I N S T I T U T D E S T A T I S T I Q U E

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

D I S C U S S I O N P A P E R

0426

ADVANCED INDICATORS OF PRODUCTIVITY OF UNIVERSITIES AN APPLICATION OF ROBUST NONPARAMETRIC METHODS TO ITALIAN DATA

BONACCORSI, A., DARAIO, C. and L. SIMAR

http://www.stat.ucl.ac.be

Advanced Indicators of Productivity of Universities. *An application of robust nonparametric methods to Italian data*[*]

Andrea Bonaccorsi
University of Pisa, Italy
bonaccorsi@sssup.it

Cinzia Daraio[†]
IIT-CNR and Scuola Superiore S. Anna, Italy
cinzia@sssup.it

Léopold Simar[‡]
Université Catholique de Louvain, Belgium
simar@stat.ucl.ac.be

October 6, 2004

Abstract

In recent years the pressure on public budgets in almost all industrialised countries has lead governments to pursue efficiency in the allocation and management of resources trying to apply to scientific research and higher education some fundamental ideas of the economic analysis such as the concepts of economies of scale and scope. This paper explores scale, scope and trade-off effects in scientific research and education. Advanced productivity methods are used to analyse the Italian system of universities. In particular, robust methods based on the concept of order$-m$ frontiers (Cazals, Florens and Simar, 2002) are really useful in this framework for their properties of not being influenced by extremes and noise in the data. Furthermore, in the field of

[*]This paper has been prepared within the AQuaMethPSR Project under the PRIME Network of Excellence supported by the European Commission, 6th Framework Programme. Previous versions of this paper have been presented to the 8^{th} S&T Indicators Conference in Leiden, 23-25 September 2004, and at the North American Productivity Workshop in Toronto, 23-25 June 2004. We would like to thank conferences' participants for their helpful comments. We would also like to thank Alessandro Daraio for his assistance in building the databases.

[†]Work partially supported by the Italian Registry of ccTLD .it.

[‡]Research support from the "Interuniversity Attraction Pole", Phase V (No. P5/24) from the Belgian Federal Government (Belgian Science Policy) is also acknowledged.

science and education, external factors and environmental conditions may be cause of heterogeneity and influence dramatically the performance of universities. Hence, we apply the Daraio and Simar (2003) full nonparametric methodology (that overcomes most limitations of previous one or two-stage approaches) to robustly take into account external-environmental factors. From a preliminary investigation on Italian data we find that economies of scale and scope are not significant factors in explaining research and education productivity. We do not find any evidence of the trade-off research vs teaching: on the contrary, increasing scientific quality improves educational efficiency; on the other hand, a good educational efficiency does not deteriorate research efficiency. About the trade-off academic publications vs industry oriented research, *local effects* of a complementarity/rivalry relation seem to emerge: it seems that initially, collaboration with industry may improve productivity, but beyond a certain level the compliance with industry expectations may be too demanding and deteriorate the publication profile. Nevertheless, the existence of an inverted $U-$shaped relation should be confirmed by more evidence. Advanced robust methods in efficiency analysis are shown as useful tools for measuring and explaining the performance of a public research system of universities. Further developments of the analysis are outlined.

Keywords: nonparametric frontier, robust estimation, economics of education, economics of science, university, scale, scope, specialisation, public sector research.

JEL Classification: C14, C50, I21, I28, I29

1 Introduction

In recent years the pressure on public budgets in almost all industrialised countries has lead governments to pursue (or at least to declare they pursue) efficiency in the allocation and management of resources in the public research sector. The increasing societal demand for accountability and transparency of science and research also makes it important to demonstrate that public funding follows clear rules. A manifestation of this trend is the effort to apply to public scientific research two very fundamental concepts drawn from economic analysis, that are, increasing returns to scale or economies of scale, and economies of scope. At the level of universities, size and scope may be reflected in appropriate measures of inputs or outputs (*e.g.* staff or enrollments or graduate students, number of schools, number of different curricula) in a range between specialist universities, covering a few fields, to generalist universities, covering almost the entire spectrum of disciplines. If these two forces were at play in scientific research and higher education, then a sound policy implication would be that in order to improve the public efficiency, resources should be concentrated into larger institutions and/or more generalist universities. On the other hand, there may be negative scope effects, implying that different activities are *rival* rather than being complementary. This paper explores scale, scope and trade-off effects in scientific research and higher education with reference to the Italian system of universities.

In this paper we apply advanced nonparametric techniques for estimating the technical efficiency of Italian universities.

In particular, *robust* methods are really useful in this framework for their properties of not being influenced by extreme values and outliers in the data. Several methodological advancements have been done in nonparametric frontier estimation with robust methods based on the concept of order-m frontiers (Cazals, Florens and Simar, 2002). These recent methods have been applied for the first time to scientific research in Bonaccorsi and Daraio (2003), who offer a comparison between two large institutions in biomedical research in two countries.

However, in the field of science and education, external-environmental variables may be cause of heterogeneity and may considerably influence the performance of universities. Several studies in the education literature have tried to face this problem by developing and applying one-stage or two or multiple-stage approaches to take into account what they define as *socio-economic differences*.[1]

The basic idea was to relate efficiency measures to some external or environmental factors which might influence the production process but that are not under the control of the managers.

Unfortunately, both one stage and multiple stage approaches are flawed by restrictive prior assumptions and/or on the role of these external factors on the analysed process. From one hand, as discussed and demonstrated by Simar and Wilson (2003), the multiple-stage approaches suffer of several methodological problems related to the complicated and unknown autocorrelations between first and second stages estimation procedures but also to the inherent bias of the first stage efficiency estimates. On the other hand, in the one stage approach one has to assume the effect of the external factors on the production process, *i.e.* the analyst should know in advance if the external factors affect positively or negatively the comprehensive performance. Of course, these problems and assumptions are very tight.

