doi.org/10.26398/IJAS.0031-015

MEASURING THE EFFICIENCY OF ACADEMIC DEPARTMENTS: AN EMPIRICAL STUDY

Lucia Buzzigoli[1], Antonio Giusti, Laura Grassini, Alessandro Viviani

Department of Statistics, Computer Science, Applications, "Giuseppe Parenti" University of Firenze, Firenze, Italy

Abstract.*The evaluation practice in Italian universities, introduced by the "Bologna process", has been renewed by a subsequent reform in 2010, that has redesigned the Italian academic governance. In this new context, the revisited role of the Departments, which have become the focus of academic activities, has motivated the objective of comparing their performance. The paper opens with a brief review of the literature on university performance evaluation with data envelopment analysis. After the presentation of the characteristics of the Italian university system and, in particular, of the University of Firenze, we syntetically describe the method and indicate the specification that will be applied, together with the set of input and output variables. The final section contains some concluding remarks and suggests directions for future research.*

Keywords: *Academic departments, relative efficiency, data envelopment analysis.*

1. INTRODUCTION

In recent years, the question on how the resources – especially the public ones - are utilized in higher learning institutions has become an important issue, because those institutions are under a general pressure to increase efficiency and improve the quality of their activities. However, it is not easy to analyze efficiency in such institutions: firstly, because only some input and output prices can be evaluated; secondly because these institutions utilize multiple inputs to produce multiple outputs (Johones, 2006) and therefore traditional parametric methods (e.g. production functions) cannnot be applied.

The nonparametric DEA (Data Envelopment Analysis) approach, which does not make explicit a specific production function, can consider a multiplicity of outputs and finds the production frontier empirically, has proven useful in assessing relative efficiency among universities or among departments of a single institution from the work of Ahn et al. (1988).

---

1 Corresponding author: Lucia Buzzigoli, email: lucia.buzzigoli@unifi.it.

These studies have always considered universities or departments as complex production units with specific characteristics. Nazarco and Saparauskas (2014) describe many examples of DEA application in academic performance measures from around the globe referring to multiple viewpoints that highlight the extent of the problems to be faced: the choice of the decision making units, the mode of fund allocation among units, the optimal size of the units, the efficiency of academic activities, the comparison among different countries, the dynamics of productivity.

As far as the Italian academic system is concerned, the efficiency of the university system is analyzed by Agasisti and Del Bianco (2006), Monaco (2012), Bergantino et al. (2012), Faggi et al. (2015); furthermore, Agasisti and Johnes (2009) and Agasisti and Perez (2010) compare Italian universities with academic institutions in different countries. Finally, Agasisti and Ricca (2016) analyzed technical efficiency of Italian public and private institutions in 2007-2010. The papers of Pesenti and Ukovich (1999), Rizzi (1999) and Buzzigoli et al. (2010) look at the departments of a single institution. Guccio et al. (2015) valued the convergence of technical efficiency of the universities in order to assess whether the reform process 2000-2010 results in an efficiency gain.

The purpose of this paper is to assess the relative performance of the departments of the University of Firenze.

A previous work (Buzzigoli et al. 2010) presented a study concerning the evaluation of the 70 departments of the University of Firenze using Data Evelopment Analysis. Several models with a different number and different sets of variables were implemented to assess departments' relative performance.

The results highlighted the need for further studies but, after the introduction of the law n. 240/2010, the set of decision making units radically changed because the number of departments has reduced to 24, and their functions changed as well. Moreover, also some of the variables used in the previuos application changed, like the one used to evaluate research activities.

Therefore, the novelty of the present work is represented substantially by two characteristics: the field of investigation and the characteristics of the information used.

The paper opens with the presentation of the characteristics of the Italian university system and – in particular – of the University of Firenze (in the following UNIFI). After a synthetic description of DEA, we present the input and output variables we chose for the application, focused on the efficiency in resource utilization. After the discussion of the analysis, the final section contains some concluding remarks and suggests directions for future research.

2. THE UNIVERSITY SYSTEM IN ITALY

The law n. 240/2010 introduced a reform of the Italian university system that implied the renewal of the organizational structure and of the composition of collegial bodies for universities (a detailed analysis in Lucianelli, 2013). Some issues are particularly relevant for our work.

First of all, departments have gained a new – and more significant – role: while the previous organization was based on departments and faculties, now departments are the unique internal structure of universities and assume direct responsibility not only for research, but also for teaching (previously assigned to faculties) and third mission activities. Therefore, university departments promote and manage research, manage teaching workloads among members, organize degree programs and doctoral programs and carry out consultancy work. While research and teaching are duties, third mission activities are not mandatory by law, but are gaining more and more attention because universities can give a decisive boost to innovation based growth by means of research valorization and the production of public goods (Baglieri, 2017).

Secondly, to reduce the number of departments and consequently simplify the organizational and management structure of universities, the law 240 has introduced a size threshold: at least 35/40 professors/researchers per department, depending on the total number of academic staff members. Each researcher belongs to one and only one department and the disciplinary areas of department members should be "homogenous".

Finally, old faculties have been replaced by optional 'connection structures' – often called schools – aimed exclusively for the rationalization of teaching by means of coordination between departments. If they exist, these structures manage common services to a number of degree courses.

To date, the Italian university system is composed by 97 higher learning institutions: 61 public universities, 19 private universities, 11 private online universities and 6 special tertiary education schools, which provide doctoral training. They are classified according to their size: there are 12 large universities with more than 40,000 students, 29 middle universities with a number of students between 15,000 and 40,000, and 56 small universities (including special schools) with less than 15,000 students (ANVUR, 2018).

According to the law 240/2010 the evaluation procedures among universities and inside universities (among departments and degree courses) is conducted by ANVUR (the Italian National Agency for the Evaluation of Universities and Research Institutes), established in 2011, which replaced both the CNVSU (National Committee for the Evaluation of Universities) and the CIVR (Italian Committee for Research Evaluation) (Turri, 2014).

