

STI 2018 Leiden

*23rd International Conference on Science and Technology Indicators
"Science, Technology and Innovation Indicators in Transition"*

STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Chair of the Conference

Paul Wouters

Scientific Editors

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

Layout

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

The articles of this collection can be accessed at <https://hdl.handle.net/1887/64521>

ISBN: 978-90-9031204-0

© of the text: the authors

© 2018 Centre for Science and Technology Studies (CWTS), Leiden University, The Netherlands



This ARTICLE is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

Tackling the heterogeneity of HEIs by combining different data sources and applying conditional benchmarking techniques¹

Cinzia Daraio*, Martina Gregori*, Giuseppe Catalano* and Henk F. Moed*

*cinzia.daraio@uniroma1.it; martina.gregori@uniroma1.it; giuseppe.catalano@uniroma1.it; henk.moed@uniroma1.it.

DIAG, Sapienza University of Rome, Via Ariosto 25, I-00185 Roma (Italy)

Introduction

The assessment of the performance of Higher Education Institutions (HEIs) at the micro (institutional), meso (regional) and macro (country) level is an important and recurrent question in the higher education's policy debate. The modernisation agenda for Higher Education in Europe (European Commission, 2016) identifies the relevance of creating effective governance and funding mechanisms for higher education. It underlines the importance to ensure greater flexibility and autonomy for institutions to specialise more easily, promoting better educational and research performance while fostering excellence within higher education systems. 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.

The measurement of academic performance is a core topic both in research policy and in the field of bibliometric and informetrics. International comparisons or rankings of institutions such as Shanghai, Times Higher Education and Leiden Ranking are being published on a regular basis with the aim to inform students, researchers and knowledge seeking external groups. Research managers use this information to benchmark their own institutions against their competitors (among many others, Hazelkorn, 2011).

However, the analysis of the performance of HE systems is far from being easy to deal with. One of the main critical issues to address properly the assessment of performance, in a multi-level (systemic) perspective, is the consideration of the heterogeneity of the HEIs involved. There are different sources of heterogeneity, including the mission, the national context, the presence or absence of medical schools, the legal status and the disciplinary orientation and degree of specialization. Among the heterogeneity factors of HEIs, disciplinary specialization or subject mix is considered one of the most relevant (López-Illescas et al. 2011, Daraio et al. 2011).

This paper addresses the issue of heterogeneity in its multidimensional faces by adopting two approaches. The first is integrating heterogeneous sources of available data, and the second the application of advanced econometric techniques, which allow comparing or benchmarking HEIs by capturing observed and *unobserved* heterogeneity.

¹ This work was supported by the Italian Ministry of Education and Research through the PRIN Project N. 2015RJARX7.

In a project funded by the European Commission, a team of five research groups created a European Tertiary Education Register (ETER), containing comparable basic information on about 2,500 higher education institutions from 34 European countries (ETER, n.d.). One of the aims of the current paper is to present a first exploration of the use of data from ETER for analysing the European higher education system. In the current study, the ETER database was combined with indicators from the Leiden Ranking 2017.

In the economics of production, as well as in operational research and management science, performance evaluation and benchmarking exercises based on quantitative methods are widely diffused. The idea of benchmarking is very common in productivity analysis because one of its key outcomes is the identification of targets and efficient peers to benchmark the Decision Making Units under evaluation. Benchmarking activity is also important to support strategic decision-making. Nonparametric efficiency estimators, such as Data Envelopment Analysis (DEA), are widely used in the context of performance evaluation and benchmarking because allow to identify and adopt best practices as a means to improve performance and increase productivity and are particularly valuable for service operations, where identifying benchmarks or standards is more difficult than in a manufacturing context.

The paper consists of two parts. The *first* part explores the use of the ETER dataset for assessing academic activity and performance. It gives an introduction to the multi-faces of the heterogeneity of HEIs, and illustrates how integrating heterogeneous datasets helps to tackle the heterogeneity of HEIs. It offers a view that is much broader than that of a purely bibliometric study or a university ranking, and that is based on data that are verified by representatives from the institutions themselves, or from the national statistical offices. It gives a descriptive characterization of the heterogeneity of HEIs, focusing on differences among national higher education systems, not only in terms of research but also related to teaching and training.

