Cristian Barra (University of Salerno)
Sergio Destefanis (University of Salerno)
Vania Sena (University of Essex)
Roberto Zotti (University of Torino)

PERFORMANCE OF FIRST-YEAR UNIVERSITY STUDENTS IN ITALY AND ACADEMIC REFORMS

This version 10/06/2018

Abstract

This paper uses Data Envelopment Analysis (DEA) to model and quantify the impact of university resources on the academic performance of 49,394 first-year students in a large public Italian university. The analysis is articulated in two stages. The first stage decomposes students' academic performance into a student-level faculty effect directly influenced by the way teaching activities are organised at the faculty level, and a student-level within-faculty efficiency score, which is based on the comparison of the students' performance with their peers in the same faculty. In the second stage, we focus on the student-level faculty effects and on their evolution across the different faculties. We find evidence that the increase in the number of contact hours for professorial staff has an adverse effect on the student-level faculty efficiency scores. In addition, the student-level within-faculty efficiency scores appear to be mostly unaffected by the reforms. Finally, these results are robust vis-à-vis different specifications of the production set.

Keywords: Data Envelopment Analysis; Decomposition of efficiency; Higher education, Education production functions.

JEL-Codes: I21, I23; C14; C67.

1. Introduction

According to Hanushek (2014, p. 1) "A simple production model lies behind much of the analysis in the economics of education. The common inputs are things like school resources, teacher quality, and family attributes, and the outcome is student achievement." Yet (ibid., p. 6) " … The existing research suggests inefficiency in the provision of schooling. It does not indicate that schools do not matter. Nor does it indicate that money and resources never impact achievement. The accumulated research surrounding estimation of education production functions simply says there currently is no clear, systematic relationship between resources and student outcomes". In this paper we bring to bear upon these issues a novel specification of the relationship between student achievement, university resources and other factors, using student-level data from a large public Italian university. We believe that this institution is a worthwhile case study for the following reasons. When in 2007, Italian higher education institutions were invited to reduce the number of offered programmes or increase the average number of contact hours for a substantial share of the academic staff, this university did not reduce the number of programmes but only increased the number of contact hours for professorial staff, allowing assessment of the impact that higher workload has on staff effectiveness. This can yield interesting evidence at a moment when governments in Western countries aim to increase students' participation rate and equity in higher education and at the same time wish to improve quality. Employers need graduates whose knowledge and skills are appropriate for the tasks they are hired to perform. Society wants higher education to increase the wealth of the nation and to solve its social problems. This increase in demands is not matched by growth in the size of the academic corps. Consequently, the quality of teaching, research, and service may be increasingly threatened and needs serious examination in relation to the increasing academic workload. There is more pressure upon academics to do more teaching (i.e. lectures, assessments tasks, assignments, students advising and counselling, supervising postgraduate students), research (i.e. conducting research, writing books, chapters, articles, reports) and community service (administration and faculty commitments, consultancy and professional activities as well as services to the community); as a consequence, an increase in the total amount of hours that academic staff is requested to allocate on their professional activities will affect the quality of academic outputs as well as the proportion of time they will spend on any specific activity. The issue of increased workload is controversial. Indeed, while university administrators tend to associate higher workload with higher productivity, academics, believing they are already overworked, associate higher workload with stress and lower productivity (see Soliman and Soliman, 1997 for a discussion on the relationship between academic workload and quality.

Moreover, increasing academic workload is also a result of efficiency measures achieved by non-replacement of retirees, voluntary separations, resignations, and non-renewal of temporary contracts. The Italian experience in this respect is quite relevant. In 2007 a ministerial decree (Ministerial Decree 16 March 2007, turned into law on 31 October 2007) was approved, linking the number of modules and programmes offered by a faculty to the number of hours taught by its permanent academic staff. This law aimed at rationalising the programmes so that they could be staffed by permanent staff (rather than by temporary staff). The reaction of academic institutions to these reforms was mixed. Most universities simply decided to increase the number of contact hours (where possible) for existing staff while trying to maintain the existing academic provision (hoping that this could help with future recruitment). In 2008, the Italian government approved another law (Law n. 133/2008), which targeted directly the turnover rate of academic staff so that each university could only replace a fraction (varying, over time, between 20% and 50%) of the retired staff. the latter measure decisively aimed at reducing the academic staff employed by Italian universities, subsequently decreasing the public resources devoted to higher education. In this respect, it has been successful. Figures from the Italian Department of Education show that the total number of academic staff[1] across the whole higher education system shrank from 62,768 to 56,449 over the period 2008–2011. At the same time, student numbers decreased from 1,809,499 to 1,781,786, with the result that the student-to-staff ratio increased from 28.82 to 31.56.

At face value, the combined impact of these two reforms of universities is rather clear: overall, universities were first induced to increase the workload of existing staff, but once turnover was frozen after 2008, the workload of staff increased even more (while the number of students increased simultaneously). Our paper quantifies the impact of these legislative measures on the academic performance of nine cohorts of first-year students (more precisely, students at the end of their first year) drawn from a large Italian university based in southern Italy and observed from academic year 2002–2003 to academic year 2010-2011. While some aggregate data exists about various measures of students'

[1] The sum of full professors, associate professors, and lecturers.

performance and satisfaction, detailed cross-university data on students registered at universities are not made available by the Italian Department of Education. The use of student-level data is needed to control for various students' characteristics and to disentangle the portion of the students' academic achievement affected by the students' personal characteristics and effort from the portion ascribable to the institution (university or faculty) activities. Focusing on several cohorts of students from the same university offers some further advantages. First, by focusing on data from one university only, we eliminate unobservable sources of cross-university heterogeneity. Second, data collected at the university level ensure consistency in the collection of information on students. This enables a sounder comparison of the first-year students' performance before and after the 2007 reform.

We employ an empirical methodology articulated into two stages. In the first stage, we adapt the procedure suggested by Portela and Thanassoulis (2001), Thanassoulis and Portela (2002), Johnes (2006a; 2006b) to a single-university setting (see Barra and Zotti, 2014; Barra et al., 2017). We use Data Envelopment Analysis (DEA) to decompose the academic performance of the students at the end of their first year into a within-faculty efficiency score (which is directly linked to their effort and capabilities) and a faculty-dependent effect (which measures the impact on students' academic performance of the efficiency of faculties compared to other faculties within the same academic area). In the second stage, we regress these student-level efficiency scores (focusing in particular on the faculty-specific effects) on faculty-level indicators such as the student-staff ratio, while controlling for some individual characteristics of the students (i.e. basic demographics, educational background and pre-enrolment characteristics, households' financial conditions). The results suggest that the student-level performance that can be attributed to the efficiency of the faculties (with respect to the other faculties within their scientific area) has on average worsened after the two reforms. The results also suggest that faculty efficiency has never reverted back to the pre-2007 values, suggesting that until 2011, the faculties have not taken any corrective action to address the combined negative impact of the two reforms on faculty performance. These results are robust in relation to a set of additional checks.

The paper has the following structure. Section 2 summarises the main literature on the measurement of efficiency in higher education and provides a short overview of the Italian university system. Section 3 describes the methodology we employ in the paper as well as the data and the production sets for our DEA model. The main results are given in Section 4. Finally, Section 5 offers concluding remarks, including some policy implications.

2. Literature review and institutional setup

2.1 Literature review

Frontier analysis is a commonly used methodology for the analysis of efficiency among universities. Its popularity is due to the fact that it is sufficiently flexible to take into account the multi-output and multi-input dimension of universities' activities (Johnes, 2006a). Generally, researchers have used frontier analysis to compare the efficiency of different universities (Abbott and Doucouliagos, 2003; Flegg et al., 2004; Agasisti and Johnes, 2010) or of different departments (Johnes and Johnes, 1995; Thursby, 2000; Tomkins and Green, 1988; Sarrico and Dyson, 2000; Kao and Hung, 2008; Tyagi et al., 2009; Halkos et al., 2012; Moreno and Tadepalli, 2002).

Given the type of analysis carried out in our paper, in this brief review we focus on the literature that measures the efficiency of departments within the same university (Halkos et al., 2012; Buzzigoli et al., 2010; Tauer et al., 2007; Kao and Hung, 2008; Tyagi et al., 2009; Moreno and Tadepalli, 2002; Koksal and Nalcaci, 2006; Rizzi, 1999; Gimenez and Martinez, 2006). Most of these papers are grounded in different institutional setups, and so it is difficult to draw some general lessons from their results. However, they also share some common concerns. First of all, some studies find that inefficiency is driven mostly by the composition of the workforce, and therefore some departments prefer to allocate their resources either to research or to teaching only. Tyagi et al. (2009) examine the efficiency at Roorkee Institute of Technology (India) and find that some of the inefficiency must be ascribed to the fact that some departments did not use their staff fully. Gimenez and Martinez (2006) focus on the cost efficiency of Autonoma University of Barcelona (Spain), and find that a higher proportion of non-tenured staff is associated with higher efficiency levels. Halkos et al. (2012) also find evidence of misallocation of resources among departments at the University of Thessaly (Greece). As for Italy, to the best of our knowledge, there are very few papers that have measured the efficiency of

Italian departments and faculties within the same institution. Rizzi (1999) presents an analysis of the efficiency at the University of Venice, obtaining similar results by applying both Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA). Buzzigoli et al. (2010) evaluate the efficiency at the University of Firenze, with the main result being that the allocation of academic staff across departments is skewed towards teaching. A second concern arising from the literature is that the disciplinary composition (or subject mix) of different universities may affect efficiency measurement. For instance, younger universities may prefer a subject mix that favours less expensive subjects (like social sciences or humanities), while universities located in former industrial areas and with historical links to the local industry may prefer a subject mix skewed towards science and engineering. In this case, comparing the performance of these two types of universities may be misleading, as the second group could appear as less efficient than the first group simply because its subject mix favours more resource-using subjects.[2] A few attempts have been made to control for the subject mix of different institutions when measuring their efficiency (or productivity). Agasisti and Bonomi (2014) suggest comparing similar departments within different universities on the grounds that departments have sufficient autonomy in managing their production process. Johnes (2006a) analyses the academic performance of a cohort of graduates from British universities using frontier analysis and suggests sorting students into subject areas that include disciplines with similar unit costs.