For this reason, Daraio and Simar (2003) generalizing the approach of Cazals, Florens and Simar (2002) propose a full nonparametric methodology to explain efficiency differentials by external-environmental factors that overcomes most limitations of previous approaches. In particular, the effect of external factors came as a result of the analysis and is not assumed.

Advanced robust methods in efficiency analysis are shown as useful tools for measuring and explaining the performance differentials of a public research system of universities. As

[1] See Ruggiero (2004) and the references cited there.

applied to our data, they shed lights on two main phenomena:

(a) the existence of scale and scope economies in the Italian system of universities;

(b) the analysis of the trade-off between teaching vs publication activities and between academic research (publication) vs industrial applied research.

Policy implications of the empirical evidence are largely discussed.

The paper is organised as follows. The next two sections review the relevant literature on scale and scope economies and trade-off effects. Section 4 describes the methodology applied and highlights the advantages of its utilisation as compared with traditional nonparametric techniques of DEA/FDH type. Section 5 defines the variables used in the analysis defining inputs, outputs and environmental-external factors used in the empirical investigations. Some descriptive statistics are also reported. In section 6 the main results of the analysis are reported, while section 7 summarizes and concludes the paper, outlining directions for further research.

2 Economies of scale and scope in universities

The literature in economics of education has long debated whether higher education institutions, such as universities, benefit from economies of scale and scope. A related problem has been discussed in the economics of research, taking into consideration specialized research institutions such as Public Research Organisations (PROs) and universities. Clearly, in the case of universities both aspects are relevant and the interplay between returns to scale and scope in the two fields of education and research is the most intriguing aspect.

Recall that in the context of manufacturing production, economies of scale refer to the fact that an increase of k times in all factors of production determines an increase in output of more than k times. Therefore the higher the scale of production (*i.e.* productive capacity of plants), the lower the unit or average cost in the long run. To claim that increasing returns to scale are at play one must increase simultaneously all factors of production, not only the variable ones (*i.e.* work). Economies of scope, on the other hand, refer to the reduction in average cost associated to the possibility to produce two or more qualitatively different outputs using the same structure of inputs. In the field of research and education, economies of scope refer to the joint production of undergraduate and postgraduate education, and to the use of common inputs for the production of different degrees (*e.g.* teaching calculus to students of physics and engineering). In the case of universities, economies of scope between teaching and research and between different types of research should also be considered.

In higher education and research it is not easy to identify the relevant unit of analysis. Research activity is carried out in laboratories and teams usually organised in institutes or departments with a disciplinary connotation. Higher education activities are carried out in schools or faculties partially defined by disciplinary boundaries but in general much broader in scope. Thus in general a school purchases the services of teachers that do research in different departments, and a department sells teaching services to several schools. Since no price system is in place between the two entities, it is difficult to allocate resources to the units of observation. As a matter of fact, economics of research (*e.g.*, Ramsden, 1994; Johnston, 1994; Adams and Griliches 2000) focuses mainly on departments as unit of observation, while economics of education considers schools or universities (*e.g.* Cohn, Rhine and Santos, 1989). The only level of observation that allows to explore the relation between the two fields is the university level. In general, the activity of researchers and teachers is reasonably well represented at the university level.

The arguments for assuming increasing returns of scale and scope are several.

- *Teaching process*

  Up to a certain point, the cost to teach to more students grows less than proportionally with respect to the number of students. Large universities may allocate resources more efficiently if they can fill classrooms. It is likely that this effect has an upper bound given by congestion phenomena. On the other hand, if teaching involves significant interpersonal relation, the maximum number of students per teacher does not depend on the size of the university, so that costs may increase proportionally with the number of students.

- *Indivisibilities in human capital for research*

  In many areas the production of scientifically meaningful output requires the combination and coordination of many scientists from different fields, bringing competencies in the substantive field and in complementary areas such as measurement techniques, statistical analysis, scientific computing, software development, data and image processing and analysis and the like. In some areas the very substantive field requires the integration of different disciplinary backgrounds. As a result the notion of minimum size of a research unit is economically sensible. However, we should be careful in defining the level of observation. First of all, indivisibility is more important at the level of team or laboratory than at the level of institute or department. Second, while the notion of indivisibility is clear in abstract terms, its empirical relevance may be highly variable. In other words, the minimum size of a team or laboratory may be extremely variable across specific areas within the same fields. In general, this means that economies of scale may be important up to a threshold level, then become irrelevant. If the threshold level is quite small (in the order of a few units or a few dozens), the practical implication is that even small institutes may be highly efficient, provided that their teams or labs meet the minimum requirement.

- *Educational and research infrastructure*

  Educational activities require fixed assets such as libraries and computer rooms, while research activities often require expensive equipment and instrumentation. To the extent that these resources are indivisible, efficient units must be large enough to ensure full utilization of assets. Furthermore, it may be assumed that students are also sensitive to the size of universities because they assume larger universities to be relatively more equipped and well-serviced. Thus in order to attract students one should have sufficient visibility.

- *Indivisibilities in administration*

  This is a classical argument from minimum efficient scale in administrative and clerical staff. Small research institutes and universities may underutilize their fixed amount of personnel. With the increase of Information and Communication Technology (ICT) use in administration, it is no clear how this argument may be important. In addition, although minimum scale is clearly conceivable for administration, this is most probably at a relatively low level, beyond which returns are constant. In addition, one should also consider the possibility of congestion of large bureaucratic institutions as a source of severe diseconomies.