The main funding system is centralized and it relies on the Ordinary Financing Fund (FFO), a single grant to finance teaching, research and other needs. In the last years the traditional, historically based allocation system (largely depending on the dimension of universities, in terms of personnel and students) has been integrated with teaching and research performance indicators (Geuna and Piolatto, 2016). The most important result related to the ANVUR evaluation is the construction of indicators on the quality of the research by scientific sector, by department and by university aimed at the allocation of a share of FFO.

Other important opportunities of research development derive from a number of national (like the Fund for Investments in Basic Research – FIRB – or the Research Projects of National Interest - PRIN) and European (like the Horizon 2020 framework scheme) sources of funding, on a competitive basis.

3. THE METHOD

The Data Envelpment Analysis (DEA)[2] was developed by Charnes et al. (1978) based on an idea by Dantzig (1951) and Farrell (1957) who introduced non-parametric methods for estimating efficiency. In general, given *n* production units, defined as DMUs (Decision Making Units), these transform multiple inputs into multiple outputs. The estimation takes place through the identification of a non-parametric linear boundary – for each unit separately – so as to identify the one that represents the best practice in the input/output transformation.

The DEA, as is known, is a non-parametric method since it does not require the specification of a model for the production process and does not require any predefined set of weights, relative to the inputs and outputs. The efficiency of each DMU is measured in relative terms and is determined by empirically comparing the input/output structure of a unit with that of all the others, finding the set of weights that maximizes its efficiency. To do this it is possible to follow two alternatives: to find the best theoretical DMU capable of producing the same quantity of outputs with a smaller amount of input (DEA input oriented) or that can produce a greater quantity of output with the same quantity of input (output oriented DEA). The relative technical efficiency is estimated through the DEA: the efficient units are the most efficient among those observed, without a comparison with average values or theoretical values.

---

[2] The literature on DEA is really vast; for the model description we basically refer to Coelli et al. (2005). Another reference for classic DEA models is Cooper et al. (2007).

From a mathematical point of view the DEA is a linear programming procedure that constructs a non-parametric (or border) surface on the data[3].

Supposing we have I DMUs with N inputs and M outputs, the input oriented approach for each unit can be represented as follows:

$$
\begin{aligned}
&\min_{\vartheta,\lambda} \vartheta \\
&s.t. Q\lambda \geq q \\
&\vartheta x \geq X\lambda \\
&\lambda \geq 0
\end{aligned}
$$

where ϑ is a scalar representing the efficiency of the DMU; Q is the $M \times I$ output matrix; X is the $N \times I$ input matrix; λ is an $I \times 1$ vector of constants; the column vectors x and q respectively contain the inputs and outputs of the DMU.

The ϑ that solves the problem is equal to 1 when the relative DMU is input efficient. If the resulting ϑ is less than 1, the $1-\vartheta$ difference represents the relative input reduction needed to project the inefficient DMU on the frontier. The vector λ contains the weights to be applied to efficient DMUs to construct the optimal theoretical producer which is the reference point for the considered DMU. The first inequality states that the theoretical DMU should produce at least as many outputs as that considered. The second inequality identifies how much the considered DMU input could be decreased, based on the efficiency value ϑ. The point ($X\lambda$, $Q\lambda$) represents the projection of the DMU on the frontier.

Instead, the output oriented approach for each unit can be represented as follows:

$$
\begin{aligned}
&\max_{\phi,\lambda} \phi \\
&s.t. Q\lambda \geq \phi q \\
&x \geq X\lambda \\
&\lambda \geq 0
\end{aligned}
$$

where $\phi - 1$ is the proportional increase in the outputs that could be reached by the DMU with the input quantities that remain constant; while $1/\phi$ is a scalar that defines the technical efficiency score, that varies between 0 and 1.

---

[3] We used DEAP, a free software developed by T. Coelli that can be downloaded at the web site www.uq.edu.au/economics/cepa/deap.htm.

In the input oriented DEA, input is minimized for a given level of output, while in the output oriented DEA, output is maximized for a given level of input. Sometimes the choice between the two models may not be evident. A general rule can be that for which the decision maker chooses the orientation based on the variables (inputs or outputs) on which he can exercise his control. If he can control both, then the choice will be determined by his objectives: for instance, whether he needs to cut costs he will choose input orientation, while if he wants to maximize production he will choose output orientation. Finally, the field of application also has its importance. For example, in the case of public services, a growing demand must often be met with strong budgetary constraints: therefore, output oriented DEA seems more adequate.

In both orientations, the units that have been ranked as efficient by the model are called 'peers', and they can be identified as possible benchmarks or targets (in the sense of best practice organizations), with which the relatively less efficient organizations are compared. For each inefficient DMU a list of efficient peers can be provided, which are the most similar to the inefficient DMU in terms of its production/input composition.

The aforementioned versions are associated with a technology with constant returns to scale (CRS). The DEA can also be applied with variable scale returns (VRS), adding a convexity constraint to the system: $\mathbf{1}'\lambda=1$, $\mathbf{1}'$ being the row unit vector. The VRS hypothesis is often more realistic, because the size of the DMU can affect its productivity and, therefore, also its efficiency. When the DMU size is heterogeneous, each DMU should be compared only with DMUs of similar size. Note that in the VRS case the estimated boundary is closer to the original points; therefore, the VRS efficiency scores are higher than the corresponding CRS scores, which may suffer from scale inefficiency.

Scale efficiency (S.E.) can be calculated as the ratio between technical efficiency under CRS and technical efficiency under VRS. As the first one is smaller than or equal to the latter, the measure of scale efficiency cannot be greater than 1. A DMU lying on both the VRS and the CRS frontier represents the optimal scale (i.e., scale efficiency=1); a DMU lying only on the VRS frontier is not technically efficient under CRS. A DMU lying below the two frontiers is both scale and technical inefficient. DEA procedure can also determine whether a DMU that does not reach the optimal scale should increase (decrease) its size: in this case, the DMU shows increasing (decreasing) returns to scale.

