Source: own elaboration based on data from ARWU and USTAT
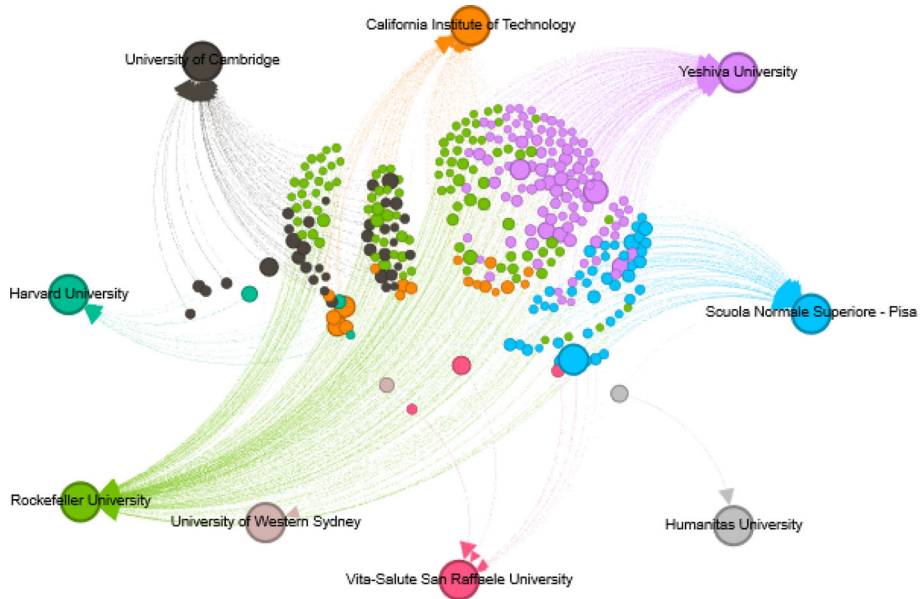
Fig. 2 Community detection network identified using the directed Louvain algorithm on 318 universities. Each colour represents a community (nine communities in total). Labelled nodes represent the efficient units in their respective communities. Source: own elaboration based on DEA and the directed Louvain analysis network. (Color figure online)

4 Results

In this section, we outline the main results of our analysis. The DEA-VRS (Banker et al., 1984) model considered physical capital (x_1) and revenues (x_2) as inputs and ARWU score as the output (y). The basic assumption was that these inputs were the main factors determining universities' ARWU scores. Moreover, by analysing academic rankings and comparing the scores with those obtained using DEA, Bougnol and Dulá (2006) found important equivalences between the DEA and academic rankings, leading them to conclude that "DEA is a suitable tool for these types of studies" (Bougnol & Dulá, 2006). Employing an output-oriented approach, given the inputs (i.e., financial resources), we aimed at maximising universities' ARWU scores and identifying peers (i.e., best in class). Subsequently, these peers were analysed using the directed Louvain community detection algorithm to identify communities of universities with similar financial resources and ARWU scores.

4.1 International comparison

Figure 2 illustrates the DEA and directed Louvain analysis network. We identified nine distinct communities (Community A: aqua green; Community B: black; Community C: orange; Community D: sand; Community E: green; Community F: violet; Community G: red; Community H: grey; Community I: light blue). Each colour in the figure corresponds to a different community (also denoted by capital letters, as specified in Table 5), and the labelled nodes represent the efficient/reference units within their respective communities (i.e., the efficient DMUs "pointed" to by the directional arcs from the inefficient units).

Table 5 Descriptive statistics on the communities identified in the international sample

Community	Colour in Figs. 2 and 3	Number of DMUs	Mean physical capital	Mean revenues	Mean ARWU score	Reference DMU
A	Aqua green	4	6,049,832,760.00	5,877,664,385.00	66.59	Harvard University
B	Black	37	2,813,239,459.00	1,992,333,948.00	33.64	University of Cambridge
C	Orange	19	1,394,417,525.00	2,282,171,655.00	31.14	California Institute of Technology
D	Sand	2	216,030,524.40	1,045,809,086.00	17.52	University of Western Sydney
E	Green	102	1,221,794,498.00	691,512,590.10	15.51	Rockefeller University
F	Violet	104	426,216,549.20	337,160,659.10	10.61	Yeshiva University
G	Red	4	69,295,754.20	220,205,929.00	10.03	Vita-Salute San Raffaele University
H	Grey	2	29,420,685.50	32,049,860.50	8.83	Humanitas University
I	Light Blue	44	244,601,540.20	141,714,426.40	8.18	Scuola Normale Superiore-Pisa