2.1 Previous empirical evidence

The evidence on returns to scale in scientific production and in higher education are ambiguous. In fact, there is lack of consensus on the existence of economies of scale in the production of research and university teaching. Amongst many others, Brinkman (1981), Brinkman and Leslie (1986), Cohn, Rhine and Santos (1989), de Groot, McMahon and Volkwein (1991), Nelson and Hevert (1992) and Lloyd, Morgan and Williams (1993) report the existence of economies of scale. Verry and Layard (1975), Verry and Davies (1976) and Adams and Griliches (2000), on the contrary, find constant returns to scale, while Narin and Hamilton (1996, p.297) review several studies and conclude “we have never found that the size of an institution is of any significance”. By concluding the review commissioned by the UK Office of Science and Technology, von Tunzelmann, Ranga, Martin and Geuna (2003) state: “there seems to be little if any convincing evidence to justify a government policy explicitly aimed at further concentration of research resources on large departments or large universities in the UK on the grounds of superior economic efficiency”. On the other hand, other studies find increasing returns to scale until a threshold level, after which constant or even decreasing returns describe better the situation. As Johnston (1994) summarises the literature, “the results of this body of work can best be characterized as ambiguous and contradictory. The majority verdict is that research output is linearly related to size with no significant economies of scale apparent. Others have argued that the relationship between output and size is more complicated - for example, that there are economies of scale up to a certain group size after which diseconomies set in” (Johnston, 1994, p.32). Despite ambiguity in empirical evidence, the notion of economies of scale and scope is often invoked to support policies of concentration of resources in larger universities or institutes, forcing small institutes to merge or disappear, or policies of merger and consolidation of scientific institutions and universities. For example, Abbott and Doucouliagos (2003) report that the Australian government, in the attempt to improve the efficiency of the university system by exploiting economies of scale and scope, consolidated a large number of higher education institutions into a small number of large multi-campus universities. The keyword for these policies is *critical mass*: the underlying idea is that universities must stay big in order to mobilize sufficient managerial resources to do fund raising and appropriate size and diversity of research teams to ensure research competitiveness.[2] A classical example is the attempt to merge Imperial College and University College in London by a vigorous team of top decision makers, some of which came from industry.

Concentration is a robust structural property of institutional systems that allocate research funds in proportion to publishing output. Since publication activity follows a strongly asymmetric distribution, it is no surprise that research funds are not allocated on a uniform basis.[3] However, public policies based on critical mass and large institutes induce levels of

[2]As it has been noted “a prominent feature of research support policy in many, though not all countries, over the last twenty years has been the espousal and implementation of resource allocation processes that provide ‘selectivity and concentration’. Implicit in these policies has been the assumption that ’bigger is better’; in other words, that scientific research benefits from economies of scale. This approach has been most pronounced in the UK and to some extent other Anglo-Saxon derivative countries, but it has been the subject of consideration and experiment in many other countries as well” (Johnston, 1994, p.25-26).

[3]In a sample of Australian researchers in 18 universities, Ramsden (1994) found that 14% of researchers was responsible for 50% of all publications in the 1985-89 period, while 40% published 80% of the total. Approximately the same ratio was found by Cole et al. (1981) and Reskin (1977) for US universities (15% of researchers published 50% of the total), while Halsey (1980) found comparable concentration ratios for British universities and polytechnics (23% of researchers published 68% of the total). As a consequence, a small number of universities that follow a consistent policy of hiring scientists with a strong publication record absorb a large share of funds: in the US more than 50% of national budget for universities is concentrated in the top 33 universities. In UK the top 6 universities absorb around 50% of the total. Policies aimed at concentration do not simply follow the structural asymmetry of distribution of publication activity, but aim to actively improve productivity.

concentration of resources that go beyond the usual level.

In examining evidence from the National Research Council and using also robust nonparametric techniques, Bonaccorsi and Daraio (2004a; 2004b) confirm the notion that returns to scale may be subject to *local* effects. This means that research activity may be subject to increasing returns up to a certain size, then to constant or even decreasing. The strong policy implication of these findings is that concentration may be counterproductive if it brings institutes to grow up to the region where constant or decreasing returns are at work. In fact, estimating economies of scale over the entire range of observations, as is standard in the production function approach and with regression techniques, will result in averaging a number of very different local size effects. The policy implication of finding, for example, economies of scale will be consolidating universities or merging research units. But if size effects are local the policy may even worsen the situation. Suppose there are several regions of returns to scale, initially increasing then constant or decreasing. Merging units means that smaller institutes, which initially benefited from economies of scale, will become larger and will enter into a region where these effects are eliminated. On the contrary, in the robust nonparametric frontier approach (based on order$-m$ frontiers, for details see section 4) it is possible to estimate separately the efficient frontier returns to scale, the global effect of scale, and the individual position with respect to returns to scale. It is possible that returns to scale are variable over a limited interval, whilst they are constant over other intervals of the observed size distribution.

3 Trade-offs

We also address a classical topic in the economics of education and research. It is well known that university people are left (relatively) free to allocate their time budget across various activities. At the same time the teaching burden may be largely variable across universities and time. The relevant question is whether competing activities are substitute or complementary activities. This question applies to two different problems:

- the trade-off between research and teaching
- the trade-off between research for publication and research for industrial use or patenting.

These are clearly critical issues for the overall organisation of universities and for policy-making.

Let us first discuss the trade-off between research and teaching.

At the individual researcher level, several arguments may be put forward in opposite directions. Teaching activities not only refers to classroom teaching but also involves tutoring of students, revision of homework, supervision of theses. If these activities refer to undergraduate students, they do not contribute very much to research record. On the other hand, working with students gives the opportunity to identify at an early stage potentially promising researchers. Exposing ideas to students is usually a challenging activity, that contributes to research creativity and discipline. Also, in some countries (as Italy) the preparation of undergraduate dissertations involves heavy work that in some cases may be useful for research purposes. Therefore the net effect is rather indeterminate.