Part 2 presents an advanced, up-to-date nonparametric econometric technique to identify *unobserved* or latent heterogeneity factors and assess their impact on productivity/efficiency of HEIs. We refer to conditional nonparametric models (Daraio, Simar and Wilson, 2018a, b) in which some *pre-conditions* observed in the data allows one to identify latent or *unobserved* characteristics or performance dimensions difficult to observe directly.

About ETER

Table 1 give an overview of the main dimensions and types of indicators in ETER. The total number of HEIs included is about 2,500. It currently covers indicators per year for the time period 2011-2014. Almost 1,300 are categorized by the ETER Project Team as universities. It is this category that the current paper puts a focus on. ISCED stands for International Standard Classification of Education (ISCED), a statistical framework for organizing information on education maintained by UNESCO. The next section analyses bachelor (ISCED 6), master (ISCED7), and PhD (ISCED 8) students and graduates. It also uses an ISCED classification into 10 fields of education (FoE).

Table 1. Principal dimensions and indicators in ETER

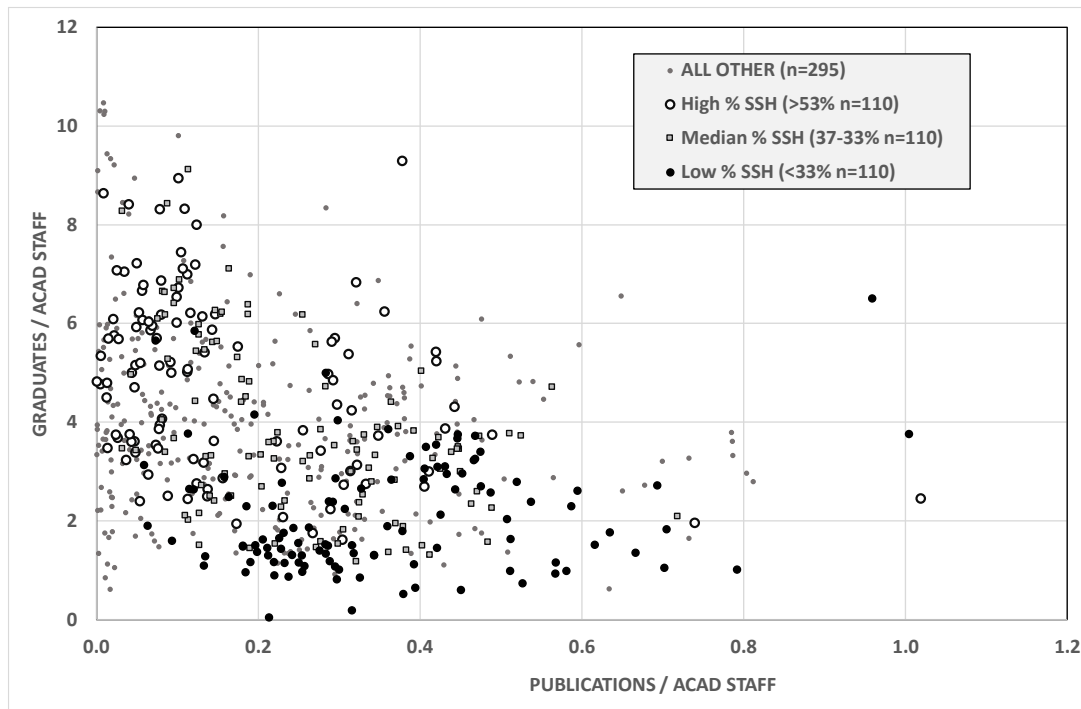
Dimensions	Variables
Identifiers	Institution name (in own language); English institution name
Basic institutional descriptors	Type of institution; foundation year; legal status; university hospital
Geographical information	NUTS 2, 3 region; city; geographical coordinates
Educational activities	Enrolled students and graduates ISCED 5, 6, 7, by field of education, gender, citizenship and mobility
Research activities	Number of PhD students and graduates (ISCED 8), by field of education, gender, citizenship and mobility; R&D expenditures
Expenditures	Personnel, non-personnel, capital expenditure
Revenues	Total revenues; core budget; third party funding; tuition fees
Staff	Number of academic staff (FTE and/or headcounts); by field of education, gender and citizenship; full professors by gender; administrative staff

Combining indicators on academic research and education

Figure 1 plots for all 627 universities for which data on publication and graduation output is available, and for the year 2014, publication productivity based on publication counts derived from Web of Science (WoS) against graduation productivity related to bachelors and master students. It also shows the degree to which the institutions are oriented towards social sciences and humanities (SSH), on the basis of their share of students active in the two UNESCO fields of education Arts and Humanities, and Social sciences, journalism and information. The category All OTHER in this graph relates to institutions for which no data on student numbers per field of education are available.