Source: own elaboration based on DEA and the directed Louvain analysis network

Description of the Communities
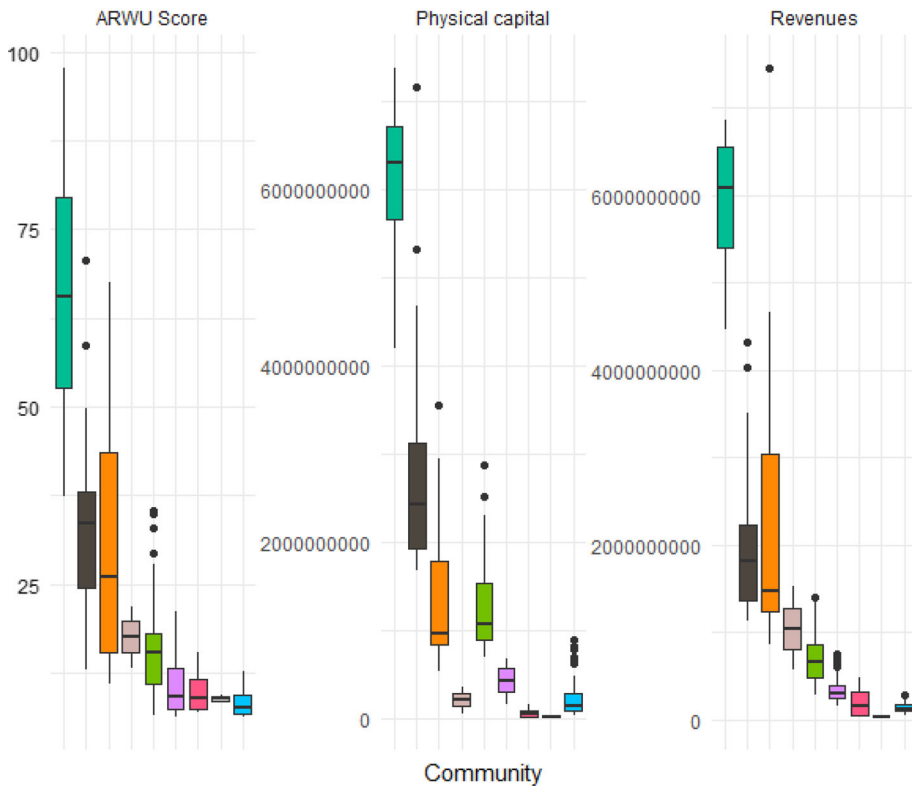
Fig. 3 Boxplots of the variables analysed using DEA for the international case (from the left to the right: ARWU score=y, physical capital=x_1 and revenues=x_2). Colours represent the different communities identified (as reported in Table 5). Source: own elaboration based on DEA. (Color figure online)

Table 5 and Fig. 3 report the summary statistics of the identified communities. Details of the results per university are provided in Supplementary Material B.

The reported information highlights the heterogeneity in universities' available resources, in relation to their ARWU scores. In particular, Community A, comprising four universities (i.e., Columbia University, Harvard University, New York University, Stanford University), emerged as the community with both the highest average ARWU score and the greatest financial resources, in terms of both average physical capital and average revenues. In this community, Harvard University served as the DMU of reference, while also claiming the top position in the overall ranking. In this analysis, each community had a similar reference peer in terms of resources. Therefore, in contrast to the classic DEA analysis, which assigns one or more reference peers to an inefficient DMU, in this scenario, each university had a "most similar" peer to emulate.

The identified communities also revealed a proportional relationship between revenues and ARWU score, with communities' average ARWU scores increasing in line with their average revenues. However, this correlation did not hold true for physical capital. In fact, Community D and Community I had very similar average values of physical capital (216,030,524.40 for Community D, 244,601,540.20 for Community I), yet significantly different ARWU scores

(17.52 for Community D, 8.18 for Community I). This suggests that universities' annual economic resources (i.e., liquid funds) may have a more significant influence on positionality in academic rankings. In contrast, long-term assets appear less impactful, as they cannot be easily changed within the same year (e.g., the acquisition of real estate or intangible assets typically requires a large investment spread over several years). Consequently, this initial finding implies that a greater concentration of total operating revenues in fewer universities could yield benefits, provided that the new resources are effectively utilised and managed.