After the introduction of the DEA, several versions appeared in the literature (Adler et al., 2002), but for our application we refer to the classic ones: the CCR (assuming constant returns to scale, Charnes et al., 1978) and BCC (assuming variable returns to scale, Banker et al., 1984). The choice of the most appropriate specifications for our application (input/output oriented, CRS or VRS) will be made based on the specific purposes of the analysis and the characteristics of the data we are going to process.

4. THE UNIVERSITY OF FIRENZE: DATA AND VARIABLES

In 2017, the UNIFI was the seventh university in Italy in terms of number of students (51678 students) and teachers (1648 considering only professors, tenure and non-tenure researchers) (MIUR, 2017).

In the last ten years the FFO funds allocated to the UNIFI decreased of about 14% and the teaching personnel also decreased of about 20% (while students decreased of only about 5.3%).

In the following, we briefly describe the departments of the UNIFI, referring to the three dimensions of their activities (teaching, research and third mission).

After the 2010 reform, the number of departments passed from 70 to 24, while the previous classification of departments in five scientific areas was mantained: bio-medical (5 departments), scientific (6), technological (6), socio-economic (3), and liberal arts (4).

Table 1 presents departments per area and per number of teaching personnel in 2016. Our measures include researchers and professors, while temporary lecturer positions, usually hired to cover teaching needs, are excluded, as they cannot be considered internal personnel and imply additional costs.

The area classification hides within it diversified realities. Once for all: the areas are not homogeneous with respect to the dimension, when dimension is measured in terms of teachers affiliated to the department. Therefore, we grouped departments according the number of teachers in three classes: small (less than 50 teachers), medium (50-89 teachers), large (more then 89 teachers).

More than a half of the departments are of medium size: in fact, the general average number of teachers per department is about 68 (69 for the medium size).

There is no clear characterization of the different areas with respect to size, althogh in the scientific area prevail small departments and in the technological and, in the bio-medical area, the medium size prevails.

Tab. 1: Departments per area and dimension.

Area	Small (<50)	Medium (50-89)	Large (>89)	Total	# of teachers	Avg. # of teachers
Technological	1	4	1	6	376	63
Scientific	3	2	1	6	379	63
Liberal arts	2	2	0	4	207	52
Bio-medical	0	4	1	5	437	87
Socio-economic	1	1	1	3	225	75
Total	7	13	4	24	1624	68
Avg. # of teachers	42	69	108	68		

The academic staff can be considered the most important variable in our analysis, because teaching, research and third mission are carried out by department members.

Teaching activity can be measured through lesson hours. In the 50% of departments, the teaching burden per teacher is between 101 and 120 hours and the overall mean value is 94 hours per teacher. Some areas have a lower average: this can be due to the different educational needs in the various reaserch areas: some professional contents are typically transmitted by high qualified non-academic professionals, appointed by contract. This is not relevant for our research, where inputs and outputs must be referred to the departments resources and only to those. Therefore, the teaching hours considered are those held by academic staff.

Another proper measure for teaching activiy could be the number of students. Nonetheless, in the UNIFI, students are ‘linked’ to degree programs, which are managed by schools, although teaching is a departmental responsibility; moreover, the degree programs derive from the collaboration between several departments. Therefore, it would be difficult to distribute students among departments, if not on the basis of heavy conventions.

As far as the evaluation of research is concerned, different measures are available, both at national and local level. The process of research evaluation conducted by ANVUR is based on a complex procedure that implies the survey of a predetermined number of publications for each researcher (usually three) during a five-year period. In this sense, the final indexes do not measure the real productivity of research units, because they suffer from a sort of ex-ante standardization (they do not consider all the products of each researcher, but only a subset of them). In order to avoid this kind of problem we choose a measure, called *R*, that enters the calculation of the final index and that can be considered a sort of ‘grade’ of the quality of research evaluated in each scientific sector in UNIFI by comparison with researchers of the same scientific sector in Italy. *R* varies around 1: values above 1 indicate a quality above average, while values below 1 represent a quality below average. Department values range from 0.83 to 1.34 and the mean value is 1.06.

Another measure connected to research achievements is the amount of research grants. We consider a variable including reaserch funds coming from a number of external sources, all distributed on competitive basis, both at national and international level.

As regards the third mission, according to ANVUR (2015), several indicators can be used to evaluate the two main dimensions of the interaction between the academic world and society. The valorization of research can be quantified by the number of spinoffs, patents, the number and amount of contracts with industries, etc., while the involvement of universities in the social and cultural life by the number of archeological sites, museum facilities, social/cultural events organized by universities, etc.. In the first case, indicators refer primarily to technological research areas, while in the second case to humanities and socio-economic ones. The measure we chose is connected only to the first dimension and is the amount of the annual turnover. We calculated an average over three consecutive years (2015, 2016 and 2017) because data are very variable from year to year. At present, there are no reliable data for the second dimension.

Finally, another important variable to be considered is the amount of ordinary endowment funds, which are a general source of funding – although quite limited – for the departments' operative expenses. They are distributed yearly by the university's governing bodies on the basis of indicators related to the number of teaching personnel and the characteristics of departments' activities.

All the data used in the application have been kindly provided by the administrative offices of UNIFI.

Some descriptive statistics for the variables used in the analysis are presented in table 2, where ordinary endowment funds, grants and turnover are measured in Euros.

Tab. 2: Variables used in DEA application: average values and first and third quartile per department.