At the university level, the net effect is even more complex to understand. On one hand, the larger the number of students per professor, the smaller the time budget to be allocated to research. Thus large and overcrowded universities are expected to discourage professors from carrying out good research. At the same time, however, most institutional systems give funds to universities in proportion to the number of students. This is not only true in private systems such as the American one, but also in the great majority of public systems, where the government usually gives money to universities using enrolled students and/or degrees as a proportionality parameter. If this is true, having more students means having more resources, part of which my be allocated to research. Large universities may more easily leverage financial resources through student fees or governmental transfers and engage in high profile investment activities that benefit research, such as building facilities and creating laboratories. For example, using an incentive-based model, Gautier and Wauthy (2004) show that if multi-department university may redistribute resources, they may induce better teaching quality and research.

Again, the net effect is difficult to determine ex ante. For example, Graves, Marchand and Thompson (1982) found a negative impact of teaching on research in economics departments, with an elasticity coefficient between .25 and 1.66, according to specifications. On the contrary, Cohn, Rhine and Santos (1989) found positive teaching/research complementarity, supporting the notion of economies of scope between the two activities.

A more recent trade-off has been explored in the literature on the so called third mission of universities and public research organisations. Put it briefly, the public research system has been pushed in many Western countries to engage in proactive exchanges with industry and to actively contribute to economic growth, mainly at national and local level. In addition to teaching and research, universities have been invited to assume another argument in their objective function, namely the transfer of results from science to industry, and to implement it through the creation of intellectual property rights over the results of research, collaborative contractual relations with industry, joint consortia and research companies with industry, or creation of spin-off companies (Martin, 2001; Thursby and Kemp, 2002). This trend has motivated deep concerns along two dimensions.

The first is the process of defining objectives of research and performance criteria. Assuming the goal of serving better the needs of industry and reducing the time lag between discovery and commercial application may induce short-termism and deteriorate the quality of research (Henderson, Jaffe and Trajtenberg, 1998; Hicks and Hamilton, 1999; Mowery and Ziedonis, 2002; Gittelman and Kogut, 2003).

On the other hand, the commercialisation of science inevitably implies the adoption of communication practices that are typical of the industry environment, including industrial secrets, restrictions to disclosure, and confidentiality. This is in sharp contrast with the practice and ethos of public science. Several authors have warned against possible risks of adopting commercial criteria in the Republic of science (Campbell and Blumenthal, 2000; David, 2000; Mowery and Sampat, 2001; Nelson, 2004).

Some recent results on the Italian case add evidence to this debate. Balconi, Breschi and Lissoni (2004) have investigated the patenting behaviour of Italian academicians and identified a large pool of inventive activity that is not registered officially as university patenting. Comparing academic inventors to colleagues of the same discipline and seniority that did not patent, Breschi, Lissoni and Montobbio (2004) find that academic inventors are much more productive. Along similar lines, Calderini and Franzoni (2004) study the patenting and publishing behaviour of 1323 researchers at universities and PROs in the fields of Engineering Chemistry and Nanotechnologies for New Materials along 30 years. Again, they find that the event of patenting not only takes place after a period of high quality scientific publications but also is followed by a similar period, suggesting positive association.

While the debate cannot be solved without additional evidence across several countries, disciplines, and institutional settings, we put forward a conjecture. We suggest that the complementarity/rivalry relation is subject to local, non-linear effects along the distribution of relevant variables. In particular, we propose that the impact of industry involvement on pure scientific productivity may follow an inverted U-shaped relation with respect to the extent of involvement. Initially, collaboration with industry in various forms may improve productivity as measured by international publications, but beyond a certain level the compliance with industry expectations may be too demanding and deteriorate the publication profile.

4 The methodology

4.1 Introduction

In efficiency analysis the main purpose is the study of how firms combine their inputs to obtain their outputs. More generally, in an activity analysis framework (see *e.g.* Debreu, 1951; Shephard, 1970), the management of a Decision Making Unit (DMU) is characterized by a set of inputs $x \in I\!R^p_+$ used to produce a set of outputs $y \in I\!R^q_+$. The set of technically feasible combinations of (x, y) is defined as:

$$\Psi = \{(x, y) \in I\!R^{p+q}_+ \mid x \text{ can produce } y\}. \tag{4.1}$$

In this setting, the Farrell measure of output-oriented[4] efficiency for a firm operating at the level (x, y) can be defined as:

$$\lambda(x, y) = \sup\{\lambda \mid (x, \lambda y) \in \Psi\}, \tag{4.2}$$

where $\lambda(x, y) \geq 1$ is the proportionate increase of outputs a DMU working at the level (x, y) should perform to achieve efficiency. The efficient frontier corresponds to those firms where $\lambda(x, y) = 1$.

See Figure 1 for a graphical representation of the production set and the situation of a DMU A which produces an output y_A using x_A input. For sake of representation we illustrate a simple univariate frontier, even if an advantage of nonparametric methods is their multi-input multi-output description of production technology.

Of course the *true* efficient frontier is not known and has to be estimated using a sample of production observations.

[4] To save place, in this paper we only present the output oriented case that we apply in the empirical analysis.

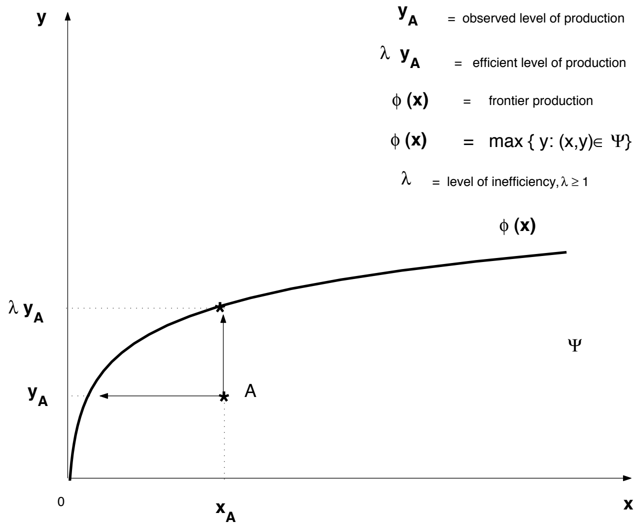

Figure 1: *Production set and efficient frontier: an illustration.*

In efficiency analysis, the nonparametric approach is based on envelopment techniques, whose main estimators are Data Envelopment Analysis (DEA, see Farrell, 1957, and Charnes, Cooper and Rhodes, 1978) and Free Disposal Hull (FDH, see Deprins, Simar and Tulkens, 1984). These estimators rely on the idea that the attainable set is defined by the set of minimum volume containing all the observations. The DEA estimator relies on the free disposability[5] and on the convexity of the set Ψ, whereas the FDH relies only on the free disposability assumption.