Figure 1 clearly illustrates the narrowness of scope of a productivity analysis based on one of the two aspects research or teaching only. But it also illustrates large differences in publication productivity between universities strongly oriented towards SSH and those with a weak level of activity in these fields. It provides evidence for the negative bias of WoS-based publication productivity counts towards institutions oriented towards social sciences and humanities.

Figure 1. Graduation and publication productivity versus disciplinary orientation



Part 1: Heterogeneity in the European HE system

The aim of this section is to illustrate differences between national higher education systems. All data are extracted from the ETER dataset and relate to the year 2014.

Figure 2. Distribution of bachelor students across fields of education for 14 countries.

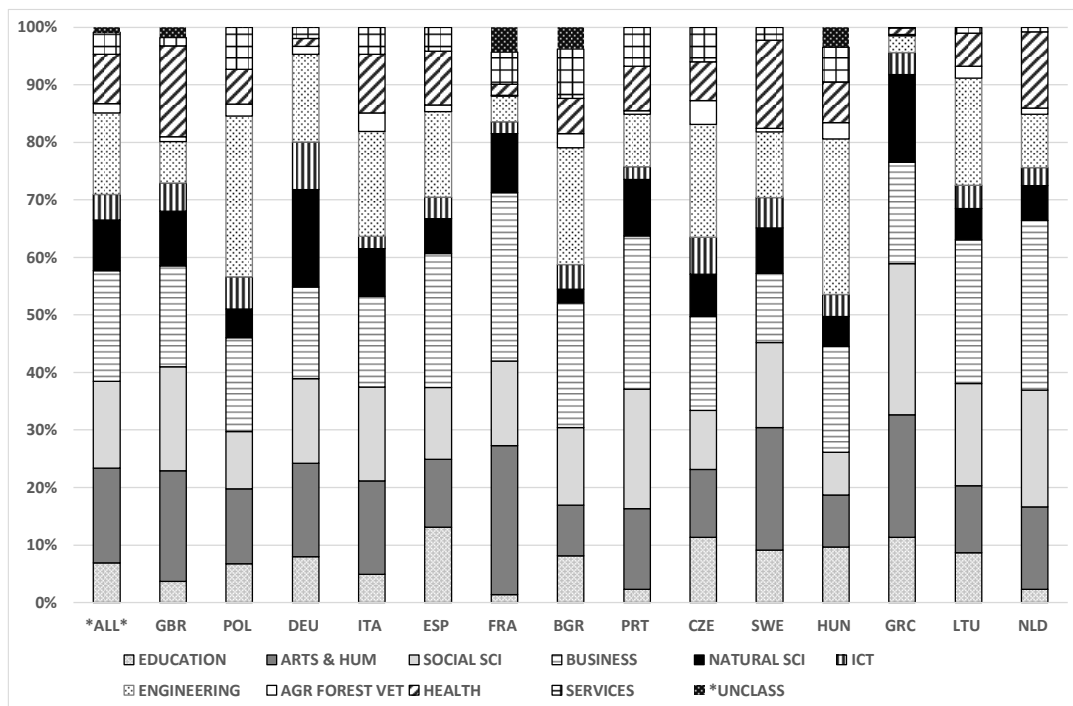
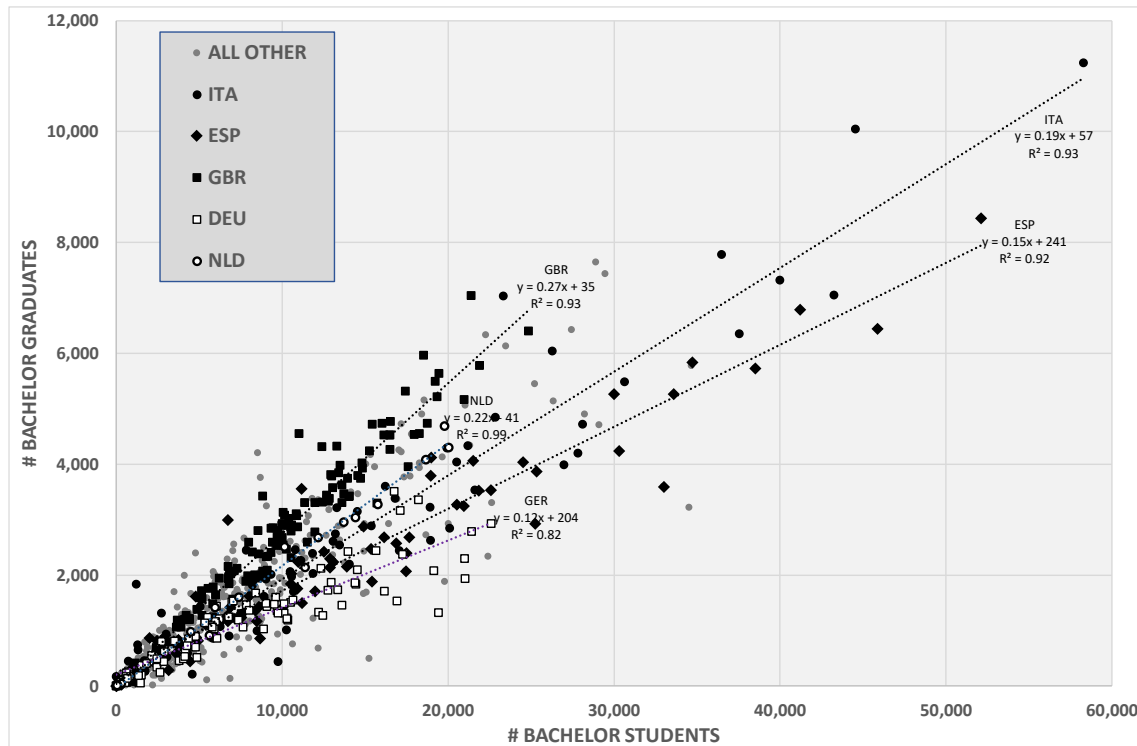