	Average	First quartile	Third quartile
Ordinary endowment funds (Euros)	243749.79	195253.00	294593.00
Teachers	67.65	48.00	84.00
Teaching hours	6342.58	4426.50	7853.00
R	1.05	0.96	1.13
Grants (Euros)	3029754.79	705790.57	3933894.19
Turnover (Euros)	433531.67	63283.30	567039.04

Other variables are available in university files, but we decided to use only the aforementioned ones, because they can be properly evaluated at department level and their definition is relevant for our analysis.

Finally, in higher learning institutions production modeling - especially when the public sector is concerned - also outcome measures should be used, where for outcome we mean the typical concept derived from public management: “something that happens in the world outside the organization […]: it is an effect ‘out there in the real world’ ” (Pollit and Dan, 2011). For example, the number of graduates who find a work one/two years after graduation. Although the recognition of the importance of learning outcomes is growing (see OECD, 2012), reliable measures of this kind cannot be proposed at department level.

5. MODEL SPECIFICATION

DEA specification implies the choice of the model orientation (input or output oriented), of the returns to scale assumption (CRS and VRS) and the choice of input and output variables.

As regards model orientation, in the literature on university or departments evaluation we can find both input and output oriented models (two examples for all: Aziz et al., 2013 and Agasisti and Ricca, 2016). We choose an output oriented model, because our aim is to verify – after the 2010 reform – which departments are more efficient than the others, given the resources. We are therefore interested in determining how much output quantities could be increased without modifying the input quantities as a mean to evaluate whether the use of financial and labour inputs – often limited – is efficient in the general framework of a rational use of public resources. This is particularly important in this historical period, when universities are required to improve the quality of their services while government funding is not increasing.

As regards the returns to scale assumption, the CRS hypothesis is adequate when the DMUs are operating at optimal scale (Coelli et al., 2005); in our case this assumption is not recommended, since the departments configuration was originally based on regulatory constraints and not on efficiency oriented organizational models. Therefore, we use a VRS specification, that also permits the computation of scale efficiency.

Other important issues are connected to the choice of variables. In fact, the choice of input and output variables in DEA specification is always essential: as in Coelli at al. (2005) we can say that “the quality and appropriateness of data used in these sophisticated techniques are just as important as the techniques themseves”. Nonetheless, this choice can be rather complicated in our context of analysis.

The first problem is the number of input and output variables related to the corresponding number of DMUs. While the number of units in our case is fixed, the number of variables is not determined a priori. After the 2010 reform, as mentioned earlier, the departments of the UNIFI decreased from seventy to twenty-four. Therefore, the number of possible variables to include in the model is limited by a much lower threshold than in Buzzigoli et al. (2010). It is well known, in fact, that the number of DMUs must be sufficiently high compared to the total number of input and output variables. Different rules of thumb have been proposed in the literature (for a review, Osman et al., 2014) but there is no general agreement on which is the best one. Our model specifications will respect the condition $I > \max\{N \times M, 3 \times (N+M)\}$ where I is the number of DMUs, N and M are – respectively – the number of inputs and outputs.

A second important issue is the different scale of variables. It is known that DEA can use variables in different units, but in our sample the magnitude of data range from numbers around 1 (the variable R) to numbers around 10^7 (the grants variable). This imbalance could produce calculation problems (Sarkis, 2007), and can be handled by dividing each input/output variable by its own mean. In our case, normalized data did not significantly modify the results produced by original data. Therefore, we can refer to original values, which produce input/output targets that are easier to interpret.

According to the previous considerations, the variables that can be used in our DEA application are: number of teaching personnel, volume of ordinary endowment funds (Euro), volume of research grants (Euro), number of teaching hours, research quality index and annual turnover (Euro).

Research grants can be interpreted both as an input variable (because the availability of financial resources is obviously of help in determining better and more numerous outputs) and as an output variable (as a result of research effectiveness). In our case, we considered them as an output for two reasons. Firstly, scientific reputation of researchers is often one of the criteria of selection of competitive projects; therefore, the amount of grants can be considered as an output of the academic 'production process'. Secondly, the choice is coherent with the classic DEA approach where smaller inputs and larger outputs are better.

Therefore, the output variables are: research grants, teaching hours, research quality index and annual turnover.

As a consequence, the input variables are: teaching personnel, which is the essential input to produce all the three typical activities of academic production process (teaching, research, third mission); ordinary endowment funds, that can be considered as a general support for all these activities. Moreover, these two variables can be considered proxies of labour and capital.

6. DEA APPLICATION AND RESULTS

The first step of our analysis is the application of a DEA model including all the aforementioned input and output variables (model M1).

Table 3 shows both CRS and VRS scores together with Scale Efficiency (S.E.) evaluation. Returns to scale (R.S.) are also reported, in order to assess whether the DMU operates with increasing (irs) or decreasing (drs) return of scale. In VRS model, 12 DMUs result efficient, while in CRS version five of the VRS efficient departments present relatively lower scores: it means that these DMUs are too big or too small relative to their "optimal" size. Nine DMUs have a VRS score less than 0.9 but, in any case, the average score is 0.942, indicating a good general level of performance.

Tab. 3: Results for model M1.

	Efficiency scores			
DMU	**CRS**	**VRS**	**S.E.**	**R.S.**
D1	0.999	1.000	0.999	drs
D2	0.922	0.942	0.979	irs
D3	0.865	0.938	0.922	drs
D4	0.747	0.815	0.917	drs
D5	1.000	1.000	1.000	-
D6	0.983	1.000	0.983	irs
D7	0.882	0.882	0.999	irs
D8	1.000	1.000	1.000	-
D9	1.000	1.000	1.000	-
D10	0.995	1.000	0.995	irs
D11	1.000	1.000	1.000	-
D12	0.674	0.709	0.950	drs
D13	0.697	1.000	0.697	drs
D14	0.707	0.852	0.830	drs
D15	0.748	0.944	0.792	drs
D16	0.717	0.833	0.860	drs
D17	1.000	1.000	1.000	-
D18	0.985	1.000	0.985	irs
D19	1.000	1.000	1.000	-
D20	0.942	0.983	0.959	drs
D21	0.972	0.981	0.991	irs
D22	0.809	0.811	0.997	drs
D23	1.000	1.000	1.000	-
D24	0.869	0.911	0.953	drs
mean	0.896	0.942	0.950	
n. of peers	7	12		

As the departments are involved in research, teaching and third mission, it is interesting to evaluate their performance for each activity separately. Therefore, a second step in our application involves the comparison of three additional DEA specifications. The respective models, called M3 (research), M4 (teaching) and M5 (third mission) share with M1 the same input variables (operating expenses and teachers), but include different outputs, as specified in table 4.