Another noteworthy finding is the considerable disparity in the number of universities per cluster, possibly influenced by the presence of outliers. For instance, Community D comprised only two universities, whereas Community F included 104. Theoretically, the DEA analysis could return an isolated cluster consisting of only one efficient unit.

The results presented here, regarding communities of diverse membership and financial resources, complement the findings on the structural heterogeneity of universities in terms of size, disciplinary subject mix and service orientation (i.e., the so-called "third mission"). Further discussion of the heterogeneity of European universities can be found in Bruni et al. (2020). As acknowledged in the literature, DEA is susceptible to outliers, which can affect the frontier and consequently influence all efficiency estimates. However, the present research aimed at identifying the best-performing universities as examples (i.e., peers) for other universities to emulate. The community clustering results reflect the reality of university funding. It would be interesting to explore, in future iterations of this research, how these communities might change if the impact of peers is mitigated through the use of robust efficiency analysis (for an introduction, see Daraio & Simar, 2007).

In our case, even the smallest communities maintained a connection (as evidenced by the graph in Fig. 2 showing no isolated nodes). Consequently, universities with very different resources were compared. Potentially, it would have been more equitable to consider or segment the ranking according to the resources available to universities. In this vein, some steps towards incorporating financial resources and promoting greater "fairness" in academic rankings have been taken, as evidenced by the THE, which, in its current version, also considers university income.

The following section will focus on the Italian case. This change in the layer of analysis, transitioning from an international to a national context, will enable us to further reflect on the results of this initial analysis, considering a specific national system.

4.2 Insight into a single university system: The Italian case

As in the international analysis, for the national case, we applied the DEA technique and the directed Louvain algorithm to the directed graph of peers and weights for the 46 Italian universities in the 2019 ARWU. Figure 4 displays the resulting graph with the identified communities. The analysis revealed three main communities and two smaller communities (one of which was practically isolated): a community of large, generalist universities (in blue), comprising nine universities, with *Sapienza University of Rome* serving as the reference DMU; a community of medium-sized universities (in light green), encompassing 14 universities, with *University of Pisa* serving as the reference DMU; a community of small, research-oriented universities (in purple), consisting of 18 universities, with *Scuola Normale Superiore - Pisa* serving as the reference DMU; a small community (in dark green) comprising two universities (*Humanitas University*, *International School for Advanced Studies*), with *Humanitas*
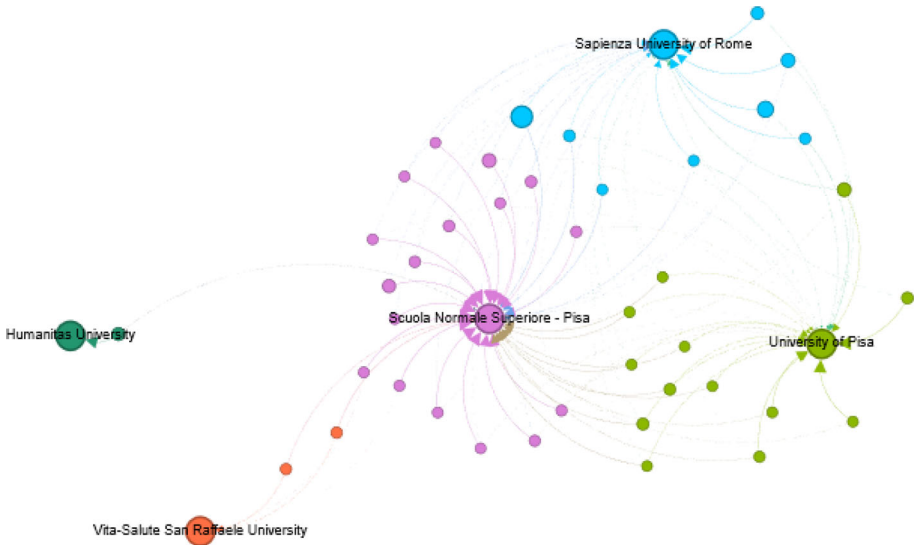
Fig. 4 Italian case: Graphical representation of communities. Labelled nodes represent the efficient units in their respective communities. Colours represent individual communities identified using the directed Louvain method. Source: own elaboration based on DEA and the directed Louvain analysis network. (Color figure online)

University serving as the reference DMU; and another small community (in orange) comprising three universities (*Vita-San Raffaele University*, *Free University of Bozen-Bolzano*, *Tuscia University*), with *Vita-San Raffaele University* serving as the reference DMU.