Figure 2 shows that in central- and east- European countries the share of students in arts & humanities, social sciences and business, administration & law is lower than that in other European countries, and the share of engineering students higher.

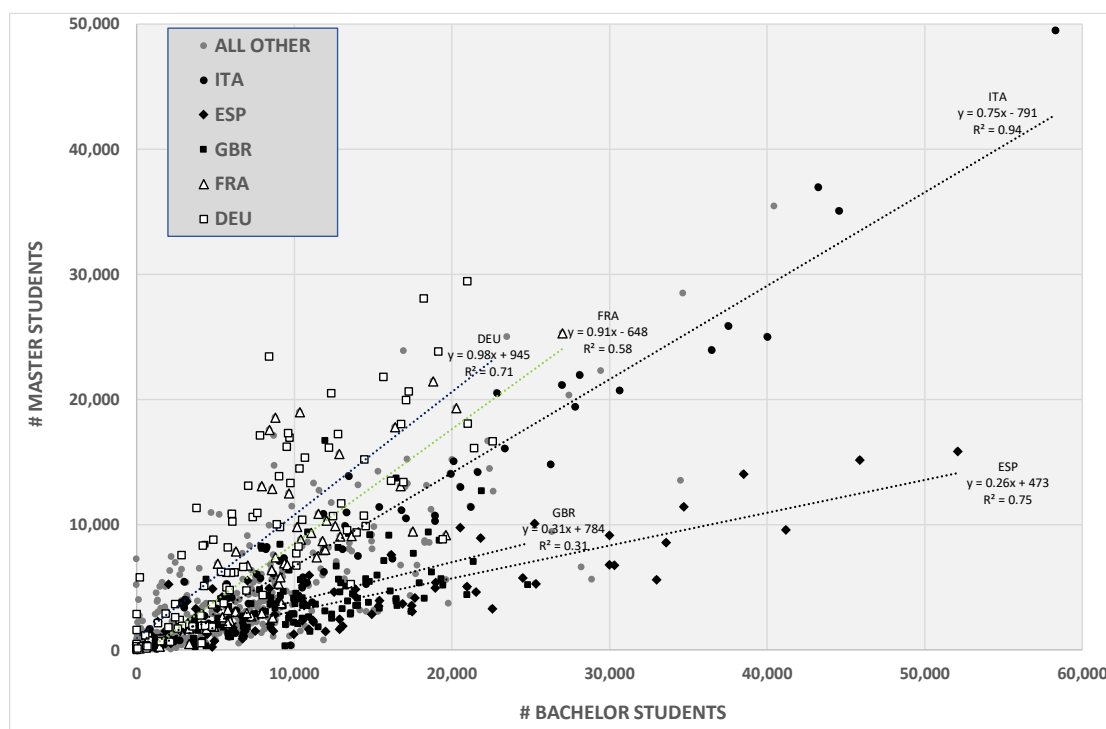
Figure 3. Number of graduates against number of enrolled students, for bachelor degree courses