The FDH estimator of Ψ, based on a sample of n observations (x_i, y_i), is the free disposal closure of the reference set $\{(x_i, y_i)|i = 1, ..., n\}$. It can be defined as:

$$\widehat{\Psi}_{FDH} = \Big\{(x, y) \in I\!R_+^{p+q} \mid y \leq y_i, x \geq x_i, i = 1, ..., n\Big\}. \tag{4.3}$$

The DEA estimator of Ψ, is the convex closure of $\widehat{\Psi}_{FDH}$:

$$\begin{aligned} \widehat{\Psi}_{DEA} \;=\; & \Big\{(x, y) \in I\!R_+^{p+q} \mid y \leq \sum_{i=1}^{n} \gamma_i\, y_i \;;\; x \geq \sum_{i=1}^{n} \gamma_i\, x_i, \\ & \quad \text{for } (\gamma_1, ..., \gamma_n) \quad \text{s.t.} \sum_{i=1}^{n} \gamma_i = 1; \gamma_i \geq 0, i = 1, ..., n\Big\}, \end{aligned} \tag{4.4}$$

The estimated output oriented FDH efficiency score of a firm (x, y) is given by:

$$\widehat{\lambda}_{FDH} = \max\{\lambda \mid (x, \lambda y) \in \widehat{\Psi}_{FDH}\}. \tag{4.5}$$

[5] A set Ψ is free disposal if $(x, y) \in \Psi$ implies $(x', y') \in \Psi$ for any $x' \geq x$ and $y' \leq y$.

Similarly, the estimated output oriented DEA efficiency score of a DMU (x, y) is given by:

$$\widehat{\lambda}_{DEA} = \max\{\lambda \mid (x, \lambda y) \in \widehat{\Psi}_{DEA}\}. \tag{4.6}$$

The nonparametric approach in efficiency analysis offers several advantages, among whose:

- absence of specification of the functional form for the input-output relationship;
- measurement of the efficiency with respect to the efficient frontier which measures the best performance that can be practically achieved;
- appropriate benchmark to be used for comparison: non requirement of any theoretical models as benchmarks;
- production of *multi-inputs multi-outputs* performance indicators.

4.2 Robust order-m frontiers

On the other hand, one of the main drawbacks of DEA/FDH nonparametric estimators is their sensibility to extreme values and outliers in the data.

It is interesting to note that mostly for the problem of outliers the DEA methodology was abandoned by the Italian Conference of Rectors (CRUI) few years ago.

To overcome this methodological limitation, Cazals, Florens and Simar (2002) propose a nonparametric estimator of the frontier, more robust to extreme values and outliers. It is based on the concept of the *expected maximum output function* of order$-m$.

See Figure 2 for an illustration of an order$-m$ frontier in a univariate setting.

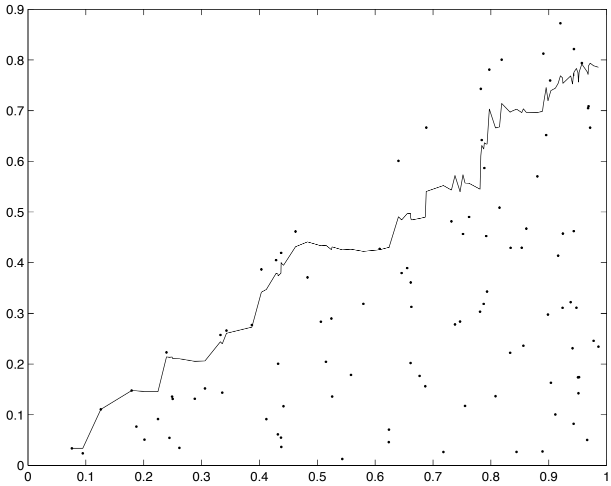

Figure 2: *An illustration of a robust order−m frontier.*

Extending these ideas to the full multivariate case, Daraio and Simar (2003) define the concept of *expected* order-m output efficiency score. This robust approach is based on a probabilistic formulation of the model. The production process is described by the joint probability measure of (X, Y) on $\mathbb{R}_+^p \times \mathbb{R}_+^q$. In this formulation, the support of (X, Y) is the attainable set Ψ and the Farrell output efficiency[6] can be characterized, under the free disposability assumption, as:

$$\lambda(x, y) = \sup\{\lambda \mid S_Y(\lambda y \mid x) > 0\}, \tag{4.7}$$

where $S_Y(y \mid x) = \text{Prob}(Y \geq y \mid X \leq x)$. A nonparametric estimator of $\lambda(x, y)$ is provided by pluging in (4.7) the empirical version of $S_Y(y \mid x)$:

$$\widehat{S}_{Y,n}(y \mid x) = \frac{\sum_{i=1}^{n} \mathbb{1}(X_i \leq x, Y_i \geq y)}{\sum_{i=1}^{n} \mathbb{1}(X_i \leq x)}. \tag{4.8}$$

It has been shown that the resulting nonparametric estimator coincides with the FDH estimator defined in (4.5).

The order-m output efficiency can be defined as in Daraio and Simar (2003): for a given level of inputs x in the interior of the support of X, consider m i.i.d. random variables

[6]Here again, we only describe the output-oriented case, for the input oriented case, for more technical details and illustrations, see Daraio and Simar (2003) and Daraio (2003). For a description of the nonparametric estimators and computational aspects see Daraio and Simar (2003).