Detailed results for each Italian university, including their efficiency score, ARWU score and community affiliation, are presented in Supplementary Material C.

An analysis of the resources used in relation to the ARWU score (Fig. 5) revealed similar findings as those observed for the intentional case: universities with higher revenues tended to achieve higher ARWU scores, with physical capital (despite exhibiting a higher average in the case of the large, generalist (blue) community) emerging as less influential than revenues. Moreover, when comparing the identified communities with the number of students enrolled in the various universities (Fig. 6) an indicator that is often used as a proxy for university size-it became evident that the different communities were associated with different university sizes. In particular, the blue community represented the largest universities, the green community represented medium-sized universities and the purple community represented small universities. The orange and dark green communities, on the other hand, represented very small universities. A similar trend was observed for the differences in academic staff across communities (Fig. 7). A further interesting finding emerged, regarding geographical distribution (Fig. 8). Despite the widespread notion of a sharp division between northern and southern Italy across various dimensions (e.g., income, development, infrastructure) (see, e.g., Abramo et al., 2016), our analysis did not reveal a significant difference in community affiliation between universities based in northern, central or southern regions. This finding aligns with a previous result obtained from a Leiden Ranking analysis, also focused on the Italian case (Daraio et al., 2023). In conclusion, our analysis demonstrated that Italian universities with greater resources also achieved higher positions in the ARWU ranking.
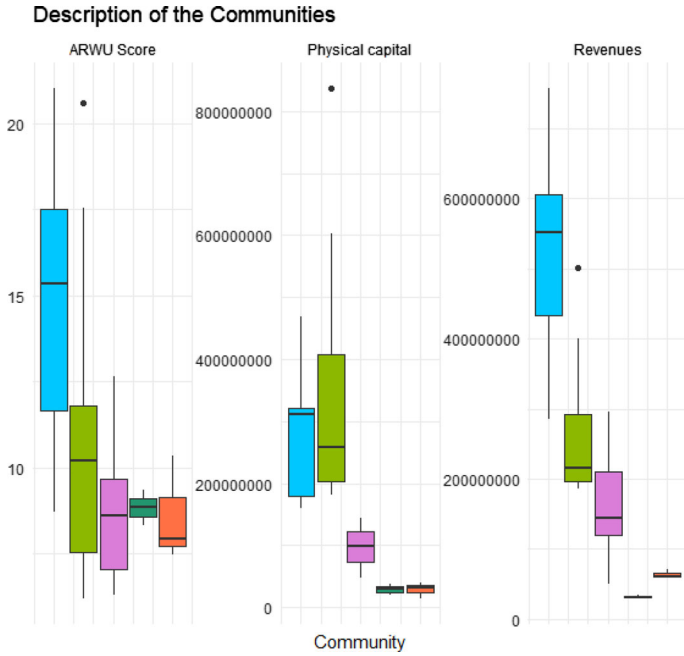
Fig. 5 Boxplot of the variables analysed using DEA for the Italian case (from the left to the right: ARWU score=y, physical capital=x_1 and revenues=x_2). Colours represent the different communities identified. Source: own elaboration based on DEA. (Color figure online)
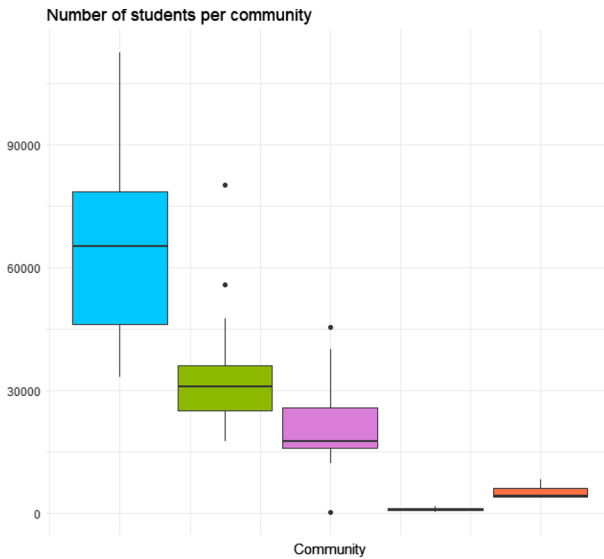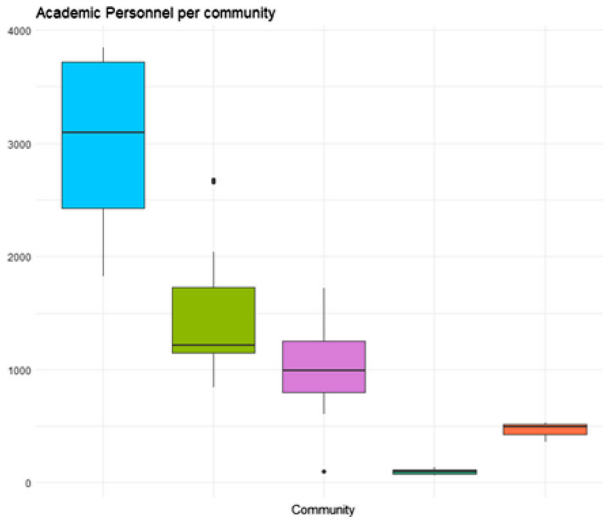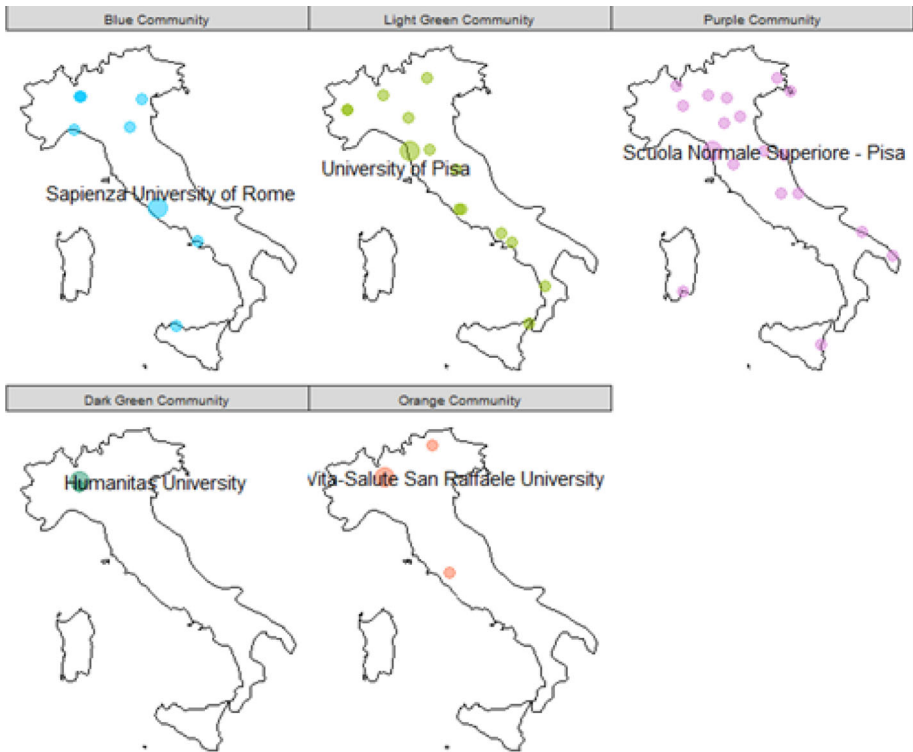