The ratio between graduates and students of a country is visualized in Figure 3 as the slope of the linear regression line through the data points related to institutions in that country. Its interpretation is complex. This ratio can be assumed to be negatively affected by a factor that is considered as indicators of the inefficiency of the system, namely the time to degree. But other factors are at stake as well, such as the trend in the number of enrolled students over the years, the degree of experience in a country with the bachelor-master model, and the level of student fees. Difference in costs per students across subject domains are large and systematic (Filippini and Lepori, 2007). About existing differences in student fees policy, De Witte and López-Torres (2017) observe that some countries, such as UK, education is more and more expensive while in others, as in the Netherlands, the costs for education is lower. Student fees are also related to different student policy measures adopted. Booi, Leuven and Oosterbeek (2012) describe different measures to support students including a basic grant provided to all students, additional grant for students from low income families and student loan schemes with a kind of mortgage type repayment. However the authors show that often basic grants are insufficient and frequently the adoption of loans is quite low and often substituted by occasional works which contribute to increase the length of study.

Figure 4 plots the number of master students against the number of bachelors. The interpretation of differences in the ratio of masters and bachelor students is as of yet unclear to the current authors. Historical differences in the higher education system among countries may play an important role. Perhaps in Spain and, to a lesser extent, the UK, the attainment of a bachelors degree is more often considered as an end point in an academic study than it is in Germany or France.

Figure 4. Number of enrolled students in Master and Long Term degree courses versus Number of enrolled students in Bachelor degree courses



The previous observations illustrate that descriptive variables about historical context and current policies related to national HE systems (Agasisti and Bonomi, 2014) would add significant value to the performance assessments, as they would provide relevant background knowledge on national or local HE structures. As these descriptive analyses on the ETER database confirm, the variance among institutions within the same country is higher than the variance of institutions of similar size or specialization between different countries. The factors that influence the HEIs performance - such as, *policies*, *institution dimensions* or *disciplinary orientation* - could be used to carry out meaningful comparisons or benchmarks on *appropriate* reference sets, among “comparable” units of analysis, in order to deal with both the heterogeneity and the multi-level perspective in HEIs. This is the subject of Part 2 of the analysis.

Part 2: Tackling heterogeneity through pre-conditions of *unobserved* dimensions of performance

De Witte and López-Torres (2017) in concluding their review of efficiency analysis applications in education state that “for future research, it seems very fruitful to undertake more research about differences across countries and educational systems”. As far as methods for comparing HEIs is concerned, Daraio (2018) describes the development of *nonparametric* models as “dissatisfaction” towards traditional *parametric* models which rely on strong hypotheses about the functional form of the relations between inputs and outputs.

Our proposed approach (Daraio, Simar and Wilson, 2018 a,b) allows for comparing performance through flexible *conditional directional distances* which accounts for the heterogeneity in the comparison through conditioning the analysis at given external factors

(*observed* : z , *unobserved*: v). These factors may be observed or unobserved. In this paper we apply a robust nonparametric approach to identify unobserved heterogeneity factors in a general nonparametric and non-separable context and estimate the impact of this unobserved heterogeneity on the performance to analyse existing *trade-offs* or *complementarities* between the unobserved heterogeneity factors and the multidimensional performance of the HEIs. Table 2 illustrates the main differences between unconditional and conditional analysis.