Discussions of the role of subject mix in influencing the measurement of efficiency in universities are linked to discussions of the correct unit of analysis within this area. A common feature of the studies mentioned above is that departments or faculties have seldom been considered to be decision-making units (DMUs) on the grounds that, ultimately, universities are organisations that have control over their production technology. Only a few studies have suggested that departments may be a more relevant DMU for the measurement of efficiency, arguing that in some institutional settings, departments or faculties have sufficient autonomy to decide on their levels of staff, equipment, and teaching programmes (that such may well be the Italian case is vouched for by Bonaccorsi and Daraio, 2007, p. 246; Buzzigoli et al., 2010, p. 26).

It must be stressed that only few studies (Johnes, 2006a; Johnes, 2006b; Barra and Zotti, 2014; Barra et al. 2017) have applied DEA models to the academic performance of

[2] Sarrico and Dyson (2004) and Sarrico et al. (2009) stress the importance of the subject mix when measuring productivity and efficiency at university level. See also Dyson et al. (2001) for a discussion on the homogeneity assumptions about the units under assessment in DEA.

university students to measure the efficiency of the departments/faculties to which they belong, even though this procedure had already been applied in the context of schools (Portela and Thanassoulis, 2001; Thanassoulis and Portela, 2002; Waldo, 2007). Johnes (2006a) analyses the performance of English universities in 1993, controlling for their subject mix, while Johnes (2006b) focuses on the teaching efficiency of British economic departments in 1993. She finds that efficiency scores measured through student-level data (unlike those relying on departmental or university data) allow the researcher to disentangle the students' effort from the characteristics of the institution to which they belong. Along similar lines, Barra and Zotti (2014) use DEA to decompose students' under-attainment into one component attributable to their faculties and another attributable to their own effort. Barra et al. (2017) provided an example of how data routinely produced and stored by universities can be used to monitor the performance of students and identify the component of the students' under-performance that can be attributed to the universities' activities and therefore can be addressed with corrective measure. In this paper, we extend their approach by focusing on a longer period (before and after the 2007 reform) by analysing faculties' performance and the role of potential determinants as well as exploring its evolution and determination related to the to the 2007 policy reform.

2.2 The Italian Institutional Setup

During the 1990s, a set of reforms allowed Italian universities (heavily regulated by the central government up until then) to have their own statutes, to create new departments and teaching programmes, and to acquire more discretionary power over the use of the budget devolved to them from the central government (Law 168/1989). At the same time, higher education funding from the central government was made conditional on a set of performance indicators in an attempt to promote an efficient allocation of resources (Law 537/1993).

Following the Bologna process, and throughout our period of analysis, Italian universities adopted the credit system (commonly used across Europe). The credit system was introduced to address one of the main problems of the Italian university system, namely, the existence of a large number of students who could not complete the degree within the notional time of completion. In line with this policy aim, and importantly for our empirical setup, among the performance indicators used by the central government to promote an efficient allocation of resources there exist two groups of students: those enrolled in the second year, having obtained a certain number of credits in the first year, and those who have not obtained any credits at the end of the first year (the current regulation is set by Ministerial Decree 18 October 2007, n. 506; CNVSU, DOC 07/2009, but similar rules were introduced earlier). Both indicators penalise universities where students do not acquire a number of credits as close as possible to those theoretically obtainable in the first year (i.e. 60 credits), preparing the ground for a speedy completion of the academic career.

In 2007 a legislative action (Ministerial Decree 16 March 2007, which became Law 544/2007) was enacted to reduce the transfer of resources from the central government to universities. This new law included rules linking the number of modules offered by a faculty to the number of hours taught by its permanent academic staff. These rules were meant as a way to stop the proliferation of new modules, as faculty, in order to command more resources for recruitment, used to increase the number of modules offered while relying on contract lecturers for their delivery. However, faculties, instead of reducing the number of modules, chose to keep it as high as possible with the result that the workload of academic staff was increased.[3] These changes were enacted during academic year 2007–2008 and became mandatory for all in academic year 2009–2010. The University under our scrutiny was one of the early enforcers. The decrees had already been adopted in 2007–2008, which crucially adds to the interest of this particular case study.

In 2008 the growing pressure on Italian public finances urged another government to promulgate a law (Law n. 133/2008) including a set of measures whose aim was to slow down the turnover of academic staff across the whole university system, in practice reducing the number of academic staff in each university. Only a fraction of the retiring academic staff—initially 20%, then, after a 2009 modification of the law, 50%—could be replaced by each university. Although the legislation did not officially interfere with the autonomy a university has in deciding how to fill a retired post, in practice it created a big push toward the reduction of personnel, as well as incentives for the replacement of more expensive positions (in particular, full professors) with lower-pay posts.

These measures were relatively successful in reducing the financial resources devoted nationally to higher education. The reduction was equal to 39.5 million euros for

[3] This basically meant increasing the workload of full professors and associate professors, who must each teach 120 hours a year, but can easily go beyond this limit. Lecturers teach up to 60 hours but often teach less than that, and their workload is not easily adjustable.

2009, 119 million euros for 2010, 198 million euros for 2011, 276 million euros for 2012 and 314 million euros from 2013 onwards. Legislative action was also successful in reducing the total number of academic staff in the university system, and in affecting its mix. Table 1, below, shows the evolution of the total number of academic staff as well as its composition (full professors, associate professors, lecturers), of the number of students enrolled, and the student-academic staff ratio over the period 2003–2011, comparing the whole Italian higher education system and the University under scrutiny. For the whole of Italy, the number of academic staff has shrunk since 2008, and the fall in the number of full professors has been bigger than the fall in the number of associate professors and (especially) lecturers. This does not change when accompanied by the information on student numbers. The ratio of students to academics has increased over the same time span (and the student-professor ratio increases even more quickly). Very similar patterns occur at the University under analysis, where academic staff numbered 1,034 in 2008, going down to 939 in 2011. Here, too, full professors is the group having experienced the largest fall, and, although the number of students enrolled has slightly decreased, the ratio of students to academics has risen from 37.72 to 38.93 over the period 2008–2011 (and from 65.99 to 66.72, when only full and associate professors are considered). We do not have comparable figures for financial resources at the national and university level. Yet the evolution in the number of academic staff, and its relationship to the number of students, show that evidence from this University about the impact of the 2007 measures is likely to be highly relevant for the whole Italian university system.

Table 1. Academic staff, Students and Student/Staff ratio – 2003-2011 – Italy and University

Year	Full professors	Associate professors	Lecturers	Academic staff	Students	Student/ Ac. Staff ratio[a]	Student/ Prof. ratio[b]
Italy							
2003	17,957	18,097	20,426	56,480	1,768,295	31.31	49.05
2004	18,070	18,103	21,229	57,402	1,814,048	31.60	50.15
2005	19,274	18,967	22,010	60,251	1,820,221	30.21	47.60
2006	19,843	19,086	23,045	61,974	1,823,748	29.43	46.85
2007	19,623	18,735	23,571	61,929	1,810,101	29.23	47.19
2008	18,929	18,256	25,583	62,768	1,809,499	28.83	48.66
2009	17,880	17,567	25,435	60,882	1,814,344	29.80	51.19
2010	15,854	16,955	24,939	57,748	1,799,542	31.16	54.85
2011	15,242	16,611	24,596	56,449	1,781,786	31.57	55.94
University (under scrutiny)							
2003	252	251	263	766	38,883	50.76	77.30
2004	263	260	277	800	38,580	48.23	73.77
2005	290	293	328	907	37,327	41.15	64.03
2006	295	314	376	985	39,165	39.76	64.31
2007	297	310	388	995	39,029	39.23	64.30
2008	290	301	443	1,034	39,003	37.72	66.00
2009	278	293	437	1,008	39,665	39.35	69.47
2010	256	285	427	968	37,065	38.29	68.51
2011	246	302	391	939	36,558	38.93	66.72

Note: Data have been sourced from the Ministero dell'Università e della Ricerca - Ufficio di Statistica. Banca dati dei docenti di ruolo; Ministero dell'Università e della Ricerca - Ufficio di Statistica. Indagine sull'Istruzione Universitaria; Nucleo di valutazione Università degli Studi – Relazioni 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011. a. The Student/Ac. staff ratio is equal to the sum of students divided by the sum of full professors, associate professors and lecturers. b. The Student/Professor ratio is equal to the sum of students divided by the sum of full professors and associate professors.

3. **The Empirical Analysis**

In this section, we adopt Data Envelopment Analysis (DEA) to measure the academic performance of each student and decompose it into a within-faculty efficiency score (related to his or her effort in the studies) and a faculty-dependent efficiency score (which measures the faculty's capability of deploying efficiently its teaching resources). In the second stage, we apply fractional regression models to assess the links between the faculty-dependent effect and some faculty-level indicators related to the effectiveness of teaching, also controlling for some characteristics of the individual students (i.e. basic demographics, educational background and pre-enrolment characteristics, households' financial conditions). We also use fractional regression analysis to test for the existence of structural breaks in faculty efficiency corresponding to the introduction of the 2007 law. The details of this two-stage methodology are provided in Section 3.1. In Section 3.2, we illustrate the data we use for the empirical analysis, while our production set is described in Section 3.3.