$Y_i, i = 1, \ldots, m$ generated by the conditional q-variate distribution function $F_Y(y \mid x) = \text{Prob}(Y \leq y \mid X \leq x)$ and define the set:

$$\Psi_m(x) = \{(x', y) \in I\!R_+^{p+q} \mid x' \leq x, Y_i \leq y, i = 1, \ldots, m\}. \tag{4.9}$$

Then, for any y, we may define:

$$\tilde{\lambda}_m(x, y) = \sup\{\lambda \mid (x, \lambda y) \in \Psi_m(x)\} \tag{4.10}$$

For any $y \in I\!R_+^q$, the (expected) order-m output efficiency measure denoted by $\lambda_m(x, y)$ is defined for all x in the interior of the support of X as:

$$\begin{aligned}\lambda_m(x, y) &= E(\tilde{\lambda}_m(x, y) \mid X \leq x) &(4.11)\\ &= \int_0^\infty [1 - (1 - S_Y(uy \mid x))^m]\, du &(4.12)\end{aligned}$$

A nonparametric estimator of $\lambda_m(x, y)$ is given by:

$$\hat{\lambda}_m(x, y) = \int_0^\infty \left[1 - (1 - \widehat{S}_{Y,n}(uy \mid x))^m\right] du. \tag{4.13}$$

Hence, in place of looking for the upper boundary of the support of $S_Y(y \mid x)$, as was typically the case for the full-frontier and for the efficiency score $\lambda(x, y)$, the order-m efficiency score can be viewed as the expectation of the maximal output efficiency score of the unit (x, y), when compared to m units randomly drawn from the population of units producing using less inputs than the level x. This is certainly a less extreme benchmark for the unit (x, y) than the "absolute" maximal achievable level of outputs: it is compared to a set of m peers (potential competitors) producing using less than its level x of inputs and we take as benchmark, the *expectation* of the maximal achievable output instead of the absolute maximal achievable output.

Then, for any $y \in I\!R_+^q$, the expected maximum level of outputs of order-m is defined as $y_m^\partial(x) = \lambda_m(x, y)\, y$ which can be compared with the full-frontier $y^\partial(x) = \lambda(x, y)\, y$.

The robust nonparametric methodology we apply in this paper (based on order-m efficiency) adds some new advantages to the traditional nonparametric approach (DEA/FDH).

- As the robust indicators are based on estimators that do not envelop all firms, they are more robust to outliers and noise in the data which may strongly influence the nonparametric estimation of efficiency. The level of robustness can be set by means of m (tuning parameter). The level of robustness is determined by the percentage of points remaining above the order$-m$ frontier. Clearly, when $m \to \infty$ this percentage goes to zero.

- The robust nonparametric indicators avoid the *curse* of dimensionality, typical of nonparametric estimators. The order-m indicators are $\sqrt{n}$-consistent estimators whereas the DEA are only $n^{2/(p+q+1)}$-consistent estimators ($n^{1/(p+q)}$ for the FDH). This indicates for the DEA/FDH the necessity of increasing the number of observations when the dimension of the input-output space increases to achieve the same level of statistical precision;

- The order-m indicators allow to compare samples with different size, in an indirect way, avoiding the *sample size bias*[7], of which nonparametric indicators (DEA/FDH) suffer. In this case, m plays an important role. The benchmark, in fact, is not made against the most efficient units in the group, but against an appropriate measure drawn from a large number of random samples of size m within the group. In this way size-dependent effects are eliminated.

- The possibility of explaining efficiency, considering the *conditional influence* of *external factors* Z on the full frontier and on its robust counterpart (see the next section).

- Using this approach, it is possible to decompose the performance of a firm (x, y), as measured by the *Conditional Efficiency* index in three main indicators: an indicator of the *internal* or managerial efficiency, an *externality index*, and finally, an *individual* index.

- Moreover, we can evaluate the effect of external/environmental Z variables on the performance of firms in different economic *scenarios*, contemplating various numbers of *potential competitors*, using the parameter m in its dual meaning.

- Parametrization and robust elasticities are available by applying the two-step estimators (based on FDH and order$-m$ frontiers) introduced by Florens and Simar (2004).

4.3 Conditional measures of efficiency

As we anticipate in the introduction, external or environmental conditions may strongly influence the efficiency evaluation of research and educational systems.

In the efficiency literature, mainly three approaches have been developed[8]: a *one-stage* approach, a *two-stage* approach and a bootstrap-based approach. Nevertheless, all of them

[7]Zhang and Bartels (1998) show formally how DEA efficiency scores are affected by sample size. They demonstrate that comparing measures of structural (*i.e.* average) inefficiency between samples of different sizes leads to biased results.

[8]See Daraio and Simar (2003) and the references cited there.

are flawed by restrictive assumptions on the data generating process and/or on the role of these *external* factors on the production process.

Based on the probabilistic formulation presented above, Daraio and Simar (2003) propose a general full nonparametric approach that overcomes most drawbacks of previous approaches. The probabilistic formulation allows an easy introduction of additional information provided by external- environmental variables $Z \in \mathbb{R}^r$. Hence, the joint distribution on (X, Y) conditional on $Z = z$ defines the production process if the external factor $Z = z$.