Table 2. Basic ideas of conditional vs unconditional analysis

Main features	Unconditional Analysis	Conditional Analysis
<i>Performance measurement based on</i>	Inputs (\mathbf{x}) -that are the resources used- and Outputs (\mathbf{y}) -the realized “products or services” coming out from the transformation of the inputs in the production process.	Inputs (\mathbf{x}), outputs (\mathbf{y}) and <i>external</i> environmental <i>variables</i> which may be <i>observed</i> (\mathbf{z}) or <i>unobserved</i> (\mathbf{v}).
<i>Level -granularity and accuracy- of the comparison</i>	<i>Global</i> comparison: All units are compared against each other’s.	<i>Contextual</i> (or <i>local</i>) comparison: All units are compared against each other’s but taking into account, <i>i.e. conditioning</i> –from this the name comes from- to the \mathbf{z} and \mathbf{v} factors.
<i>Assessment of the impact of external factors on the performance</i>	Can be done by assuming the “separability condition” that is: the external factors \mathbf{z} do not affect the “efficient benchmarking frontier” but only the distance of the units from this (benchmarking) frontier	Can be done without assuming the separability condition. It allows to disentangle and testing the impact of \mathbf{z} and \mathbf{v} on the efficient boundary of the benchmarking frontier.
<i>Analysis of the trade-offs vs complementarity of efficiency vs heterogeneity factors \mathbf{z} and \mathbf{v}</i>	Not available	Available

Figure 5. A simple illustration of conditional analysis (that includes heterogeneity factors) vs unconditional analysis

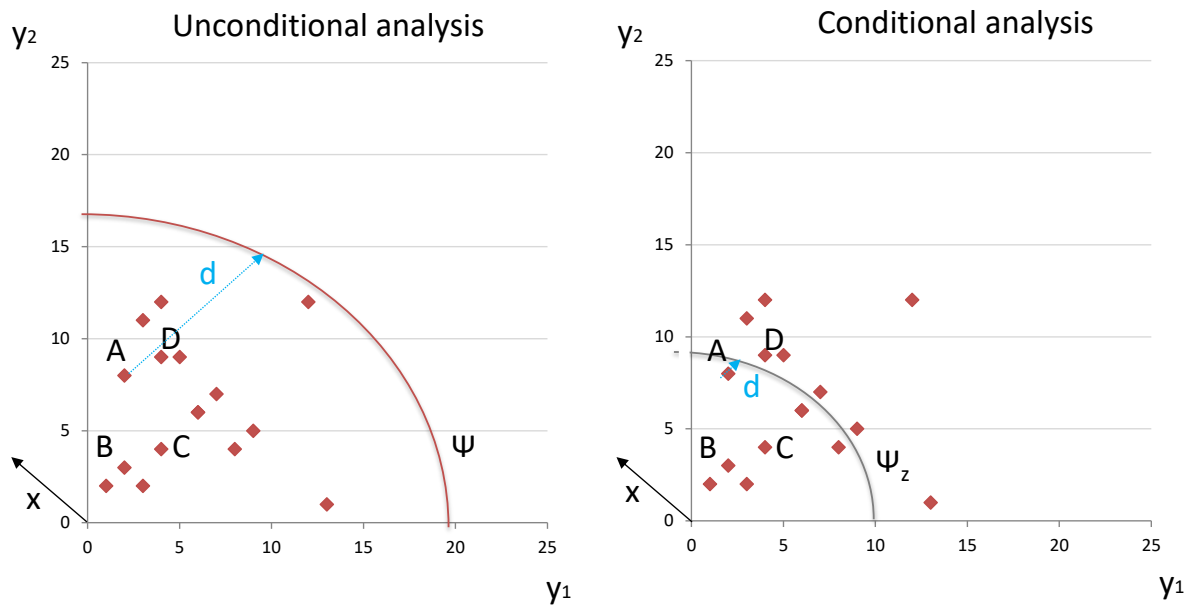


Figure 5 illustrates the assessment of the performance of HEI A with respect to other HEIs including B,C, and D and the other HEIs represented by a diamond along the direction indicated by d . HEIs' production process is characterized by two outputs (y_1 and y_2 in Figure 5) realized by using the input x .