Tab. 4: Output variables of the different models.

Model	Output variables
M1	grants, R, teaching hours, turnover
M3	grants, R
M4	teaching hours
M5	turnover

Tab. 5: VRS efficiency scores for model M3, M4 and M5.

DMU	M3	M4	M5
D1	0.679	1.000	0.243
D2	0.920	0.900	0.141
D3	0.858	0.858	0.167
D4	0.761	0.742	0.011
D5	1.000	0.936	0.209
D6	1.000	1.000	1.000
D7	0.746	0.859	0.361
D8	0.786	0.953	1.000
D9	1.000	0.836	0.026
D10	1.000	0.952	0.002
D11	0.687	1.000	0.008
D12	0.709	0.420	0.265
D13	0.791	0.857	0.720
D14	0.776	0.665	0.374
D15	0.903	0.708	0.166
D16	0.828	0.608	0.331
D17	1.000	0.950	1.000
D18	1.000	0.963	0.159
D19	0.955	1.000	0.011
D20	0.828	0.953	0.051
D21	0.853	0.959	0.075
D22	0.703	0.786	0.251
D23	1.000	1.000	0.075
D24	0.858	0.861	0.052
mean	0.860	0.865	0.279
n. of peers	7	5	3

The VRS efficiency scores for each of the three models are in table 5, together with the number of peers (7, 5 and 3 for, respectively, M3, M4 and M5). The Spearman Coefficients between the different rankings – including also model M1 – are in table 6.

Tab. 6: Spearman coefficients between the rankings.

	M1	M3	M4	M5
M1	1.00			
M3	0.51	1.00		
M4	0.74	0.27	1.00	
M5	-0.09	-0.15	-0.23	1.00

The results show that only department D6 (a technological one) is a peer in all models.

In model M3, 16 of the 17 non-efficient departments present decreasing returns of scale, while in model M4, 14 of the 19 inefficient departments shows increasing returns to scale. Although models M3 and M4 show a similar VRS average score (0.860 for M3 and 0.865 for M4), the average scale efficiency (not reported) is rather different (0.794 for M3 and 0.956 for M4), and the Spearman coefficient between M3 and M4 (0.27) shows that these two rankings are not stable. As far as the efficiency scores of M3 and M4 are concerned, the scatterplot in fig.1 – including the bisector of the quadrant – highlights an 'arrow shape' confirming the findings in Buzzigoli et el. (2010). Only two DMUs are efficient in both models, while, more often, efficient units in M3 model are not efficient in M4 model and vice-versa. Therefore, we can hypothesize that in the departments of UNIFI there is a trade-off between research and teaching efficiency. One department (belonging to the bio-medical area) is off the cloud of the other departments, because it shows a particularly low score for M4. This is probably due to the characteristics of teaching in the biomedical departments, most of them sharing low scores in M4.

Model M5 produces only three peers and shows very low efficiency scores when compared with the other two models (the average VRS score is 0.279).

The comparison of M5 with M4 (which has a single output variable, like model M5) is in Fig. 2: all DMUs, except the peers, have higher efficiency values in M4 (therefore, they are under the bisector) and most of the scores are lower than 0.4.

These results confirm the problems connected to the nature of third mission and the peculiarity of the measure used to evaluate it. The measure adopted quantifies only one dimension of the third mission, that is the amount of contracts of departments with both private and public non-academic subjects. Therefore, it shows noticeable differences among departments because, at present, only some of them have a strong vocation of this kind.

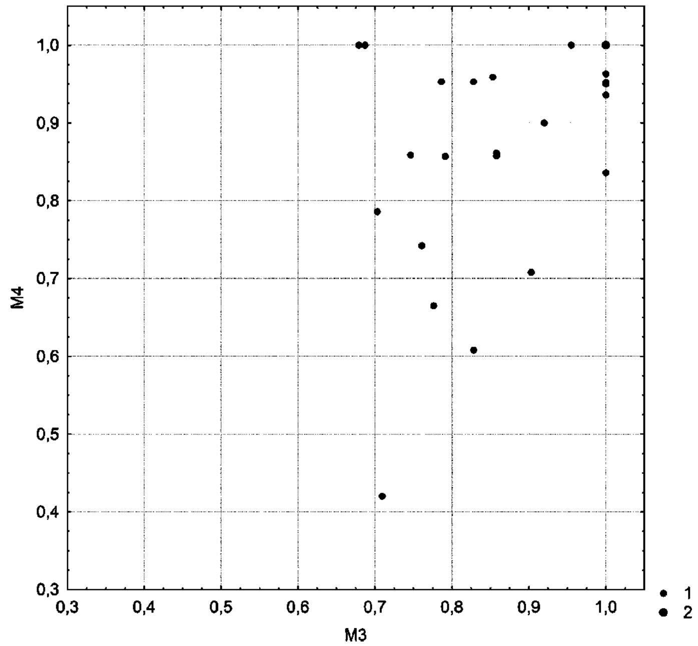

Fig. 1: Efficiency scores: M3 vs M4 model.