3.1. The Empirical Methodology

As mentioned in the Introduction, the academic performance of university students can be seen as the result of two components (once the quality or innate ability of the students are controlled for): the effort students put into their studies and the efficiency with which departments (or faculties) organise their teaching resources. This decomposition method builds upon Thanassoulis and Portela (2002), who have proposed an empirical methodology that allows decomposition of pupils' performance in British schools exactly in these two components. The main intuition behind this procedure is that any cross-sectional variation of the students' performance may be due either to differences in the efficiency between the faculties or to differences in the amount of effort students put into their studies. In practice, the components measure the performance of the pupils with respect to two groups: the whole population of pupils across different schools (stage 1) and the pupils registered in the same school (stage 2). The ratio between these two indicators of performance is a measure of the pupils' under-attainment that is attributable to the way the school manages its teaching resources. The procedure can be easily adapted to the higher education context, although some changes need to be implemented in order to control for variations in the subject mix across different faculties. Indeed, although departments or faculties within the same university share the same environment, teaching across subjects differs because of the use of labs, the reliance on classes and the patterns of contact hours. This implies that faculties should be grouped in areas that share a similar organisation of teaching (Johnes, 2006a; 2006b and Sarrico et al., 2009). Hence the first stage of the decomposition procedure should be modified in such a way that the performance of students taught by a given faculty member can be compared to the performance of the students in the same subject area (and not to the whole population of students). Following the guidelines of the Italian Department of Education, we group faculties into three main areas: Pure and Applied Sciences (including Fac. n. 1, Fac. n. 2 and Fac. n. 3), Humanities (including Fac. n. 4, Fac. n. 5 and Fac. n. 6), and Social Sciences (including Fac. n. 7 and Fac. n. 8). For reasons of confidentiality, we do not disclose the actual names of faculties.

To illustrate the procedure, we adapt the example presented in Barra and Zotti (2016). Let us consider two faculties, 1 and 2 (see Figure 1, below).

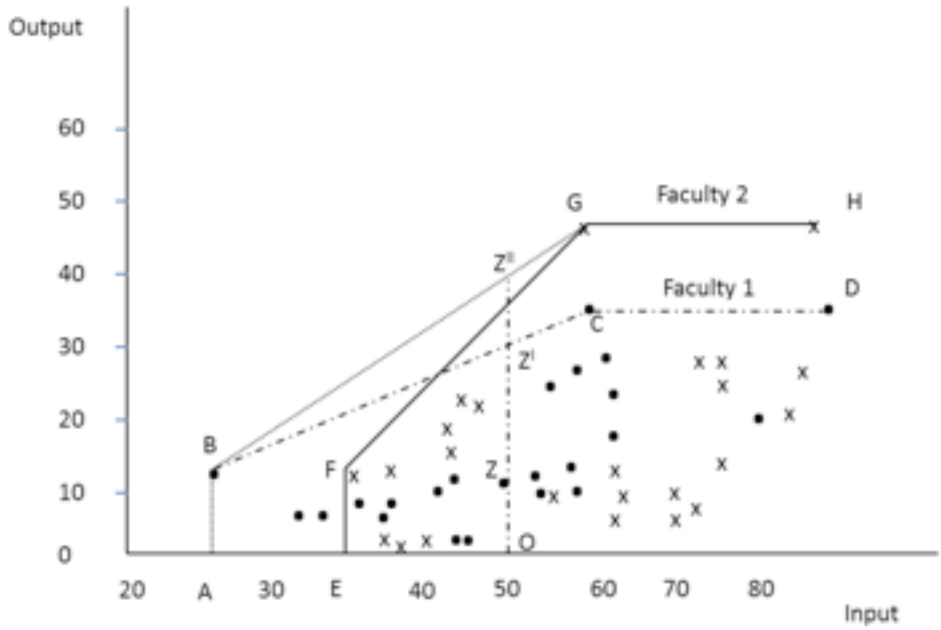

Figure 1 – Disentangling student and faculty effects

We construct the following measures of efficiency:

a) A measure that captures the performance of each student against the best performance defined by a frontier calculated for the students' population in the same subject area (EFF). This student-specific measure is defined as the radial distance of student Z from the boundary ABGH and is measured by the ratio (OZ/OZ''), where Z'' is the maximum output that student Z could have attained, given her input level. The student-specific efficiency represents the proportion of the score that student Z receives, conditional upon her input and relative to the best-performing students from the same scientific area;

b) A measure that captures the performance of each student in relation to a frontier computed for the students' population in the same faculty. This measure is labelled as the "within-faculty efficiency score" (WFEFF) and is defined as the radial distance of student Z from the boundary ABCD (in case of faculty 1). It is calculated as the ratio

(OZ/OZ'), where Z' is the within-faculty maximum output that student Z could have attained, given her input level. Faculty effects are not involved in this measure, as we are comparing students belonging to the same faculty. Any differences in the outcome observed is the result of differences in effort made by the student (by definition, faculty effectiveness is identical across students);

c) The distance between the two above-considered frontiers evaluated at the input level of student Z. This is labelled as the faculty-dependent efficiency measure (FDEFF) and is defined as the ratio (OZ'/OZ''). This efficiency measure is faculty-specific and reflects the component of the students' performance affected by the extent to which a faculty member is efficient compared to the other faculty members in his or her own scientific area.

Using the above-defined efficiency measures, the overall measure of student under-attainment can be defined as (OZ/OZ'')=(OZ/OZ')(OZ'/OZ'') where (OZ/OZ') is the measure of under-attainment attributable to the student's own effort while (OZ'/OZ'') is the measure of under-attainment attributable to the faculty to which the student belongs.

The frontiers against which the academic performance of each student is measured are computed by using DEA, first proposed by Charnes et al. (1978, 1981). DEA is a well-established methodology for the measurement of efficiency. It does not require a specific functional form, although it imposes some assumptions about the production technology (for more theoretical details on DEA, see Coelli et al., 1998). In this paper, we focus on technical efficiency computed by using an output-oriented DEA model with variable returns to scale (DEA-VRS).[4] DEA-VRS is to be preferred in our case, as suggested by Agasisti (2011), who argues that the assumption of constant returns to scale is restrictive in a university setting because it is reasonable to assume that the "dimension (number of students, amount of resources, etc.) plays a major role in affecting the efficiency" (Agasisti, 2011, p. 205). Johnes (2006a) also points out that "measurement scales of attainment are arbitrary in the educational context, and if one student's A level score is n times another's, and if both students are efficient, then there is no reason to expect that the degree result of the first student will also be n times that of the second" (Johnes, 2006a, pp. 91-92). An output-oriented model has been adopted following Agasisti and Dal Bianco (2009), who claim that "as Italian universities are increasingly concerned with reducing the length of studies and improving the number of graduates, in order to compete for public resources,

[4] Also, Portela and Thanassoulis (2001), Thanassoulis and Portela (2002), Johnes (2006a; 2006b), and Barra and Zotti (2014) rely on an output-oriented DEA-VRS.

the output-oriented model appears the most suitable to analyse higher education teaching efficiency" (Agasisti and Dal Bianco, 2009, p. 487). Johnes (2006a) is also in favour of the output-oriented model, as "in a given year, once an individual student is at university, his characteristics (both social and academic) are fixed, and therefore his efficiency (in terms of academic achievement at university) is maximised by maximizing outputs subject to his given level of inputs" (Johnes, 2006a, p. 91).

Mathematically, the DEA-VRS model we use is:

$$Maximize_{\mu,v} \quad \theta_0 = \sum_{J=1}^{J} v_j x_{j0} \tag{1}$$

$$Subject\ to \quad \sum_{i=1}^{I} \mu_i y_{i0} = 1 \tag{2}$$

$$\sum_{i=1}^{I} \mu_i y_{ik} - \sum_{j=1}^{J} v_{jk} x_{jk} \leq 1\ for\ all\ k = 1,2,\dots,n \tag{3}$$

$$I'\mu_i = 1; \;\; \mu\, \epsilon\, [0,1];\ \forall \mathrm{i} = 1,\dots,\mathrm{I}; \tag{4}$$

$$I'v_j = 1; \;\; v\, \epsilon\, [0.1]; \forall j = 1,\dots,J; \tag{5}$$

where θ_0 denotes the efficiency score of the DMU_0 under analysis, n is the number of DMUs under analysis, I and J represent respectively the number of outputs and inputs; $Y_k = \{y_{1k}, y_{2k}, \dots, y_{ik}, \dots, y_{Ik}\}$ is the vector of outputs for DMU k with y_{ik} being the value of output i for DMU k; $X_k = \{x_{1k}, x_{2k}, \dots, x_{jk}, \dots, x_{Jk}\}$ is the vector of inputs for DMU k with x_{jk} being the value of input j for DMU k; finally, μ_i and v_j denote the weights for output i and for input j, respectively. DMU k is efficient if $\theta_k = 1$. In order to estimate the DEA models, we rely on the Benchmarking package from freeware R.

