European Journal of Operational Research 244 (2015) 918–930

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Innovative Applications of O.R.

# Rankings and university performance: A conditional multidimensional approach ☆

Cinzia Daraio [a,*], Andrea Bonaccorsi [b], Léopold Simar [c]

[a] Department of Computer, Control and Management Engineering Antonio Ruberti (DIAG), University of Rome "La Sapienza", Via Ariosto 25, Rome 00185, Italy
[b] Dipartimento di Ingegneria dell'Energia, dei Sistemi, del Territorio e delle costruzioni (DESTEC), University of Pisa, Italy and National Agency for the Evaluation of Universities and Research Institutes (ANVUR)
[c] ISBA, Université Catholique de Louvain, Louvain-la-Neuve, Belgium and DIAG University of Rome "La Sapienza", Italy

ARTICLE INFO

*Article history:*
Received 18 June 2014
Accepted 2 February 2015
Available online 9 February 2015

*Keywords:*
Rankings
European universities
Conditional directional distances
Robust frontiers
Bootstrap

ABSTRACT

University rankings are the subject of a paradox: the more they are criticized by social scientists and experts on methodological grounds, the more they receive attention in policy making and the media. In this paper we attempt to give a contribution to the birth of a new generation of rankings, one that might improve on the current state of the art, by integrating new kind of information and using new ranking techniques. Our approach tries to overcome four main criticisms of university rankings, namely: monodimensionality; statistical robustness; dependence on university size and subject mix; lack of consideration of the input–output structure. We provide an illustration on European universities and conclude by pointing on the importance of investing in data integration and open data at European level both for research and for policy making.



## 1. Introduction and research questions

University rankings are the subject of a paradox: the more they are criticized by social scientists and experts on methodological grounds, the more they receive attention in policy making and the media. Rather than adding to the large literature on the methodological shortcomings of the existing rankings, this paper tries to give a contribution to the birth of a new generation of rankings, one that might improve on the current state of the art both in substantive and methodological bases. We provide two contributions: integrating new kind of information and using new ranking techniques.

The main criticisms (that we report in their historical order of introduction in the literature) addressed to university rankings, which we examine in detail in Section 2, can be summarized as follows:

(a) Monodimensionality
(b) Statistical robustness
(c) Dependence on university size and subject mix
(d) Lack of consideration of the input–output structure.

According to several authors, world rankings suffer from focusing only on the research dimension, which is more visible and easier to measure using external observations. A call for integrating the existing rankings with the educational perspective is in order. Yet several studies call into question the statistical properties of the rankings, irrespective of their substantive content, while others show that rankings systematically distort the representation in favour of large and established universities, and of universities in which scientific and technological disciplines, with particular reference to medical disciplines, are dominant. Finally, a few authors have raised the issue of whether it is acceptable to rank universities worldwide, without any consideration of the differences in resources made available to them by their respective national governments, or their input–output structure.

In this paper we provide an experiment that addresses all these issues, with reference to universities in Europe. The experiment might be replicated in USA and in several Asian countries, which have data comparable to the ones we use here.

First, we reduce monodimensionality by integrating data on research output (basically, scientific publications) with data on the teaching mission of universities. This is a major departure from existing rankings. The integration has been made possible by the creation of the Eumida (European Universities Micro Data) census of Higher

☆ Research supported by the University of Rome "La Sapienza", "Progetto di Ateneo 2013" and IAP Research Network P7/06 of the Belgian State (Belgian Science Policy). The opinions expressed in this paper are personal and do not involve the institution of the second author (ANVUR).

* Corresponding author. Tel.: +39 06 77274 068, +393488709425; fax: +39 06 77274 074.
*E-mail addresses:* daraio@dis.uniroma1.it (C. Daraio), a.bonaccorsi@gmail.com (A. Bonaccorsi), leopold.simar@uclouvain.be (L. Simar).

http://dx.doi.org/10.1016/j.ejor.2015.02.005
0377-2217/

Education Institutions (HEIs) in Europe, a project supported by the European Commission and Eurostat. In addition, we use data that refer to the quality of research. Thus by integrating data on education and research, and by including data not only on students but on degrees, we address the monodimensionality issue. In future studies other indicators (not available for this study) might be included, such as third mission, regional engagement and research infrastructures, leading to even more comprehensive analyses.

Second, we propose a ranking technique that is based on estimators that are robust to extreme values and outliers (as illustrated in Section 4) and delivers confidence intervals for the estimates (as illustrated in Appendix B), allowing the analyst to fully understand the statistical properties of the ranking score we propose.

Third, we address the dependence of rankings on size and subject mix by using a novel technique, called directional conditional efficiency analysis. As illustrated in the methodological section, this technique permits the estimation of efficiency measures net of the impact of size of universities (as proxied by the number of students) and net of the subject mix. This is another major departure from existing rankings. While our data do not allow any estimation in the fields of Social Sciences and Humanities (SSH), due to the limitations of current databases, for the first time we consider the subject mix of universities, as proxied by the specialization index of universities.