In order to evaluate whether turnover has a decisive influence on the general model, we applied a sort of sensitivity analysis defining a fifth model, called M2, which is obtained from model M1 eliminating the turnover from the outputs. The results are in table 7.

The comparison between Table 3 and Table 7 shows that the number of peers in M2 decreases (from 12 to 10) and all the M2 peers are also M1 peers. The average score ranges from 0.942 to 0.933; the rankings show a high Spearman coefficient (0.94). The scatterplot in Figure 3 points out that M2 scores are always not greater than M1 scores.

As noted, in the M5 model the third mission is important only for a limited number of DMUs; the comparison between models M1 and M2 shows a situation of moderate relevance of this activity in the DMUs together with the problem related to the adopted measure.

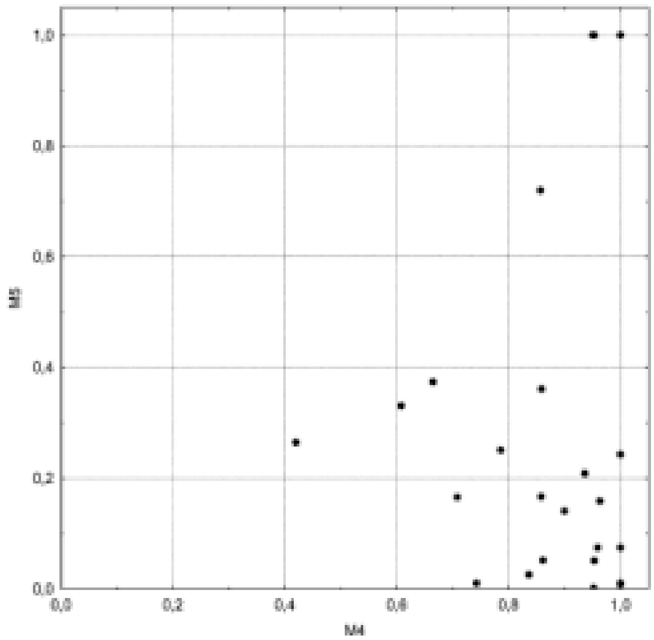

Fig. 2: Efficiency scores: M5 vs M4 model.

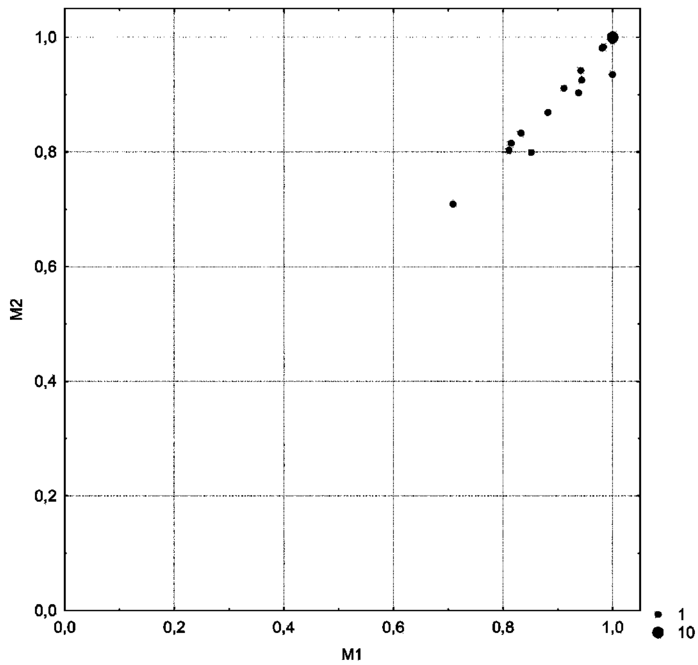

Fig. 3: Efficiency scores: M1 vs M2 model.

Tab. 7: Results for model M2.

DMU	Efficiency scores CRSTE	VRSTE	S.E.	R.S.
D1	0.979	1.000	0.979	drs
D2	0.922	0.942	0.979	irs
D3	0.851	0.903	0.942	drs
D4	0.747	0.815	0.917	drs
D5	1.000	1.000	1.000	-
D6	0.964	1.000	0.964	irs
D7	0.863	0.869	0.993	irs
D8	0.988	0.994	0.995	irs
D9	1.000	1.000	1.000	-
D10	0.995	1.000	0.995	irs
D11	1.000	1.000	1.000	-
D12	0.645	0.709	0.91	drs
D13	0.628	0.935	0.672	drs
D14	0.67	0.799	0.838	drs
D15	0.738	0.925	0.797	drs
D16	0.683	0.833	0.82	drs
D17	1.000	1.000	1.000	-
D18	0.985	1.000	0.985	irs
D19	1.000	1.000	1.000	-
D20	0.938	0.983	0.955	drs
D21	0.972	0.981	0.991	irs
D22	0.802	0.803	0.998	irs
D23	1.000	1.000	1.000	-
D24	0.864	0.911	0.949	drs
mean	0.885	0.933	0.945	

Finally, we can compare the peers of the different models according to some characteristics: typology (as specified below), scientific area and dimension.

First of all, we consider the number of DMUs for which each peer is of reference. If this number is high (in our case greater than three) we can consider this peer as ‘properly efficient’; if the DMU is efficient but do not represent a peer for any inefficient DMU, then its efficiency should be viewed with caution. In this case, we labelled the peer as ‘weak’. Furthermore, in each model we identified the peers with scale efficiency and, among them, the properly efficiency peers.

Table 8 reports the results of the classification for each model. Only in model M4 all the peers are properly efficient, showing a sort of specialization in teaching. In model M2, which has a large number of peers (10), most of ‘properly efficient’ DMUs are scale efficient.

Tab. 8: Number and typology of peers per model.