Once the efficiency scores are calculated, in a second stage of the analysis we apply the fractional regression model (FRM) described in Ramalho et al. (2010) in order to assess the links between these scores and some of their potential determinants and to specifically explore their evolution and determination related to the to the 2007 policy reform. We focus in particular on the nexus between faculty-specific effects and a set of faculty-level indicators. Very often, in the literature, DEA efficiency scores are related to exogenous factors through a regression model. Ramalho et al. (2010) argue that the traditional linear or Tobit approaches are not based on the data-generating process appropriate for DEA scores, whereas the FRM avoids this pitfall. This model was proposed in Papke and Wooldridge (1996) to treat dependent variables bounded between zero and one, regardless of whether these boundary values are observed. As DEA scores are relative measures of efficiency, they are bounded variables resulting from a normalising data-generating process mapping them onto the interval $]0;1]$ (see McDonald, 2009). Hence, under the assumption that DEA scores can be treated as descriptive measures of the relative performance of units in the sample, the FRM is the most natural way of modelling DEA scores. The FRM only requires the assumption of a functional form for y that imposes the desired constraints on the conditional mean of the dependent variable. Let y be a binary or a fractional outcome, respectively, defined as $y \in \{0,1\}$ or $y \in [0,1]$, and x a vector of exogenous variables. The FRM is defined by the following conditional expectation of y given x:

$$E(y|x) = G(x\theta) \tag{6}$$

where $G(\cdot)$ is some nonlinear function satisfying $0 \leq G(\cdot) \leq 1$. Papke and Wooldridge (1996) suggest as possible specifications for $G(\cdot)$ any cumulative distribution function, such as those commonly used to model binary data. The most obvious choices are the logit and probit functional forms. However, there are other functional forms such as the loglog and complementary loglog (hereafter cloglog) specifications, where $G(x\theta)$ is respectively equal to $G(x\theta) = e^{e^{-x\theta}}$ and $1 - e^{e^{-x\theta}}$. In all cases the (conditional) partial effects of a unit change in x_j are given by:

$$\partial E(y|x)/\partial x_j = q_j g(x\theta) \tag{7}$$

While the traditional logit and probit models approach zero and unity at the same rate, the (asymmetric) cloglog (loglog) model increases slowly (sharply) at small values of G $(\cdot)$ and sharply (slowly) when $G(\cdot)$ is near unity. More precisely, the maximum partial effects produced by logit and probit models are achieved at $E(y|x) = 0.5$ and are symmetric around that point: for example, the effect of x_j on $E(y|x)$ is the same for $E(y|x) = 0.05$ and $E(y|x) = 0.95$. On the contrary, in the cloglog (loglog) model, the greatest impact of a change in x_j occurs on DMU's with $E(y|x) > 0.5$ ($E(y|x) < 0.5$). All FRM specifications can be estimated consistently through the Quasi Maximum Likelihood estimator (Papke and Wooldridge, 1996). For this purpose, we relied on the FRM module implemented in Stata 13 (Ramalho et al., 2010).

Ramalho et al. (2010, 2014) also discuss several specification tests, which can help researchers choose among alternative specifications of the FRM. Here we rely on the RESET test, customarily used to detect functional form misspecifications.

3.2. The Data

We use a unique administrative dataset of 49,394 first-year students enrolled in a large public university in the southern Italy for each academic year starting from 2002–2003 until 2010-2011. The dataset is not a panel, but a repeated cross section. We do not follow the same students over the years; on the contrary, every year a new cohort of first-year students enters our data-set.[5]

More precisely, we have data over the academic years 2002/2003 (year 2003), 2003/2004 (year 2004), 2004/2005 (year 2005), 2005/2006 (year 2006), 2006/2007 (year 2007), 2007/2008 (year 2008), 2008/2009 (year 2009), 2009/2010 (year 2010), and 2010/2011 (year 2011). This dataset covers a rather remarkable time span; for our purposes, the main point is that it has a non-negligible number of observations before and after the 2007 reform. We have chosen to focus on the performance of first-year students for two reasons: the transition between the first and the second year has been considered as one of the weaknesses of the Italian higher education system and therefore as one of the main checkpoints for evaluating the regularity of the educational path (CNVSU, 2011). Moreover, there is evidence suggesting that the academic performance at the end of the first year is a good predictor of future academic performance and the students' likelihood of obtaining a good degree (see again CNVSU, 2011).

This university is multi-campus, and over the sample years, about 40,000 students were registered in total while the number of academic staff was about 900. In the period

[5] This makes it impossible to carry out some kinds of traditional dynamic DEA analysis, such as calculating a Malmquist index based upon individual observations.

under scrutiny, it included nine faculties[6] with around 50 teaching programmes. In order to give a further idea of the size and the financial commitment of the institution, in the last decade about 90,000,000 euros have been invested every year on human resources (both academic and non-academic). The total university turnover has been fluctuating in the same period at around 100,000,000 euros. Students mostly come from the neighbouring area and are from a middle-class background. The University has its headquarters in a mid-size city that lies a few kilometres east of the main city in the area—a city whose population is slightly above 100,000 inhabitants, and whose income per capita lies around the national mean value—to which it is well connected through a motorway. In terms of structure, in the period under consideration the University was organised into departments and faculties, with departments overseeing research and faculties being in charge of teaching provision and management.

The dataset gathers information about the students' basic demographics (gender, age), educational background and pre-enrolment characteristics (type of high school attended, score gained on the high school final exams), households' financial conditions (family's self-declared income), and general information about the university careers and performances (exams passed and credits acquired). Some descriptive statistics are provided in Table 2.

Table 2 – Descriptive statistics (unless otherwise stated, values are 2033-2011 averages)

High school grades (mean value)	78.90	
High school grades (F) (mean value)	80.69	
High school grades (M) (mean value)	76.75	
Females (n/%)	28,969	54.5%
Males (n/%)	24,190	45.5%
Non-vocational high school (lyceum) (n/%)	26,058	49.0%
Vocational high school (technical) (n/%)	15,696	29.5%
Vocational high school (professional) (n/%)	11,405	21.5%
Household Income (low – Median value € 7,000.00) – (n/%)	13,675	25.7%
Income (medium-high – Median value € 21,500.00) (n/%)	39,484	74.3%
Credits (mean)	23.72	
Credits - weighted by grades (mean)	595.31	

Note: All data have been sourced from the Statistical Office of the University under analysis.

[6] In 2007, faculties of Law throughout Italy decide to introduce a separate regime of the credit system. Subsequently, all these faculties (including the one in our University) were excluded from our administrative dataset from 2007 onwards. In the empirical analysis of this paper, we chose to exclude the faculty of Law altogether from our sample.

3.3 The production set

The specification of our production set (see Table 3) is quite simple and is in line with the empirical literature in this field (Johnes, 2006a; Barra and Zotti, 2014). Our input side basically relates to the quality of students at admission, as proxied by a) the grade each student has obtained at the national exam held at the end of the high school cycle and managed by the Department of Education for the whole country (HSG), and b) a binary variable, related to the student secondary school track, taking the value of 1 in case the student attended a non-vocational (i.e. lyceum) high school and 0 otherwise (i.e. technical and professional high school). Many studies show that the secondary school track and the final grade at the end of secondary school cycle is correlated with students' performance at university (Boero et al., 2001; Smith and Naylor, 2001; Des Jardins et al., 2002; Di Pietro, 2004; Arulampalam et al., 2004; Di Pietro and Cutillo, 2008; Lassibille, 2011). Hence both variables can be considered as a good proxy of the knowledge and skills of students when entering tertiary education.

Furthermore, other personal characteristics (mainly income of the family, gender, and age; see Montmarquette et al., 2000; Smith and Naylor, 2001; McNabb et al., 2002) are known to be related to students' academic performance, albeit in a less systematic manner. We use two of these variables, gender (GEN) and the level of self-reported household income (INC) for sorting students in smaller groups to which we separately apply DEA (see Thanassoulis, 1999; Portela and Thanassoulis, 2001; Thanassoulis and Portela, 2002; Johnes, 2006a, 2006b; Barra and Zotti, 2014; for a similar choice). In order to avoid excessive fragmentation of the sample, we neglect other potential separating variables available in our dataset. Our sample basically includes only students of White Italian background under a full-time regime, and information about the distance between students' residence and the university campus was not reliable. Furthermore 78,2% of the students is concentrated between 18 and 20 years of age.

Our baseline measure of output is the *number of credits obtained at the end of the first year (CREDITS)*. The choice of this variable is strictly related to the institutional setup of Italian universities. Credits obtained at the end of the first year matter for a single university more than any other measure of students' academic achievement, as funding from central government is conditional upon them.[7]

[7] We recall that main parameters that the Italian Ministry of Education has been using in the teaching quality assessment include: a) the share of students enrolled in the second year, having already obtained a given

Table 3 – Specification of outputs and inputs in DEA (baseline model)

Inputs	**Output**
HSG: high school grades HST: type of high-school track (binary variable) GEN: gender INC: Income class (binary variable differentiating low- from middle-high household incomes).	CREDITS: sum of credits at the end of the first year

Note: Gender (GEN) and household income (INC) are not proper inputs but are used to categorise the observations so that DEA can be run on the separate subgroups (Thanassoulis, 1999).

4. The main results

4.1 The DEA Results

To repeat, we rely on an output-oriented DEA-VRS model in order to compute:

1) a within-faculty efficiency score measure of student performance, which has been computed by comparing its own achievement at the end of the first year to the performance of other students registered in the same faculty (WFEFF);
2) a student-level measure of performance computed by comparing the students attainment at the end of the first year to the performance of the other first year students registered in the same academic area (EFF).

The ratio between 2) and 1) provides a measure of the so-called faculty-dependent effect (FDEFF, a student-level measure of under-attainment due to the under-performance of the student's faculty with respect to the other faculties in the same subject area). In this Section, we shall present evidence about the evolution and the determination of the FDEFF and the WFEFF scores, linking them to the 2007 policy reform.

Before computing our DEA scores, we have tested for the presence of outliers in the sample, as DEA is very sensitive to their presence in the sample. For this purpose, we have used a procedure similar to the approach used by Thanassoulis (1999), which identifies and excludes students with exceptional achievement (i.e. outliers) from the

number of credits in the first year, and b) the share of students who do not obtain any credits or pass any exam (i.e. inactive students) at the end of the first year (Ministerial Decree 18 October 2007, n. 506; CNVSU, DOC 07/2009).

sample by relying on the concept of super-efficiency (Andersen and Petersen, 1993). Specifically we computed the super-efficiency score (see also Thanassoulis et al. 2011) for each student in our sample in order to identify those who could be considered as overachievers. We have then computed Tørgersen's rho (Tørgersen et al., 1996), which measures the share of the production set spanned by an efficient observation. Finally, we searched for observations with a super-efficiency score below 0.5 (as a signal of anomalous achievement) *and* a rho above 0.05 (observations spanning more than 5% of the production set can be considered as influential). Around 37 observations were identified as exceptional and therefore excluded from the estimation of DEA models.