Fourth, the ranking we propose is based on an explicit input–output structure. We take benefit from the data in the Eumida dataset, that include academic and non-academic staff and personnel and non-personnel expenditures, to compute technical efficiency indicators in a multi-input multi-output framework. In this framework a university ranked high is one that makes the best possible use of its resources, on which it may have little discretionary power.

A consolidated literature has applied Data Envelopment Analysis (DEA) in the education sector (see e.g. Sarrico & Dyson, 2000; Sarrico, Teixeira, Rosa, & Cardoso, 2009 and Grosskopf, Hayes, & Taylor, 2014 and the references cited therein).

From a methodological point of view, this paper implements in the context of universities rankings the conditional directional distance approach by Daraio and Simar (2014) extending it to derive confidence bounds on the "managerial" efficiency scores robustly estimated. Indeed, as rightly emphasized by Grosskopf et al. (2014, p. 24): "Policy makers are interested in using efficiency scores [...] so it is crucially important to strengthen existing strategies for generating confidence bands around efficiency scores [...]".

Recently, Daraio, Bonaccorsi, and Simar (2015) propose a robust directional distance approach to analyze economies of scale and specialization in European universities and find that both size and specialization have a statistical significant effect on the efficiency. In this paper we make a step further and estimate the efficiency in the production of research quality taking into account also the volume of scientific production and the teaching realized. Research quality is hence the main output of interest. It is measured by a factor built taking into account international collaborations, normalized impact of research, high quality and excellence rate of publications. By applying a robust directional distance technique, we consider as non-discretionary outputs the volume of teaching and research carried out as well. We examine how European universities can improve their efficiency in the production of research quality, given the resources they are using and taking into account the level of teaching and research they produce while moving along a direction which is representative of the median case at European level.

Summing up, we believe that by integrating new data and adopting a novel technique there might be a leap forward in the way in which the activities and performances of universities are examined.

The paper unfolds as follows. Section 2 proposes an outline of the critical literature on university rankings. Section 3 introduces the main sources of data and lists the variables analyzed. Section 4 illustrates the methodology and is complemented by Appendix B. Finally, Section 5 presents the main results, while Section 6 concludes the paper.

## 2. University rankings: a guided tour of the critical literature

In this section we present the main lines of criticism to university rankings in the four chapters anticipated in Section 1. Other classifications are certainly possible. For the sake of clarity, criticisms classified in categories (a, monodimensionality) and (d, lack of input–output structure) deal with the substantive content of rankings, i.e., the data included (or missing), while studies under (b, statistical robustness) and (c, dependence on size and subject mix) mainly address methodological issues, i.e. how the data are processed in order to arrive at a ranking. Our classification clearly does not exhaust other lines of criticism: for example, we do not have any solution to the issue of English language bias, as well as for the lack of appropriate inclusion of Social Sciences and Humanities in rankings. Also we do not address the more general criticism according to which rankings are a disciplinary device created to impose neoliberal market-oriented values and practices onto an institution, the university, hitherto governed by the public ethos. At the same time our classification is reasonably comprehensive.

### 2.1. Monodimensionality

The argument is that universities all around the world perform several institutional missions: teaching, research, and third mission. Rankings that programmatically focus only on research outputs of universities are therefore biased. Even admitting that the third mission has been legitimized and institutionalized more recently, and is certainly less relevant (quantitatively) than the other two missions, it is felt that ignoring the teaching output altogether severely distorts the reality. Therefore, there is a demand for including information on teaching as well as research outputs of universities. Existing rankings include only a small set of indicators, whose meaning in terms of overall education activity of universities is questionable: the Alumni Nobel and Field prizes (10 percent) in ARWU (Academic Ranking of World Universities), student/staff ratios (20 percent weight), international students (5 percent) and international staff (5 percent) in QS (Quacquarelli Symonds) World University Rankings, and income per academic (2.25 percent), undergraduates admitted per academic (4.5 percent), ratio of international to domestic students (2.5 percent), ratio of international to domestic staff (2.5 percent) in THE (Times Higher Education Rankings). These proxies are considered unreliable and highly volatile by most analysts, as it is witnessed by the lack of consistency across various rankings, with the exception of the few top positions (Saisana, D'Hombres, & Saltelli, 2011; Salmi & Saroyan, 2007).

In fact, several authors have questioned the correspondence between rankings and quality of education, stating that in general "what is incorporated into the rankings is what is measurable, not what is valid" (Cremonini, Westerheijden, & Enders, 2008). The over reliance on research indicators may induce biased decisions (Bastedo & Bowman, 2010).

It is well known that the Shanghai ranking, the first global university ranking, originated from a specific need to provide information on research quality of universities which were considered target for Chinese students and decision makers (Liu, 2009). Therefore it did not incorporate any consideration of the teaching dimension, with the exception of prizes to Alumni, which is however biased toward large and old universities. Other rankings, such as Times Higher Education Supplement, introduced a few items related to education. However, the criticism hits the point: global league tables are largely based on the research output and ignore or underestimate the importance of education (Moed, Burger, Frankfort, & van Raan, 1985).