Model	# of peers	'Weak' peers	'Properly efficient' peers	Peers with scale efficiency	'Properly efficient' peers with scale efficiency
M1	12	1	5	7	4
M2	10	0	6	5	5
M3	7	3	2	3	2
M4	5	0	5	2	2
M5	3	0	2	1	1

A supplementary analysis considers the characteristics of our particular DMUs, that belong to different scientific areas and have different dimensions.

Table 9 presents the distribution of peers per area. There is no relevant evidence, except that there is only one efficient department in the biomedical area. This result is probably due to the particular role of biomedical departments in the Italian health system, that – as we have already underlined in §4 – makes their activities partially incomparable whith respect to the 'traditional' academic ones of the other departments. This aspect should be further investigated, but the number of biomedical units is too small to make a specific analysis.

On the contrary, the classification of peers according to their dimension (small/medium/large) shows some remarkable evidence. For each model, table 10 classifies peers per dimension and table 11 displays the overall average performance of DMUs and the average performance for small, medium and large DMUs. In all models, small DMUs outperform the others, due to the high percentage of peers among small departments. While large DMUs are more efficient that the average in four models, medium DMUs always show an efficiency lower than the average.

Tab. 9: Peer distribution per model and area.

	Scientific	Biomedical	Technological	Humanities	Social science	Total
# of depts.	6	5	6	4	3	24
Model						
M1	3	1	4	3	1	12
M2	3	0	3	3	1	10
M3	2	0	2	3	0	7
M4	2	0	2	0	1	5
M5	1	0	2	0	0	3

Tab. 10: Peer distribution per model and department dimension.

	Small (<50)	Medium (50-89)	Large (>89)	Total
# of depts.	7	13	4	24
Model				
M1	5	5	2	12
M2	5	4	1	10
M3	5	2	0	7
M4	2	2	1	5
M5	2	1	0	3

Tab. 11: Mean efficiency measures per model and department dimension.

Model	Small (<50)	Medium (50-89)	Large (>89)	All
M1	0.989	0.904	0.980	0.942
M2	0.989	0.897	0.955	0.933
M3	0.967	0.824	0.789	0.860
M4	0.960	0.798	0.917	0.865
M5	0.350	0.236	0.295	0.279

A synthetic graphical representation of the peers in the various models is the Venn diagram in Fig. 4, where capital letters represent the peers and indicate their dimension. Each set is formed by the peers of the corresponding model. Set M2 is not shown because it is the union of set M3 and set M4.

The various characteristics could be analysed together. For instance, in the more ‘specialistic’ model M4, where all the peers are ‘properly efficient’, peers are not characterized by area (two are from technological area, two from the scientific one, one from social science), nor by dimension (two are small departments, two are medium and one is large). On the contrary, in model M3 we see a prevalence of small peers, but three of them are‘‘weak’ and the two properly efficient peers are also scale efficient and belong to the scientific area.

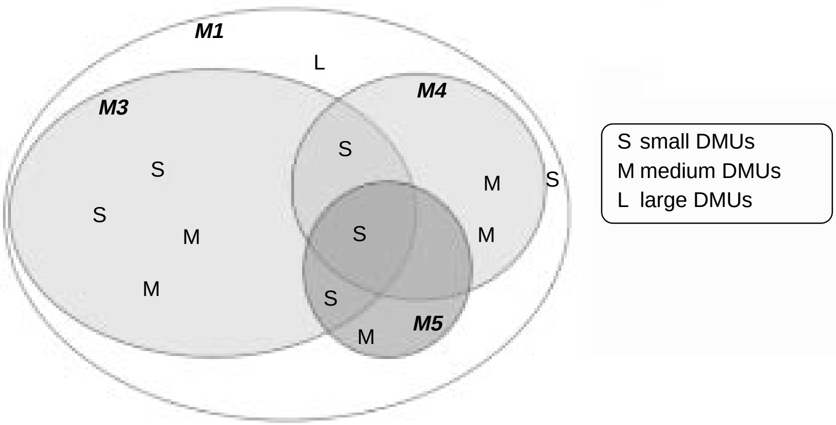

Fig. 4: Venn diagram representation of peers.

7. CONCLUDING REMARKS

DEA has been used frequently for comparing higher learning institutions. In the literature on academic performance DMUs are often universities; the analysis of academic efficiency based on departments is less common, but, on the contrary, is of great interest, because all dimensions of academic activity can be referred to the department level. Nonetheless, the evaluation among departments has the problem of identifying input and output measures related to a "production process" with a "technology" that is difficult to define, due to the complexity of the activities carried out. The definition of the variables and the identification of adequate measures represent an essential step for assessing the quality and relevance of the analysis, as data are often derived from administrative sources. This imposes an adequate and consistent path of definition and measurement of the variables employed, together with the need for a sensitivity analysis to evaluate the robustness of results.

In this paper, we presented an analysis of the departments of the UNIFI after the 2010 reform, in order to evaluate their relative efficiency with regard to some characteristics: academic activity, scientific area and dimension.

We proposed several models with the same input variables, but different outputs.

Firstly, we evaluated the efficiency of departments with respect to a general production process including all available outputs (model M1) and secondly we proposed a set of models where the output is the specific target of a single academic activity (models M3, M4 and M5 for, respectively, research, teaching, third mission). The results for the M5 model suggested a new specification of the general model (model M2), where the output variable related to the third mission was omitted.

In general, we found a good average value for VRS efficiency scores in the more comprehensive models (M1 and M2), while in the specific models the average is smaller (partly due to the smallest number of output variables) with a particular low value for M5.

More specifically, the analysis shows remarkable differences in the values of input and output variables across and within the scientific areas. Probably, an alternative classification should be proposed in order to distinguish and enhance the pecularities of departments and, in particular, in confirming some specific vocations (for instance, third mission for the technological area).

The classification according DMUs' dimension seems to produce more informative results: in particular, small departments outperforms the others with respect to efficiency measures.