The output efficiency measure under the condition $Z = z$ can be defined as:

$$\lambda(x, y \mid z) = \sup\{\lambda \mid S_Y(\lambda y \mid x, z) > 0\}, \tag{4.14}$$

where $S_Y(y \mid x, z) = \text{Prob}(Y \geq y \mid X \leq x, Z = z)$. A nonparametric estimator of $\lambda(x, y \mid z)$ that requires some smoothing in z, is provided by the following kernel estimator of the empirical version of $S_Y(y \mid x, z)$:

$$\widehat{S}_{Y,n}(y \mid x, z) = \frac{\sum_{i=1}^{n} 1\!\!I(X_i \leq x, Y_i \geq y) K((z - z_i)/h_n)}{\sum_{i=1}^{n} 1\!\!I(X_i \leq x) K((z - z_i)/h_n)}. \tag{4.15}$$

where $K(\cdot)$ is the kernel and h_n is the bandwidth of appropriate size. Daraio and Simar (2003) propose a simple data-driven method for the choice of the bandwidth.[9]

In a similar way, *mutatis mutandis*, Daraio and Simar (2003) introduce also the conditional order$-m$ measures of efficiency with their nonparametric estimators. For the output-oriented case, the order$-m$ measure of efficiency is defined as:

$$\lambda_m(x, y|z) = \int_0^{\infty} [1 - (1 - S_Y(uy \mid x, z))^m] \, du \tag{4.16}$$

and its nonparametric estimator is obtained as follows:

$$\hat{\lambda}_{m,n}(x, y|z) = \int_0^{\infty} \left[1 - (1 - \widehat{S}_{Y,n}(uy \mid x, z))^m\right] du \tag{4.17}$$

When $m \to \infty$, we recover the full frontier conditional measures, but for finite m, $\hat{\lambda}_{m,n}(x, y|z)$ provides a more robust estimator of the frontier, robust to extremes or outliers.

[9] In a *first step*, a bandwidth h is selected which optimizes the estimation of the density of Z based on the likelihood cross validation criterion, using a k-NN (Nearest Neighborhood) method. This allows to obtain bandwidths which are localized, insuring we have always the same number of observations Z_i in the local neighbor of the point of interest z when estimating the density of Z. In a *second step*, taking into account for the *dimensionality* of x and y, and the sparsity of points in larger dimensional spaces. the local bandwidths h_{Z_i} are expanded by a factor $1 + n^{-1/(p+q)}$, increasing with $(p + q)$ but decreasing with n. For more details, see Daraio (2003).

The procedure for evaluating the effect of Z on the production process is based on the comparison of the conditional FDH measure $\hat{\lambda}_n(x, y \mid z)$ with the unconditional FDH measure $\hat{\lambda}_n(x, y)$. Accordingly, the same comparison is done for the robust order$-m$ efficiency measures. In particular, the ratios $Q^z = \hat{\lambda}_n(x, y \mid z)/\hat{\lambda}_n(x, y)$ (and their robust version Q^z_m) are useful to investigate on the effects of Z on performance: if $Q^z = 1$, then the conditional and unconditional efficiency measures are equal: this means that Z does not affect the performance of the analysed firm; if Q^z is much lower than 1, this means that the firm has been *highly* influenced by Z.

When Z is univariate, the scatterplot of these ratios against Z and its smoothed nonparametric regression line is also very helpful in describing the effect of these external-environmental variables on the production process. In the output oriented case (used in this paper), we have:

- An *increasing* smoothed nonparametric regression line denotes a Z that is *favorable* to the production process. In this framework, a favorable Z means that the environmental variable operates as a sort of “extra” input *freely available*: for this reason the environment is “favorable” to the production process. Consequently, the value of $\hat{\lambda}_n(x, y \mid z)$ will be much smaller (greater efficiency) than $\hat{\lambda}_n(x, y)$ for small values of Z than for large values of Z: the ratios $\hat{\lambda}_n(x, y \mid z)/\hat{\lambda}_n(x, y)$ will increase with Z, on average.

- A *decreasing* smoothed nonparametric regression line indicates a Z that is *unfavorable* to the production process. In this case, the environmental variable works as a “compulsory” or *unavoidable* output to be produced to face the negative environmental condition. Z in a certain sense penalizes the production of the outputs of interest. In this situation, $\hat{\lambda}_n(x, y|z)$ will be much smaller than $\hat{\lambda}_n(x, y)$ for large values of Z. As a result, the regression line of $\hat{\lambda}_n(x, y \mid z)/\hat{\lambda}_n(x, y)$ over Z will be decreasing.

Here we point out the clear interpretability of these kind of scatterplots. As we will see later in the empirical results section, the analyst has an immediate view on the global effect of external factors on the performance: an *increasing* line indicates a positive influence of the factor, a *decreasing* line points to a negative effect and a *straight* line reveals no influence of the factor on the performance.

4.4 A decomposition of conditional efficiency

This nonparametric approach, with its robust counterpart, offers a rigorous methodology to identify the factors that might influence firms efficiency by measuring their *global* effect on the performance. It gives also the possibility to analyse the effect of these variables on each individual firm. The performance of a DMU (x, y), measured by the *Conditional Efficiency* index (CE), $\widehat{\lambda}_n(x, y \mid z)$, can be decomposed in three main indicators (see Daraio (2003) for more details):

1. An *unconditional* efficiency score (UE), $\widehat{\lambda}(x, y)$ that represents the *internal or managerial efficiency*;

2. An *externality index* (EI) defined as $\widehat{E}(Q^z|Z = z)$. It is the *expected value* of the ratios Q^z given the value of z owned by the firm. It is given by the nonparametric fitted value of Q^z obtained by some appropriate nonparametric regression of Q^z on Z:

$$\widehat{E}(Q^z|Z = z) = \frac{\sum_{i=1}^{n} Q^{z_i} K((z - z_i)/h)}{\sum_{i=1}^{n} K((z - z_i)/h)}, \tag{4.18}$$

   where $K(\cdot)$ is the Kernel and h an appropriate bandwidth.

3. An *individual index* (II) defined as $Q^z/\widehat{E}(Q^z|Z = z)$. It compares the value of Q^z with the value we would expect for it, given its value of Z. It represents the firm's expected *intensity* in catching the opportunities or threats by the environment (external factor).