Table 4 presents the values of the FDEFF scores averaged over year and faculty cells. We provide mean faculty values (at the bottom of the table) and, for each year, the absolute deviations from these mean faculty values. This format has two advantages. It first highlights the significant mean efficiency differences among faculties. Interestingly both the worst and the best performing faculties are in Pure and Applied Sciences. Fac. n. 1 records the lowest values of the FDEFF scores while the opposite is true for Fac. n. 3 (obviously, this does not necessarily imply that credits are lower in Fac. n. 1 than in Fac. n. 3, but that the latter puts to better use its inputs). Secondly, this presentation clearly suggests that the faculty-dependent measures of the students' performance have systematically (although only mildly) worsened from 2007 onwards.

Table 4. FDEFF scores from baseline DEA model, mean values by year and faculty

	PURE AND APPLIED SCIENCES			HUMANITIES			SOCIAL SCIENCES		*Annual means*
Year	*Fac. n. 1*	*Fac. n. 2*	*Fac. n. 3*	*Fac. n. 4*	*Fac. n. 5*	*Fac. n. 6*	*Fac. n. 7*	*Fac. n. 8*	
2003	0.005	0.016	-0.001	0.007	-0.001	0.001	0.002	-0.006	**0.004**
2004	0.022	0.011	-0.001	0	0.003	0.004	-0.009	-0.004	**0.004**
2005	0.01	0.006	-0.001	0.005	0	0.015	-0.001	-0.002	**0.005**
2006	0.006	-0.002	-0.002	-0.002	0.008	-0.003	-0.001	-0.006	**0.001**
2007	-0.034	0.003	-0.001	0.001	-0.008	0.004	0.008	-0.007	**-0.004**
2008	-0.013	-0.003	0	0.002	0	0.001	0.004	-0.001	**-0.001**
2009	-0.04	-0.009	0.002	-0.009	0.004	-0.006	0	0.002	**-0.008**
2010	0.025	-0.005	0.001	-0.005	-0.004	-0.004	-0.002	0.011	**0.001**
2011	-0.003	-0.011	0.003	-0.012	0.001	-0.014	-0.001	0.019	**-0.005**
Fac. means	**0.631**	**0.562**	**0.993**	**0.835**	**0.945**	**0.821**	**0.947**	**0.866**	

Notes: The figures in the cells are the values of the FDEFF indexes averaged by year and faculty. The DEA model employed to compute these indexes uses as inputs the high school grades and the type of school track, and the sum of credits at the end of the 1st year as output. Gender and Income are used to categorise the observations so that DEA can be run on the separate subgroups (Thanassoulis, 1999).

4.2 The basic FRM results

In order to shed some light on the links between the worsening of FDEFF after 2007 and the institutional reform that occurred at the same time in Italian higher education, and, more generally, to gain some knowledge about the determinants of FDEFF, we now proceed to the second part of the empirical analysis. We test whether student-level FDEFF scores over time and faculties are associated with faculty and year dummies, as well as with some indicator of faculty resources.

For the latter, we relied on a variable rather customary in the literature, the student-professor ratio, that is the ratio between number of students and total academic staff. More specifically, we use the following weighted student-professor ratio (WSPR):

$$WSPR_i = \left(\frac{Students\ enrolled_i}{(Equivalent\ Personnel_i) * (1 + External\ Resources_i)}\right) \tag{8}$$

where $Students\ enrolled$ is the total number of students enrolled in faculty i; $Equivalent\ Personnel$ is the sum of academic staff (disentangled in four categories, namely, full professors, associate professors, assistant professors, and lecturers). In adding up the number of various components of academic staff, full and associate professors were weighed twice as much as lecturers reflecting the institutional rules presiding to the attribution of teaching load to these categories of academic staff. Attempts based on other measures of students or staff, or on teaching-related expenses, will be presented in the next sub-section as robustness checks; $External\ Resources$ is a weighted measure of the resort to external teaching measured as the number of course units taught by external staff over the total number of course units.

In the analysis of faculty-dependent efficiency, we should also allow for the possibility that different faculties may recruit students with different ability (or other characteristics), and this may affect the variation of the faculty effects over time and across faculties. In order to assess the relevance of these effects, we also included some individual characteristics of the students in the estimates: HSG (the high-school grade), plus dummies for high-school track, gender, and income class.[8]

We relied on a Quasi Maximum Likelihood approach to estimate various types of FRMs. In particular, we focused on a fractional logit, a fractional loglog and a fractional cloglog. We report only results from our preferred models[9] (qualitatively, other models give the same results), and present the conditional partial effects evaluated at the mean (or, for dummy variables, at the mode) and the associated z-values. Table 5 illustrate the estimates obtained for a simple specification, both without and with student-level variables.

Table 5. Baseline FRM estimates on FDEFF scores.

Variables	dy/dx	z	dy/dx	z
	Loglog model		Cloglog model	
WSPR	-0.0003	-1.74	-0.0002	-1.64
Faculty n. 1	0.133	23.28	0.089	20.42
Faculty n. 2	-0.034	-9.91	-0.004	-0.55
Faculty n. 3	0.127	23.62	0.088	15.26
Faculty n. 4	-0.042	-13.53	0.001	0.15
Faculty n. 5	-0.167	-35.79	-0.150	-14.86
Faculty n. 6	-0.198	-50.72	-0.244	-38.21
Faculty n. 7	0.414	32.67	0.212	41.51
Year 2004	-0.003	-1.14	-0.003	-1.38
Year 2005	-0.001	-0.56	-0.011	-3-34
Year 2006	-0.007	-2.95	-0.014	-5.17
Year 2007	-0.004	-1.70	-0.013	-4.25
Year 2008	-0.004	-2.23	-0.010	-3.70
Year 2009	-0.007	-2.71	-0.010	-3.49
Year 2010	-0.004	-1.34	-0.008	-2.54
Year 2011	-0.006	-1.54	-0.010	-2.37
HSG			0.001	7.00
Classical Lyceum			-0.004	-2.31
Linguistic Lyceum			0.009	1.81
Technical Institution			-0.019	-3.41
Professional Institution			-0.019	-3.38
Other Institutions			-0.024	-4.25
Female			-0.055	-2.65
Low income			0.065	4.19
R^2 -type measure	0.549		0.621	
RESET test	0.069		0.299	

Notes: Quasi Maximum Likelihood estimator. The first columns of each estimate refer to the regressor partial effects evaluated at the mean (or, for dummy variables, at the mode), while the second columns include the z-values. Standard errors are clustered by faculty and year.
Fac. n. 8, Year_2002, scientific lyceum, male, middle-high income, are the reference groups for dummy variables.
The RESET test is the standard test of misspecification of the functional form. We only report (in brackets) the p-values of this test.

[8] We also made attempts with age, nationality, distance between residence and campus, and part-time regime, but these indicators were always highly insignificant.
[9] Changes in the preferred models may depend on the shape of the distribution function for the conditional efficiency scores.

In both specifications, the partial effects show a negative and (weakly) significant association between the FDEFF scores and WSPR. It hence appears that reducing the number of staff per student has a detrimental (although not strongly so) effect on faculty-specific efficiency. The partial effects associated with the year dummies convey the impression that FDEFF decreases over time, even before the 2007 reform. Yet, to shed full light on the role that the 2007 reform may have played in driving the evolution of the FDEFF scores, it is necessary to allow for the possibility that the reform (by increasing the workload of professors) has affected the link between WSPR and faculty effectiveness. To see this, we interact WSPR with a dummy variable (REFORM) equal to one for all the years after 2007 and to zero otherwise.

The new estimated partial effects are shown in Table 6, respectively without and with student-level variables.

*Table 6. FRM estimates on FDEFF scores (estimates with WSPR*Reform interaction).*

Variables	dy/dx	z	dy/dx	z
	Loglog model		Cloglog model	
WSPR	-0.0006	-2.93	-0.0005	-2.37
WSPR*Reform	0.0003	-2.37	0.0002	1.71
Faculty n. 1	0.130	22.58	0.088	19.98
Faculty n. 2	-0.036	-9.98	-0.005	-0.65
Faculty n. 3	0.124	22.06	0.086	14.53
Faculty n. 4	-0.043	-14.52	0.001	0.12
Faculty n. 5	-0.170	-33.08	-0.152	-14.60
Faculty n. 6	-0.202	-45.05	-0.246	-38.84
Faculty n. 7	0.410	30.58	0.210	35.83
Year 2004	-0.004	-1.97	-0.005	-1.80
Year 2005	-0.003	-1.36	-0.012	-3.82
Year 2006	-0.009	-4.10	-0.016	-5.90
Year 2007	-0.006	-2.14	-0.014	-4.70
Year 2008	-0.015	-3.29	-0.018	-3.53
Year 2009	-0.017	-3.74	-0.018	-3.42
Year 2010	-0.014	-3.36	-0.016	-3.09
Year 2011	-0.016	-4.02	-0.017	-2.99
HSG			0.001	6.99
Classical Lyceum			-0.004	-2.33
Linguistic Lyceum			0.009	1.81
Technical Institution			-0.019	-3.41
Professional Institution			-0.019	-3.40
Other Institutions			-0.024	-4.25
Female			-0.054	-2.64
Low income			0.065	4.19
R^2 -type measure	0.549		0.622	
RESET test	0.445		0.297	

Notes: Quasi Maximum Likelihood estimator. The first column of each estimate includes the regressor partial effects evaluated at the mean (or, for dummy variables, at the mode), while the second column includes the z-values. Standard errors are clustered by faculty and year.

Fac. n. 8, Year_2002, scientific lyceum, male, middle-high income, are the reference groups for dummy variables.

The RESET test is the standard test of misspecification of the functional form. We only report (in brackets) the p-values of this test.