Finally, although DEA technique is a powerful tool and have been used in many issues for this purpose, it should be noted that its use is linked to the nature and the characteristics of utilised data. Therefore, the use of relevant and exhaustive data on academic activities will be an essential condition for obtaining more detailed analysis in the future. In this way DEA could also be an important tool to support the management of resources within universities.

ACKNOWLEDGEMENT

The authors wish to thank the anonymous referees for their valuable comments and suggestions.

REFERENCES

Adler, N., Friedman, L. and Sinuany-Stern, Z. (2002). Review of ranking methods in the data envelopment analysis context. *European Journal of Operational Research*, 140: 249–265.

Agasisti, T. and Dal Bianco, A. (2006). Data envelopment analysis of the Italian university system: theoretical issues and policy implications. *Int. J. Bus. Perform. Manag.* 8(4): 344-367.

Agasisti, T. and Johnes, G. (2009). Beyond frontiers: Comparing the efficiency of higher education decision-making units across more then one country. *Educ. Econ.* 17(1): 59-79.

Agasisti, T. and Perez-Esparrels, C. (2010). Comparing efficiency in a cross-country perspective: The case of Italian and Spanish state universities. *High. Educ.* 59(1): 85-103.

Agasisti, T. and Ricca, L. (2016). Comparing the efficiency of Italian public and private universities (2007-2011): an empirical analysis. *Italian Economic Journal*. 2: 57-89.

Ahn T., Charnes, A. and Cooper. W.W. (1988), Some statistical and DEA evaluations of relative efficiencies of public and private institutions of higher education, *Socio-Economic Planning Sciences*. 6: 259-269.

ANVUR (2015). *La valutazione della terza missione nelle università italiane. Manuale per la valutazione*. Available on ANVUR website.

ANVUR (2018). *Self-Assessment Report*. September 2018.

Aziz, N.A.A., Janor, R.M. and Mahahadi, R. (2013). Comparative departmental efficiency analysis within a university: a DEA approach. *Procedia – Social and Behavioural Sciences*. 90: 540-548.

Baglieri, D. (2017). Italian universities and the third mission. *Workshop "Academic rankings and civic universities: Italian higher education institutions at the challenge of third mission"*. Padua, November 2nd, 2017.

Banker, R.D., Charnes, A. and Cooper, W.W. (1984). Models for estimation of technical and scale inefficiencies in data envelopment analysis. *Management Science*. 30: 1078–1092.

Bergantino, A.S., Capozza, C. and Porcelli, F. (2012).'Measuring the efficiency of the Italian university system: the role of market structure. A Twostep DEA Analysis at Faculty level. *XXV Conferenza della società italiana di economia pubblica*.

Buzzigoli, L., Giusti, A. and Viviani, A. (2010). The evaluation of university departments. A case study for Firenze. *International Advances in Economic Research*. 16: 24-38.

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

Coelli, T.J., Prasada Rao, D.S., O'Donnell, C.J. and Battese, G.E. (2005). *An introduction to efficiency and productivity analysis*. Springer, New York.

Cooper, W.W., Seiford, L.M. and Tone, K. (2007). *Data envelopment analysis, a comprehensive text with models, applications, references and DEA-solver software*. 2nd Edition. Springer, New York.

Dantzig, G.B. (1951). Maximization of a linear function of variables subject to linear inequalities. In T. Koopman, editor, *Activity analysis of production and allocation*. John Wiley and Sons, Inc, New York.

Faggi, F., Gregori, G.M. and Madoni S. (2015). Indagine sull'efficienza delle università pubbliche italiane. *Economia e Politica*. 2015.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society. A*. CXX, Part, 3, 253–290.

Geuna, A. and Piolatto, M. (2016). Research assessment in the UK and Italy: costly and difficult, but probably worth it (at least for a while). *Research Policy*. 45: 260–271.

Guccio, C., Martorana, M.F. and Monaco, L. (2015). Efficiency and productivity trends of Italian universities: bootstrapped DEA approach. *Journal of Productivity Analysis*. 45(3): 275-298.

Johones, J. (2006). Measuring technical efficiency in higher education: an application of DEA to economic graduates from UK universities. *European Journal of Operational Reserch*. 174(1): 443-456.

Lucianelli, G. (2013). University governance at the crossroads: the Italian case. *International Journal of Business Research and Development*. 2(2): 20 34.

MIUR (2017). *Portale dei dati dell'istruzione superiore, ufficio statistica e studi, ministero dell'istruzione*. Università e Ricerca.

Monaco, L. (2012). Measuring Italian university efficiency: a non-parametric approach. *MPRA Paper*. n. 37949.

Nazarco, J. and Saparauskas, J. (2014). Application of DEA method in efficiency evaluation of public higher education institutions. *Technological and Economic Development of Economy*. 20(1): 25-44.

OECD (2012). *Assessment of higher education learning outcomes. AHELO. Feasibility study report*.

Osman, I.H., Anouz A.L. and Emrouznejad, A. (2014). *Handbook of research on strategic performance management and measurement using data envelopment analysis*. IGI Global.

Pesenti, R. and Ukovich W. (1999). Evaluating academic activities using DEA, *Internal Report*, DEEI-Università di Trieste.

Pollit, C. and Dan, S. (2011). The impacts of the new public management. In *Europe: a Meta-Analysis, Cocops Work Package 1 – Deliverable 1.1*. European Commission, European Research Area.

Rizzi, D. (1999). L'efficienza dei dipartimenti dell'Università Ca' Foscari di Venezia via DEA e DFA. *Economia pubblica*. 31(3): 103136.

Sarkis, J. (2007). Preparing your data for DEA. In J. Zhou and W.D. Cook (2007), editors, *Modeling data irregularities and structural complexities in data envelopment analysis*. Springer.

Turri, M. (2014). The new Italian agency for the evaluation of the university system (ANVUR): a need for governance or legitimacy? *Quality in Higher Education*. 20(1): 64-82.