In the panel on the left, a benchmarking comparison of HEI A is done with all the others in the comparison set including B,C and D. The best performing frontier Ψ represents those that are producing more outputs given their level of input. The right panel of Figure 5 reports the same HEIs of the left panel, but now the benchmarking is done considering the outputs produced given the input used *and also accounting for the external* (or environmental factor) or conditions z which may be the subject mix, the size or the localization of the units, or the different unobserved dimensions of performance. Hence, Ψ_z is the attainable set of $x,y|z$ that is considering the units that are comparable or similar with respect to the Z factor, taking in the comparison set of unit A those HEIs that have a value of Z "close" to the value of the Z of unit A. This conditioning allows us to compare units that are *comparable* according to the heterogeneity factor. Hence, the points B,C and those included in the production possibility set Ψ_z have a value of Z "close" to those of point A, while point D and the other points outside Ψ_z have a value much different than the Z value of unit A. The nonparametric approach applied in this paper estimates automatically, in a data-driven way (through the *bandwidth* selection), the units (HEIs) that are included in the conditional reference set.

This approach allows us to investigate the *pre-conditions* of unobserved "quality or excellence", or variety of governance regimes, on which a more careful and balanced comparison of HEIs may be done, by means of advanced benchmarking techniques, considered as a learning tool for the institutions themselves and as a useful policy analysis tool for policy makers.

Concluding remarks

The exploratory results presented in this paper illustrate the potential of the use of multiple data sources for the measurement of academic performance, especially the combination of publication databases revealing the research function, and the ETER database related to input indicators and teaching performance. The use of ETER potentially broadens – at least for European universities – the perspective of university performance compared to that of a traditional, publication- and purely research-based approach.

Much more work has to be done to further explore the ETER database, including a systematic analysis of its missing values and outliers as well as ways to properly deal with these. It could lead to recommendations as to how one could enhance the quality of its data, and how the data despite their incompleteness allow for valid conclusions.

The paper also show how an advanced econometric approach assessing the effect of observed and unobserved heterogeneity upon productivity/efficiency may substantially deepens rather elementary approaches calculating ratios of output and input indicators.

References

- Agasisti, T., Bonomi, F. (2014). Benchmarking universities' efficiency indicators in the presence of internal heterogeneity. *Studies in Higher Education*, 39(7), 1237-1255.
- Agasisti, T., Catalano, G. (2006). Governance models of university systems—towards quasi-markets? Tendencies and perspectives: A European comparison. *Journal of Higher Education Policy and Management*, 28(3), 245-262.
- Badin L., Daraio C., Simar L. (2012), How to Measure the Impact of Environmental Factors in a Nonparametric Production Model, *European Journal of Operational Research*, 223, 818–833.
- Booij, A. S., Leuven, E., Oosterbeek, H. (2012). The role of information in the take-up of student loans. *Economics of Education Review*, 31(1), 33-44.
- Capano, G., Howlett, M., Ramesh, M. (Eds.). (2015). *Varieties of governance: Dynamics, strategies, capacities*, Palgrave MacMillan, UK.
- Daraio C. et al. (2011), The European University landscape: A micro characterization based on evidence from the Aquameth project, *Research Policy*, 40, 148–164.
- 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., Simar L. (2007), *Advanced Robust and Nonparametric Methods in Efficiency Analysis. Methodology and Applications*, Springer, New York (USA).
- Daraio C., Simar L., Wilson P.W. (2018a), Central Limit Theorems for Conditional Efficiency Measures and Tests of the “Separability” Condition in Nonparametric, Two-Stage Models of Production, *The Econometrics Journal*, 21, 170-191.

Daraio C., Simar L., Wilson P.W. (2018b), Quality and its Impact on Productivity, Discussion Paper *to appear*.

De Witte, K., López-Torres, L. (2017). Efficiency in education: a review of literature and a way forward. *Journal of the Operational Research Society*, 68(4), 339-363.

ETER (n.d.). <https://www.eter-project.com/>

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

Filippini, M., Lepori, B. (2007). Cost structure, economies of capacity utilization and scope in Swiss higher education institutions, in *Universities and strategic knowledge creation: Specialization and performance in Europe*, Bonaccorsi and Daraio (eds), Edward Elgar, Cheltenham, 272-304.

Hazelkorn, E. (2011). Measuring world-class excellence and the global obsession with rankings. *Handbook on globalization and higher education*, 497-515.

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