The RESET tests are now more satisfactory, suggesting that allowing for a 2007 structural break improves the specification. The partial effects associated with the interactions between WSPR and REFORM are similar across equations, and imply that the link between faculty-specific efficiency and WSPR was partially severed after 2007. Overall, the partial effect of the WSPR variable continues to be negative and is more significant than before. Importantly, the partial effects associated with the time dummies after 2007 are now negative and larger in absolute size. This means that for a given level of the WSPR, the increasing workload following the 2007 reform brought about a reduction in faculty-specific efficiency. Hence, no unused resources appear to have existed at the time of the reform (a state of affairs that the 2008 turnover freeze is likely to have perpetuated). Furthermore it seems that the reform has strongly weakened the link between faculty-level efficiency and WSPR (most likely because hours per professor increased, lowering the marginal productivity of professors as well their efficiency). Hence, at least part of the decrease of the FDEFF scores across all the faculties can be ascribed to the increase of the WSPR ratio over time.

The above evidence suggests that the effects of the 2007 reform are permanent. However, further research is needed to know whether the decrease in faculty efficiency is permanent or transient. Indeed, faculties may have decided to change the organisation of programmes and modules after the reform, leading to a recovery in faculty-specific efficiency. For instance, faculties may decide to cut the number of modules so that academic staff may reduce its workload (eventually concentrating) on supporting the teaching of first-year students. In order to see whether this is the case, we provide below a more complex characterisation of the impact of policy-related changes. We interact WSPR with annual time dummies for all years starting from 2004. This specification is reminiscent of the empirical strategy adopted in Autor (2003). It is useful to assess whether the effects of a policy change are permanent or transitory, as well as to gauge whether policy outcomes are already anticipated in the behaviour of the economic agents under scrutiny. Table 7 considers two alternative specifications without and with student-level variables.

*Table 7. FRM estimates on FDEFF scores (estimates with WSPR*annual dummies interaction)*

Variables	dy/dx	z	dy/dx	z
	Loglog model		Cloglog model	
WSPR	-0.0006	-3.11	-0.0004	-2.41
WSPR*Year_2004	-0.0002	-1.12	-0.0003	-0.89
WSPR*Year_2005	0.0002	0.79	0.0001	0.24
WSPR*Year_2006	-0.0001	-0.26	0.0000	0.03

WSPR*Year_2007	0.0002	0.81	0.0002	1.06
WSPR*Year_2008	0.0003	2.09	0.0003	1.77
WSPR*Year_2009	0.0004	3.35	0.0002	1.34
WSPR*Year_2010	0.0004	2.43	0.0003	1.76
WSPR*Year_2011	0.0003	1.72	0.0002	1.24
Faculty n. 1	0.130	19.52	0.088	17.68
Faculty n. 2	-0.036	-8.47	-0.005	-0.62
Faculty n. 3	0.124	18.80	0.086	13.24
Faculty n. 4	-0.043	-15.34	0.001	0.08
Faculty n. 5	-0.170	-28.23	-0.152	-14.23
Faculty n. 6	-0.202	-33.95	-0.246	-35.91
Faculty n. 7	0.410	28.28	0.209	29.03
Year 2004	0.001	0.27	-0.005	
Year 2005	-0.010	-1.63	-0.018	
Year 2006	-0.008	-2.12	-0.019	
Year 2007	-0.012	-2.17	-0.023	
Year 2008	-0.014	-4.87	-0.021	
Year 2009	-0.018	-5.04	-0.020	
Year 2010	-0.018	-4.21	-0.020	
Year 2011	-0.017	-3.52	-0.020	
HSG			0.001	7.00
Classical Lyceum			-0.004	-2.33
Linguistic Lyceum			0.009	1.80
Technical Institution			-0.019	-3.41
Professional Institution			-0.019	-3.39
Other Institutions			-0.024	-4.24
Female			-0.054	-2.64
Low income			0.065	4.19
R^2-type measure	0.549		0.622	
RESET test	0.405		0.297	

Notes: Quasi Maximum Likelihood estimator. The first column of each estimate includes the regressor partial effects evaluated at the mean (or, for dummy variables, at the mode), while the second column includes the z-values. Standard errors are clustered by faculty and year.
Fac. n. 8, Year_2002, scientific lyceum, male, middle-high income, are the reference groups for dummy variables.
The RESET test is the standard test of misspecification of the functional form. We only report (in brackets) the p-values of this test.

According to the empirical evidence summarized in Table 7 (both specifications tell the same story) there are no policy anticipation effects and the 2007 reform is likely to have had at least a fairly long-term negative effect on faculty-specific efficiency. No unused resources appear to have existed at the time of the 2007 reform, and the increasing workload following the 2007 reform brought about a reduction in the faculty-specific effect. The 2007 reform did not target recruitment, but was followed by the 2008 law that decisively froze the turnover of the academic staff. The result was that the workload of professors could not be readily brought back to the pre-2007 levels, with the implication that the negative impact on students' performance of the 2007 reform has lasted for a long time.

4.3 Some robustness checks

We now proceed to assess whether the story that emerges from the above estimates is robust with respect to a series of robustness checks. In doing so, we shall concentrate on a particular specification, with interaction terms between resource indicators and the 2007 reform and with student-level variables.

As a first robustness check, we rely on different indicators in order to measure the faculty resources. In Table 8 we consider SPR, a student-professor ratio indicator simply based on the ratio between number of students and total (unweighted) academic staff, and SPR-1st, the ratio between number of first-year students (our actual focus of interest) and number of professors. We do not have measures of academic staff teaching in the first year, but in this university only professors (as opposed to lecturers) usually take first-year teaching.

Table 8. FRM estimates on FDEFF scores (with various resource indicators among regressors)

Variables	dy/dx Cloglog model	z	dy/dx Cloglog model	z
SPR	-0.0002	-1.95		
SPR*Reform	0.0001	1.53		
SPR-1st			-0.0006	-2.26
SPR-1st*Reform			0.0003	1.60
Faculty n. 1	0.089	20.05	0.089	19.03
Faculty n. 2	-0.005	-0.57	-0.007	-0.80
Faculty n. 3	0.087	14.08	0.087	15.50
Faculty n. 4	0.001	0.06	0.003	0.27
Faculty n. 5	-0.151	-14.27	-0.153	-14.91
Faculty n. 6	-0.245	-39.89	-0.246	-35.58
Faculty n. 7	0.210	35.03	0.211	34.45
Year 2004	-0.004	-1.71	-0.006	-1.74
Year 2005	-0.012	-3.70	-0.012	-3.59
Year 2006	-0.016	-5.55	-0.014	-5.56
Year 2007	-0.014	-4.54	-0.011	-4.11
Year 2008	-0.018	-3.25	-0.013	-3.62
Year 2009	-0.018	-3.21	-0.013	-3.39
Year 2010	-0.016	-2.91	-0.012	-2.95
Year 2011	-0.018	-2.88	-0.014	-2.96
HSG	0.001	6.99	0.002	6.99
Classical Lyceum	-0.004	-2.34	-0.005	-2.32
Linguistic Lyceum	0.009	1.81	0.001	1.82
Technical Institution	-0.019	-3.42	-0.019	-3.40
Professional Institution	-0.019	-3.39	-0.019	-3.42
Other Institutions	-0.024	-4.26	-0.023	-4.24
Female	-0.054	-2.64	-0.055	-2.64
Low income	0.065	4.19	0.066	4.19
R^2-type measure				
RESET test	0.298		0.297	

Notes: Quasi Maximum Likelihood estimator. The first column of each estimate includes the regressor partial effects evaluated at the mean (or, for dummy variables, at the mode), while the second column includes the z-values. Standard errors are clustered by faculty and year.

Fac. n. 8, Year_2002, scientific lyceum, female, middle-high income, are the reference groups for dummy variables.
The RESET test is the standard test of misspecification of the functional form. We only report (in brackets) the p-values of this test.

Although SPR is less significant than for the baseline specification of § 4.2, this indicator and its interaction with the 2007 reform follow the pattern set by our preferred indicator. SPR-1st is also significant and the qualitative properties of these estimates are very close to those obtained in § 4.2.

Secondly, we proceed to assess FRM with different dependent variables. In the previous analysis we relied on DEA scores computed from a model in which output was measured by the total number of credits awarded to a student at the end of the first year. This indicator can be strongly justified on institutional grounds. It matters for a single university more than any other measure of students' academic achievement, as funding from the central government is basically made conditional upon it. However, it can be argued that in order to measure the academic achievements of first-year students, a proper indicator of academic performance should consider not only the total number of credits but also the grades awarded during the first year. Therefore, we also relied on an output-oriented DEA-VRS model with a different output indicator: the sum of credits obtained at the end of the first year weighted by the grades associated with each first-year exam (this indicator also has some kind of institutional justification, as students can obtain a reduction in fees if they reach a given value for this indicator). We accordingly recomputed our student-level measures of academic performance using our new measure of output.

Another interesting issue is whether the within-faculty efficiency score component of the academic performance of first-year students has increased after 2007 to compensate for the decrease of the faculty-dependent component. In other words, first-year students may have reacted to the increasing WSPR and to the worsening of teaching efficiency at the University by increasing their own effort, with the result that overall, their academic performance at the end of the first year may not have worsened following the reforms. To be able to test this hypothesis, the dependent variable should the within-faculty efficiency score performance of the first-year students (WFEFF).

In Table 9 we show FRM estimates with, respectively, FDEFF scores from the DEA model with credits weighted by grades as output, and WFEFF scores from the baseline DEA model, as dependent variables.