Fig. 6 Boxplot of Italian university student enrolment. Colours represent the different communities identified. Source: own elaboration based on DEA and the directed Louvain analysis network. (Color figure online)

Fig. 7 Boxplot of Italian communities' number of academic personnel. Colours represent the different communities identified. Source: own elaboration based on DEA and the directed Louvain analysis network. (Color figure online)

Fig. 8 Geographical representation of the number of Italian universities per community. Labelled nodes represent the efficient units in their respective communities. Colours represent the different communities identified. Source: own elaboration based on DEA and the directed Louvain analysis network

5 Conclusions

In this paper, we

i) introduced a novel approach combining the DEA-VRS methodology with the (directed) Louvain algorithm to identify communities of universities with similar financial resources and ARWU scores;

ii) demonstrated that universities' positioning in an academic ranking is significantly influenced by their available financial resources;

iii) grouped universities with similar financial resources and similar positions in the ranking;

iv) identified universities with efficient utilisation of financial resources as practical targets for other universities to emulate; and

v) raised concerns regarding the fairness of the ranking.

The results should be considered heuristic, requiring further verification and systematic analysis. However, they have the potential to stimulate further study and provide greater insight into university rankings and university management, pointing to alternative ways of allocating research funds.

We analysed 318 international universities and a sample of Italian universities to explore the prevalent role played by financial resources in determining universities' positioning in the ARWU ranking. Annual revenues appeared more significant than physical capital in determining higher ARWU scores. Subsequently, we explored the Italian case, identifying how resources and ranking scores aligned with university size, in terms of student enrolment (a dimension not considered by the ARWU), the number of academic personnel and geographical location. No significant differences were observed in terms of teaching load or geographical location in the Italian territory.

The results offer valuable insight into the role played by financial resources in determining university positioning within the ARWU. They also prompt further reflection on the fairness of comparing universities with vastly different financial resources. Moreover, we emphasise that the proposed methodology is not intended to provide a definitive evaluation or ranking of universities, but rather to compare universities (i.e., peers) and generate key insights into the dynamics of resource management and how such dynamics impact university positioning in academic rankings.

From a policy perspective, our results may prompt policy makers to reconsider criteria for resource allocation, particularly in countries with limited public spending per capita on higher education and research (e.g., Italy). Indeed, the concentration of resources may play a key role in determining both international and national competition. Thus, policy makers could opt to target efficient analysis units (i.e., those with DEA scores of 1) for optimal resource utilisation. Alternatively, they might consider optimising and revising processes at less efficient universities, to improve resource allocation. Furthermore, policy makers could choose to allocate funding disproportionately to specific communities (e.g., in the Italian case, the purple community), to amplify the impact of funding concentration and propel universities in those communities up the ranking. However, it is important to recognise that improvements reflected by ARWU indicators may not necessarily persist in subsequent years, due to various factors, including academic processes and university inefficiencies or resource gaps. Moreover, any changes in resource allocation or increased funding should be systematic, rather than "one-off," as fluctuations in resources may result in universities reverting to previous situations. Furthermore, our findings hold potential utility for university management tasked with devising reliable strategies for moving up in academic rankings and making the institution more competitive and efficient. In particular, managers may draw

inspiration from the best practices observed in leading universities within their respective communities (i.e., the most efficient universities), thereby realistically improving their ability to compete and, consequently, their positions in academic rankings.

One weaknesses of the proposed approach is that the optimal λ_j^* values in the DEA formulation were not necessarily unique. This is a critical consideration, as our approach relied on these values to construct a directed weighted graph for further investigation using network analysis. As explained in the "Methodology and data" Section, the issue of multiple λ_j^* values primarily affects DEA CRS models. Thus, it did not affect the DEA VRS model applied in the present research. Nonetheless, we conducted robustness tests to confirm that our application yielded a single optimal solution. Should this problem arise in future research, it could be addressed using the bootstrap approach proposed in the recent paper by Kang et al. (2024). Another limitation concerns our modelling of revenues as an input of the analysed process. Traditionally, revenues are considered as an output of the production process of universities, whereas we have assumed that the financial resources available to an institution (proxied by its revenues) directly influence its ability to increase its ARWU ranking. Future works could explore the use of operating costs as an input instead of revenues or carry out comparative analyses of models that use revenues as an input versus those that use it as an output.

A potential limitation of this work concerns the influence of possible outliers on the DEA efficiency scores. Although our analysis focused on extremely high-performing universities (considered peers, and therefore exemplars for inefficient universities to emulate), we believe it would be beneficial to investigate and attenuate the impact of outliers using the robust methods developed in efficiency analysis. This remains a consideration for future iterations of this work.

Additional avenues for future research could involve decomposing revenues between public and private shares, incorporating other financial variables in the DEA model and expanding the dataset to encompass (almost) all ARWU universities (which is not currently possible, due data limitations). Moreover, the proposed approach could be applied to other university rankings (e.g., the THE or QS), to ascertain the robustness of the present results across rankings.

Finally, in a subsequent analysis, it would be interesting to develop conditional DEA models considering financial resources as control variables. This would enable the impact of financial resources on university performance to be estimated, facilitating the identification of appropriate peers for individual universities, based on the output of the conditional DEA model.

Availability of data and materials The data that support the findings of this study are available from the corresponding author upon request.

Code Availability Not applicable.

Declarations

Conflict of interest All authors certify that they have no affiliations with or involvement in any organization or entity with any financial or non-financial interest in the subject matter or materials discussed in this manuscript.

Financial interests: The authors declare that they have no financial interests.

Ethics approval Not applicable.

Consent to participate Not applicable.