Table 9. FRM estimates on various types of efficiency scores

	Dep. Var.: FDEFF scores from DEA model with credits weighted by grades as output		Dep. Var.: WFEFF scores from baseline DEA model	
Variables	dy/dx	z	dy/dx	z
	Cloglog model		Cloglog model	
WSPR	-0.0003	-2.89	0.000	0.02
WSPR*Reform	0.0005	6.16	0.001	1.47
Faculty n. 1	0.137	45.05	-0.042	-1.85
Faculty n. 2	0.075	13.53	0.014	0.73
Faculty n. 3	0.043	11.06	-0.015	-0.58
Faculty n. 4	0.125	14.95	-0.005	-0.31
Faculty n. 5	-0.076	-10.05	0.146	5.67
Faculty n. 6	-0.147	-29.29	0.054	1.99
Faculty n. 7	0.207	85.59	-0.066	-2.10
Year 2004	-0.003	-1.50	-0.009	-0.35
Year 2005	-0.003	-1.75	-0.001	-0.04
Year 2006	-0.006	-4.23	-0.029	-1.36
Year 2007	-0.002	-1.68	-0.029	-1.43
Year 2008	-0.013	-4.28	-0.068	-1.77
Year 2009	-0.015	-4.81	-0.085	-2.11
Year 2010	-0.013	-4.57	-0.083	-2.02
Year 2011	-0.014	-4.19	-0.059	-1.43
HSG	0.002	8.17	0.005	22.06
Classical Lyceum	0.003	1.29	-0.006	-1.13
Linguistic Lyceum	-0.001	-0.49	-0.049	-7.20
Technical Institution	-0.025	-4.62	-0.064	-9.51
Professional Institution	-0.029	-4.23	-0.115	-12.78
Other Institutions	-0.037	-4.39	-0.072	-10.00
Female	-0.017	-1.26	0.046	10.78
Low income	0.006	1.22	0.050	7.70
R^2-type measure				
RESET test	0.0008		0.1033	

Notes: Quasi Maximum Likelihood estimator. The first column of each estimate includes the regressor partial effects evaluated at the mean (or, for dummy variables, at the mode), while the second column includes the z-values. Standard errors are clustered by faculty and year.
Fac. n. 8, Year_2002, scientific lyceum, male, middle-high income, are the reference groups for dummy variables.
The RESET test is the standard test of misspecification of the functional form. We only report (in brackets) the p-values of this test.

When the dependent variable is FDEFF scores from the DEA model with credits weighted by grades as output, the results are once more similar to the ones obtained in 4.2. As before, WSPR and its interaction with REFORM are significant (with an opposite sign). In this case, however, the link between faculty-specific efficiency and WSPR is totally severed after 2007. Concomitantly, the time dummies after 2007 become negative and highly significant. For a given level of the WSPR, the increasing workload following the 2007 reform brought about a strong reduction in faculty-specific efficiency even under this specification of the production set. A note of caution must be sounded in the sense that the RESET test is suggestive of some kind of misspecification.

When the dependent variable is WFEFF scores from the baseline DEA model, the partial effects associated with WSPR are now not significant (which makes sense, as these factors are equally available to all students within a faculty). The time dummies are not very significant, but tend to have a negative sign. We then conclude that first-year students did not react to the increasing WSPR and the worsening of faculty-dependent efficiency at the University by increasing their own effort.

Thirdly, it could be asked whether WSPR is an exogenous variable, since it can be manipulated in order to respond to past faculty performance (Carrieri et al. 2015). The literature does not often confront this problem, which in principle can be straightforwardly analysed and faced through instrumental variable techniques. Allowing and testing for endogeneity is not straightforward in FRM's. However McDonald (2009) points out that FRM's can be approximated to a satisfactory extent by OLS. We thus proceed to compare OLS and IV estimates in order to appraise the endogeneity of WSPR (and WSPR*Reform). We use as instrumental variables the one-year lagged values of WSPR, and WSPR*Reform, plus the current and lagged values of teaching expenses per academic staff unit. The latter should be valid instruments, inasmuch as there is no a priori reason why they should be correlated with faculty efficiency, while they may be good predictors of the student-professor ratio because of their association with the seniority structure of academic staff. In Table 10, we compare the performance of OLS and IV estimation for our baseline equation.

Table 10. OLS and IV estimates on FDEFF scores.

Variables	Coef.	t	Coef.	t
	OLS		IV	
WSPR	-0.0006	-2.80	-0.0008	-2.49
WSPR*Reform	0.0003	2.24	0.0004	2.29
Faculty n. 1	0.081	20.44	0.077	13.44
Faculty n. 2	-0.022	-3.47	-0.026	-3.51
Faculty n. 3	0.079	10.44	0.074	8.35
Faculty n. 4	-0.012	-1.07	-0.010	-0.95
Faculty n. 5	-0.217	-17.53	-0.222	-15.25
Faculty n. 6	-0.331	-50.74	-0.338	-45.34
Faculty n. 7	0.110	23.44	0.104	14.92
Year 2004	-0.004	-1.40	-0.006	-1.40
Year 2005	-0.015	-4.03	-0.017	-4.03
Year 2006	-0.019	-5.78	-0.021	-5.78
Year 2007	-0.017	-4.99	-0.019	-4.99
Year 2008	-0.023	-3.99	-0.028	-3.99
Year 2009	-0.024	-4.21	-0.029	-4.21
Year 2010	-0.021	-3.66	-0.025	-3.66
Year 2011	-0.024	-3.89	-0.028	-3.89
HSG	0.001	4.80	0.002	4.66
Classical Lyceum	-0.002	-1.59	-0.003	-1.61
Linguistic Lyceum	0.004	1.15	0.005	1.16
Technical Institution	-0.019	-4.20	-0.020	-4.23
Professional Institution	-0.020	-3.80	-0.020	-3.94
Other Institutions	-0.027	-4.34	-0.028	-4.41

Female	-0.054	-2.67	-0.054	-2.49
Low income	0.077	5.04	0.078	4.76
R^2	0.605			
RESET test	0.128			
HANSEN test			0.855	
C-Endogenity test			0.259	
Underidentification test			0.000	
Stock-Yogo ID test 5% critical values			11.04	

Notes: the first column of each estimate includes the regressor coefficients, while the second column includes the t-ratios. Standard errors are clustered by faculty and year.
Fac. n. 8, Year_2002, scientific lyceum, male, middle-high income, are the reference groups for dummy variables.
The RESET test is the standard test of misspecification of the functional form. We only report (in brackets) the p-values of this test.

The diagnostics presented in Table 11 clearly indicate that instruments are valid and that endogeneity is not a significant problem. What is more important, IV and OLS estimates are reasonably close to the FRM estimates from Table 5.

We have performed additional robustness tests (not reported here) that focus an indicator of faculty resources based on teaching-related expenses per student, as well as on DEA scores based on constant-returns-to-scale models. The gist of our results does not change.

5. Concluding Remarks

The fiscal austerity programmes introduced in Europe after the sovereign debt crisis in 2011 have spurred some interest in the literature, but no quantitative estimate of the impact that funding cuts to universities may have on the performance of students and the quality of their educational experience. Apparently, there is a consensus on the negative impact that funding cuts have on the students and their academic performance through increasing student-staff ratios. Very little attention has yet been paid to the fact that in reality, this may not be the case if universities can find ways to better use their academic resources. Crucially, this negative outcome really hinges on the absence of unused academic resources, which could be deployed when financial resources shrink. Furthermore the literature indicates that some governments, in an effort at rationalisation, have introduced legislation regulating workload that is regarded as problematic by many academics. For example, in the US, Cage (1995) reports on Ohio legislation requiring more time to be spent by academics in the classroom. This legislation was seen as unnecessary by many higher education institutions which did not accept the claim and also argue that early retirement and reductions in the number of full-time academics is already causing academics to work harder. Not much empirical analysis was carried out in this field, however.

To shed light upon these issues, we have focused on the Italian experience in this area and quantified the impact that two reforms (over the period 2007–2008) have had on the academic performance of students at the end of their first year in a large Italian university from 2008 onwards. These are two interesting reforms: the first one (in 2007) induced universities to increase the workload of their professors. The second one (in 2008) froze the turnover of staff, with the result that very little new staff could be recruited.

The empirical analysis has been carried out on a sample of 49,394 first-year students from academic year 2002–2003 to academic year 2010–2011; we have adapted the procedure suggested by Thanassoulis and Portela (2002) to a single-university setting and used DEA to decompose the academic performance of first-year students into a student-level within-faculty efficiency score component and a faculty-dependent component. In the second stage, we have regressed the faculty-dependent effect on a set of faculty-level indicators while controlling for the individual characteristics of the students.

The results suggest that, allowing for SPR and student-level individual controls, the faculty-specific efficiency scores have worsened after the two reforms. The increase in professors' workload has been detrimental to this efficiency, which never subsequently reverted to the pre-2007 values. Until 2011, the faculties have not implemented any action to reduce the negative impact of the two reforms on students' performance. Indeed, universities were not in a position to accommodate the reduction in staff by using existing slack because there was none left following the previous reform. Additional tests show that similar results are obtained with different specifications of the production set underlying the DEA model. At the same time, we do not observe changes in the within-faculty efficiency score component of the academic performance of the first-year students, indicating that increases in the students' effort are not compensating for the increased student-staff ratios.

What conclusions can we draw from this study? It emerges from our results that the most serious impact of the various reforms came from the increased teaching load of professors after 2007. This evidence is in agreement with the existing literature, saying that faculty inefficiency is driven mostly by the way in which the workforce is organised and utilised. This paper also supports the common view that higher student-staff ratios can have a detrimental (if small) impact on university students' performance but also suggests that this is really the case only if there is no slack in the system and if no corrective action is taken by the universities to compensate for the increasing student-staff ratios. This last point is particularly important and raises questions about the incentives faculties (and universities) have in setting up these corrective measures. As the administration and the management of the teaching programmes are devolved to faculties (in a decentralised manner), it is commonly believed that their organisation should be flexible enough to allow them to change the organisation of the programmes in such a way that students are not damaged by increasing student-staff ratios. In this respect, the paper can offer an interesting lesson, as the Italian experience indicates that the institutional link between number of programmes (and modules) and the availability of resources for recruitment of academic staff has created the conditions for perverse incentives, which have ultimately affected negatively the students' performance once more stringent funding rules were introduced. The decentralised decision structure (and the link between the recruitment and number of modules) has created the conditions for a "prisoners' dilemma" situation where no faculty is keen on changing programmes in the hope that through this route, they can have access to a larger share of the shrinking pot of available resources.

References

Abbot M, Doucouliagos C. 2003. The efficiency of Australian universities: a data envelopment analysis. *Economics of Education Review* 22: 89-97.

Agasisti T. 2011. Performances and spending efficiency in Higher education: a European comparison through non-parametric approaches. *Education Economics* 19 (2): 199-224.

Agasisti T, Dal Bianco A. 2009. Reforming the university sector: effects on teaching efficiency. Evidence from Italy. *Higher Education*, 57 (4), 477-498.

Agasisti T, Johnes G. 2010. Heterogeneity and the evaluation of efficiency: the case of Italian universities. *Applied Economics* 42 (11): 1365-1375.

Agasisti T, Bonomi F. 2014. Benchmarking universities' efficiency indicators in the presence of internal heterogeneity. *Studies in Higher Education* 39 (7), 1237–1255.

Andersen P, Petersen NC. 1993. A Procedure for ranking efficient units in data envelopment analysis. *Management Science* 39: 1261-1264.

Arulampalam W, Naylor R, Robin A, Smith JP. 2004. Hazard model of the probability of medical school drop-out in the UK. *Journal of the Royal Statistical Society*, 167 (1): 157-178.

Autor D. (2003), Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing, *Journal of Labor Economics*, 21, 1, 1-42.

Barra C, Zotti R. 2016. Managerial efficiency in higher education using individual versus aggregate level data. Does the choice of decision making unit count? *Managerial and Decision Economics* 37: 106–126.

Barra C, Destefanis S, Sena V, Zotti R. 2017. Disentangling Faculty Efficiency from Students' Effort. In Vanthienen, J., De Witte, K. (eds.). *Data Analytics Applications in Education*. Taylor and Francis: London.

Boero G, McNight A, Naylor R, Smith J. 2001. Graduates and graduate labour markets in the UK and Italy. *Lavoro e Relazioni industriali* 2: 131-172.

Bonaccorsi, A., Daraio C., 2007, University autonomy and structural constraints in the Italian system. In Bonaccorsi, A., Daraio C. (eds.), Universities and Strategic Knowledge Creation: Specialization and Performance in Europe. Cheltenham, UK: Edward Elgar.

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.

Cage, M. C. (1995) Regulating faculty workload, *Chronicle of Higher Education*, 41(19), pp. 30, 33.

Carrieri, V., D'Amato, M. and Zotti, R. (2015). On the causal effects of selective admission policies on students' performances: evidence from a quasi-experiment in a large Italian university. *Oxford Economic Papers*, 67(4), 1034–1056.

Charnes A, Cooper WW, Rhodes E., 1978, Measuring the efficiency of decision making units. *European Journal of Operational Research* 2: 429-444.

Charnes A, Cooper WW, Rhodes E. 1981. Evaluating program and managerial efficiency. An application of data envelopment analysis to program follow through. *Management Science* 27: 668-697.

CNVSU 2011. *Undicesimo Rapporto sullo Stato del Sistema Universitario*, Roma, Italy.

Coelli T, Prasada Rao DS, Battese GE. 1998. *An Introduction to Efficiency and Productivity Analysis*. Kluwer Academic Publishers: Boston.

Davidson, R. and MacKinnon, J.G., 1981, "Several tests for model specification in the presence of alternative hypotheses", *Econometrica*, 49(3), 781-793.

DesJardins SL, Ahlburg DA, McCall B. 2002. A temporal investigation of factors related to timely degree completion. *The Journal of Higher education* 73 (5): 555-581.

Di Pietro G. 2004. The determinants of university dropout in Italy: a bivariate probability model with sample selection. *Applied Economics Letters* 11 (3): 187-191.

Di Pietro G. Cutillo A. 2008. Degree flexibility and university drop-out: The Italian experience. *Economics of Education Review* 27 (5): 546-555.

Dyson RG, Allen R, Camanho AS, Podinovski VV, Sarrico CS, Shale EA. 2001. Pitfalls and protocols in DEA. *European Journal of Operational Research* 132: 245-259.

European University Association. Annual report, 2011. http://www.eua.be/activities-services/publications/eua-activity-reports.aspx. Last access: 12.04.2016.

Flegg AT, Allen DO, Field K, Thurlow TW. 2004. Measuring the efficiency of British universities: A multi-period data envelopment analysis. *Education Economics* 12: 231-249.

Gimenez VM and JL Martinez. 2006. Cost Efficiency in the University: A Departmental Evaluation Model. *Economics of Education Review* 25(5): 543-553.

Halkos G, Tzeremes NG, Kourtzidis SA. 2012. Measuring public owned university departments' efficiency: a bootstrapped DEA approach. *Journal of Economics and Econometrics* 55 (2): 1-24.

Hanushek E. A., 2014, Education production functions, The New Palgrave Dictionary of Economics - Second Edition, S. N. Durlauf, L. E. Blume, Palgrave Macmillan, The New Palgrave Dictionary of Economics Online.

Johnes G, Johnes J. 1995. Research funding and performance in UK university departments of Economics: a frontier analysis. *Economics of Education Review* 14 (3): 301–314.

Johnes J. 2006a. Measuring efficiency: a comparison of multilevel modelling and data envelopment analysis in the context of higher education. *Bulletin of Economic Research* 58 (2): 75-104.

Johnes J. 2006b. Measuring teaching efficiency in higher education: An application of data development analysis to economics graduates from UK Universities 1993. *European Journal of Operational research* 174 (1): 443-456.

Kao C, Hung HT. 2008. Efficiency analysis of university departments: an empirical study. *Omega* 36: 653-664.

Köksal G, Nalçaci B. 2006. The Relative Efficiency of Departments at a Turkish Engineering College: A Data Envelopment Analysis. *Higher Education*, 51 (2), 173-189.

Koulouris, A., Moniarou-Papaconstantinou, V. and Kyriaki-Manessi, D. (2014). Austerity measures in Greece and their impact on higher education. *Procedia - Social and Behavioral Sciences*, 147, 518 – 526.

Lassibille G. 2011. Student progress in higher education: what we have learned from large-scale studies. *The Open Education Journal* 4: 1-8.

McDonald, J. (2009), "Using least squares and tobit in second stage DEA efficiency analyses", *European Journal of Operational Research*, 197(2), 792-798.

McNabb R, Pal S, Sloane P. 2002. Gender differences in educational attainment: the case of university students in England and Wales. *Economica*, 69: 481-503.

Montmarquette C, Mahseredjian S, Houle R. 2000. The determinants of university dropouts: a bivariate probability model with sample selection. *Economics of Education Review* 20: 475-484.

Moreno AA, Tadepalli R. 2002. Assessing academic department efficiency at a public university. *Managerial and Decision Economics* 23: 385-397.

Papke, L.E. and Wooldridge, J.M., 1996, “Econometric methods for fractional response variables with an application to 401(k) plan participation rates”, *Journal of Applied Econometrics*, 11(6), 619-632.

Portela MCS, Thanassoulis E. 2001. Decomposing school and school-type efficiency. *European Journal of Operational research,* 132: 357-373.

Ramalho EA, Ramalho JJS, Henriques, PD. 2010. Fractional regression models for second stage DEA efficiency analyses, *Journal of Productivity Analysis* 34 (3): 239-255.

Ramalho, E.A., Ramalho, J.J.S. and Murteira, J. (2014), A Generalized Goodness-of-Functional Form Test for Binary and Fractional Regression Models, *Manchester School*, 82(4), 488-507.

Rizzi D. 2000. L'efficienza dei dipartimenti dell'Università Ca' Foscari di Venezia via DEA e DFA*, Economia Pubblica.* XXXI, 3, 103-136.

Sarrico CS, Dyson RG. 2000. Using DEA for planning in UK universities. An institutional perspective. *Journal of the Operational Research Society* 51: 789-800.

Sarrico CS, Dyson RG. 2004. Restricting virtual weights in data envelopment analysis. *European Journal of Operational Research* 159: 17–34.

Sarrico CS, Teixeira PN, Rosa MJ, Cardoso MF. 2009. Subject mix and productivity in Portuguese universities. *European Journal of Operational Research* 197: 287–295.

Smith J, Naylor R. 2001. Determinants of degree performance in UK universities: A statistical analysis of the 1990s student cohort. *Oxford Bulletin of Economics and Statistics* 63: 29–60.

Soliman I., Hani Soliman H. (1997). Academic Workload and Quality, Assessment & Evaluation in Higher Education, 22:2, 135-157.

Tauer LW, Fried, HO, Fry WE 2007. Efficiencies of Academic Departments within a College. *Education Economics* 15, (4), 473–489.

Thanassoulis E, Kortelainen M, Johnes G, Johnes J. 2011. Costs and efficiency of higher education institutions in England: a DEA analysis. *Journal of the Operational Research Society* 62: 1282–1297.

Thanassoulis E, Portela MCS. 2002. School outcomes: Sharing the responsibility between pupil and school. *Education Economics* 10 (2): 183–207.

Thanassoulis E. 1999. Setting achievement targets for school children. *Education Economics* 7 (2): 101-119.

Thursby J G. 2000. What do we say about ourselves and what does it mean? Yet another look at Economics department research. *Journal of Economic Literature* 38: 383–404.

Tomkins C, Green R. 1988. An experiment in the use of data envelopment analysis for evaluating the efficiency of UK university departments of accounting. *Financial Accountability & Management* 4, 147–64.

Tørgersen AM, Førsund FA, Kittelsen SAC. 1996. Slack-adjusted efficiency measures and ranking of efficient units. *Journal of Productivity Analysis* 7: 379-398.

Tyagi P, Yadav SP, Singh SP. 2009. Relative performance of academic departments using DEA with sensitivity analysis. *Evaluation and Program Planning* 32: 168-177.

Waldo S. 2007. On the use of student data in efficiency analysis – Technical efficiency in Swedish upper secondary school. *Economics of Education Review* 26: 173-185.