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

Abramo, G., D'Angelo, C. A., & Rosati, F. (2016). The north-south divide in the Italian higher education system. *Scientometrics, 109*, 2093–2117.

Avenali, A., Daraio, C., Di Leo, S., & Wolszczak-Derlacz, J. (2024). Heterogeneity of national accounting systems, world-class universities and financial resources: What are the links? *Journal of Informetrics, 18*(2), 101502.

Baltaru, R. D., Manac, R. D., & Ivan, M. D. (2022). Do rankings affect universities' financial sustainability? Financial vulnerability to rankings and elite status as a positional good. *Studies in Higher Education, 47*(11), 2323–2335.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science, 30*(9), 1078–1092.

Banker, R. D., & Thrall, R. M. (1992). Estimation of returns to scale using data envelopment analysis. *European Journal of Operational Research, 62*(1), 74–84.

Benito, M., Gil, P., & Romera, R. (2019). Funding, is it key for standing out in the university rankings? *Scientometrics, 121*(2), 771–792.

Berne, O. (2020). What does the Shanghai Ranking really measure? In *Working paper*. https://hal.science/hal-02918290/document

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment, 2008*(10), P10008.

Bougnol, M. L., & Dulá, J. H. (2006). Validating DEA as a ranking tool: An application of DEA to assess performance in higher education. *Annals of Operations Research, 145*, 339–365.

Bruni, R., Catalano, G., Daraio, C., Gregori, M., & Moed, H. F. (2020). Studying the heterogeneity of European higher education institutions. *Scientometrics, 125*(2), 1117–1144.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429–444.

Cooper, W. W., Ruiz, J. L., & Sirvent, I. (2011). Choices and uses of DEA weights. In W. Cooper, L. Seiford, & J. Zhu (Eds.), *Handbook on data envelopment analysis.* Springer: Boston, MA.

Daraio, C., Bonaccorsi, A., & Simar, L. (2015). Rankings and university performance: A conditional multidimensional approach. *European Journal of Operational Research, 244*(3), 918–930.

Daraio, C., Di Leo, S., & Leydesdorff, L. (2023). A heuristic approach based on Leiden rankings to identify outliers: Evidence from Italian universities in the European landscape. *Scientometrics, 128*(1), 483–510.

Daraio, C., Iazzolino, G., Laise, D., Coniglio, I. M., & Di Leo, S. (2022). Metachoices in ranking knowledge-based organizations. *Management Decision, 60*(4), 9951016.

Daraio, C., & Simar, L. (2007). *Advanced robust and nonparametric methods in efficiency analysis: Methodology and applications*. New York: Springer.

de Blas, C. S., Martin, J. S., & Gonzalez, D. G. (2018). Combined social networks and data envelopment analysis for ranking. *European Journal of Operational Research, 266*(3), 990–999.

Di Carlo, F., Modugno, G., Agasisti, T., & Catalano, G. (2019). Changing the accounting system to foster universities' financial sustainability: First evidence from Italy. *Sustainability, 11*(21), 6151.

Dugúe, N., & Perez, A. (2015). *Directed Louvain: Maximizing modularity in directed networks* [Doctoral dissertation, Université d'Orléans].

Etzkowitz, H., & Leydesdorff, L. (1997). Introduction to special issue on science policy dimensions of the triple helix of university-industry-government relations. *Science and Public Policy, 24*(1), 2–5.

Fauzi, M. A., Tan, C. N. L., Daud, M., & Awalludin, M. M. N. (2020). University rankings: A review of methodological flaws. *Issues in Educational Research, 30*, 79–96.

Fortunato, S., & Castellano, C. (2007). Community structure in graphs. arXiv: 0712.2716.

Hazelkorn, E. (2017). *Rankings and the reshaping of higher education: The battle for world-class excellence*. Palgrave Macmillan.

Hazelkorn, E., & Mihut, G. (Eds). (2021). *Research handbook on university rankings: Theory, methodology, influence and impact*. Edward Elgar.

Ho, W., Dey, P. K., & Higson, H. E. (2006). Multiple criteria decision-making techniques in higher education. *International Journal of Educational Management, 20*(5), 319–337.

Joro, T., & Viitala, E. J. (2004). Weight-restricted DEA in action: From expert opinions to mathematical models. *Journal of the Operational Research Society, 55*, 814–821.

Kang, H. J., Kim, C., & Choi, K. (2024). Combining bootstrap data envelopment analysis with social networks for rank discrimination and suitable potential benchmarks. *European Journal of Operational Research, 312*(1), 283–297.

Kim, J. (2018). The functions and dysfunctions of college rankings: An analysis of institutional expenditure. *Research in Higher Education, 59*(1), 54–87. https://doi.org/10.1007/s11162-017-9455-1

Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM), 46*(5), 604–632.

Kneip, A., Simar, L., & Wilson, P. W. (2016). Testing hypotheses in nonparametric models of production. *Journal of Business & Economic Statistics, 34*(3), 435–456.

Lavigne, C., De Jaeger, S.,& Rogge, N. (2019). Identifying the most relevant peers for benchmarking waste management performance: A conditional directional distance benefit-of-the-doubt approach. *Waste Management, 89*, 418–429.

Lepori, B., Geuna, A., & Mira, A. (2019). Scientific output scales with resources. A comparison of US and European universities. *PloS One, 14*(10), e0223415.

Lepori, B., Geuna, A., & Veglio, V. (2017). A typology of European research universities. Differentiation, layering, and resource distribution. *SWPS, 1*.

Liu, J. S., Lu, W. M., Yang, C., & Chuang, M. (2009). A network-based approach for increasing discrimination in data envelopment analysis. *Journal of the Operational Research Society, 60*(11), 1502–1510.

Liu, N. C., Cheng, Y., & Liu, X. (2016). Academic ranking of world universities. In *Encyclopedia of international higher education systems and institutions* (pp. 1–6). Springer.

Malliaros, F. D., & Vazirgiannis, M. (2013). Clustering and community detection in directed networks: A survey. *Physics Reports, 533*(4), 95–142.

Merisotis, J., & Sadlak, J. (2005). Higher education rankings: Evolution, acceptance, and dialogue. *Higher Education in Europe, 30*(2), 97–101.

Moed, H. F. (2017). A critical comparative analysis of five world university rankings. *Scientometrics, 110*(2), 967–990.

Nepomuceno, T. C. C., Seixas Costa, A. P. C., & Daraio, C. (2023). Theoretical and empirical advances in the assessment of productive efficiency since the introduction of DEA: A bibliometric analysis. *International Journal of Operational Research, 46*(4), 505–549.

Newman, M. E. (2016). Community detection in networks: Modularity optimization and maximum likelihood are equivalent. arXiv:1606.02319.

Olcay, G. A., & Bulu, M. (2017). Is measuring the knowledge creation of universities possible? A review of university rankings. *Technological Forecasting and Social Change, 123*, 153–160.

Pedraja-Chaparro, F., Salinas-Jimenez, J., & Smith, P. (1997). On the role of weight restrictions in data envelopment analysis. *Journal of Productivity Analysis, 8*, 215–230.

Podinovski, V. (2015). DEA models with production trade-offs and weight restrictions. In J. Zhu (Ed.), *Data envelopment analysis* . Springer: Boston, MA.

Podinovski, V. V. (2016). Optimal weights in DEA models with weight restrictions. *European Journal of Operational Research, 254*(3), 916–924.

Podinovski, V. V., & Athanassopoulos, A. D. (1998). Assessing the relative efficiency of decision making units using DEA models with weight restrictions. *Journal of the Operational Research Society, 49*(5), 500–508.

Sarrico, C. S., & Dyson, R. G. (2000). Using DEA for planning in UK universities: An institutional perspective. *Journal of the Operational Research Society, 51*(7), 789–800.

Seiford, L. M., & Zhu, J. (1998). On alternative optimal solutions in the estimation of returns to scale in DEA. *European Journal of Operational Research, 108*(1), 149–152.

Thanassoulis, E., Portela, M. C., & Allen, R. (2004). Incorporating value judgments in DEA. In: W. W. Cooper, L. M. Seiford,& J. Zhu (Eds.), *Handbook on data envelopment analysis*. Springer: Boston, MA.

Turri, M. (2016). The difficult transition of the Italian university system: Growth, underfunding and reforms. *Journal of Further and Higher Education, 40*(1), 83–106.

Valmorbida, S. M. I., & Ensslin, S. R. (2017). Performance evaluation of university rankings: Literature review and guidelines for future research. *International Journal of Business Innovation and Research, 14*(4), 479–501.

Wilson, P. W. (2008). FEAR: A software package for frontier efficiency analysis with R. *Socio-economic Planning Sciences, 42*(4), 247–254.