In formulae:

$$CE(x, y) = UE(x, y) * EI(x, y) * II(x, y); \tag{4.19}$$

equivalently, for the output oriented case:

$$\widehat{\lambda}_n(x, y|z) = \widehat{\lambda}_n(x, y) * \widehat{E}(Q^z|Z = z) * \frac{Q^z}{\widehat{E}(Q^z|Z = z)}. \tag{4.20}$$

Accordingly, the same decomposition can be done in the case of conditional order$-m$ efficiency score.

This decomposition is particularly useful in the interpretation of the firms' performance. It offers the possibility for analyzing *individual* and *localized* effects of external factors and interpret them together with their global influence on the firms production activity.

5 Data description

In this paper we analyse data coming from two datasets collected by the Italian conference of university rectors (CRUI). The first one collects data for the academic years 1995/96 to 1997/98 and reports three sets of variables:

- Financial variables: financial resources and the structure of expenses at the university level.
- Human resource variables (teaching staff, administrative and technical staff).
- Number of students enrolled and graduated.

The second database contains data on the Italian universities' research outputs; it has been constructed using the Information Science Institute (ISI) data and contains the cumulative number of publications and citations obtained in the period 1995-1999, by university. In particular, it contains detailed information for the following scientific areas:

- Mathematics and computer science;
- Physics;
- Chemistry;
- Geology;
- Biological sciences;
- Medical sciences;
- Agricultural sciences;
- Civil Engineering and Architecture;
- Industrial Engineering;
- Information Technology.

We have data on almost the universe of the Italian university system, composed by 69 units. For more than 80% of them we have data over the total period (1995-1999). We eliminated private universities in order to ensure comparability and leave out those universities with uncomplete data. We end up with a sample of 45 universities, whose names are listed in the following Table 3.

The Italian university system is characterized by universities that deeply differ by size and their teaching supply, but basically have the same legal status and offer degrees and diplomas with exactly the same legal value. With respect to balance sheet data, it emerges that the biggest university has a level of financial resources 88 times higher than the smallest; similar differences are at place for the teaching and student populations. The teaching supply is widely differentiated: the university of Bologna has the highest number of departments ("facoltá") with 18; around 28% of our sample have at least 10 departments, 35% have a number of departments between 5 and 9, and 37% of our sample have less than 5 departments, although in this group more than one third is composed by private universities.

In Table 1 we describe the variables used in the empirical analysis, while Table 2 reports a few summary statistics of inputs, outputs and external factors.

In the empirical application that follows, we apply an output oriented framework as, at least in the short run, the resources used by universities (human, financial and physical capital) are fixed. For the robust estimation we set a level of robustness at 10% and obtain a value of $m = 75$, which we use in the order-m efficiency computations. We adopt a triangular kernel, but we notice that the efficiency scores where stable using also other kernels with compact support.

Table 1: *Definition of inputs outputs and environmental-external factors.*

INPUTS	**Description**
Human capital	
TOTDOC	Sum of full professors, associate professors assistant professors and researchers (average - academic years 1995/96-97/98)
TECHADM	Number of Technical and Administrative staff (average - academic years 1995/96-97/98)
Financial capital	
CUMEXP	Total cumulated expenses years 1995-1999 (in million of Italian lire)
CUMEXp100iscr	Cumulated expenses per 100 enrolled students
CUMEXp1doc	Cumulated expenses per 1 scholar
Physical capital	
SPACE	Number of places in the lecture-halls (average - academic years 1995/96-97/98)
OUTPUTS	
Research	
PUB	Cumulated sum of publications years 1995-1999
CIT	Cumulated sum of citations years 1995-1999
PUBp100doc	No. of publications per 100 scholars (average values)
Teaching	
LAUCUM	Cumulated sum of graduated (and with *diploma*) academic years 1995/96-97/98
LAUCUMp100iscr	No. of graduated per 100 enrolled students (average values)
EXTERNAL FACTORS	
ISCR	No. of enrolled students - average academic years 1995/96-97/98
FACULT	No. of schools within the university
LOAD	no. of curricula (or *courses of specialisation*) activated per 100 scholars
CITPUB	Ratio of CIT over PUB
TRASFPRIV	Percentage of private contracts over total university budget (average values - 1995/99)

Table 2: *Summary statistics of inputs outputs and environmental-external factors.*

INPUTS	**Mean**	**Max**	**Min**	**St. Dev.**
Human capital				
TOTDOC	893.93	2643.53	85.47	707.84
TECHADM	972.08	4859.5	76.33	888.87
Financial capital				
CUMEX	1022933	3130092	136203	755922.1
CUMEXp100iscr	3841.88	6859.47	1328.30	1356.69
CUMEXp1doc	1246.42	2730.70	753.79	325.29
Physical capital				
SPACE	11803.19	31558	1739	7599.94
OUTPUTS				
Research				
PUB	2494.6	11360	10	2468.70
CIT	12807.18	65046	147	14449.08
PUBp100doc	253.47	498.89	3.48	112.42
Teaching				
LAUCUM	6329.76	24942	610	5965.49
LAUCUMp100iscr	20.55	32.98	10.30	5.74
EXTERNAL FACTORS				
ISCR	29529.44	102253	4898.33	25494.81
FACULT	7.18	15	2	3.26
LOAD	4.76	9.10	2.23	1.79
CITPUB	5.33	23.3	1.97	3.18
TRASFPRIV	2.18	6.99	0.02	1.70

In Table 3 we report some partial ranks of the Italian universities in our sample. The second column of Table 3 reports the rank of universities by their Cumulated expenses per 100 enrolled students (CUMEXp100iscr). The university of Reggio Calabria is the first in this ranking meaning that its CUMEXp100iscr are the highest of the sample, while the university of Salerno is the last (it occupies the place 45) in the sample, meaning that its CUMEXp100iscr are the lowest of our sample. The same interpretation has to be given to the third, fourth and fifth columns of Table 3 that report the partial ranks for, respectively, the number of graduated per 100 enrolled students (LAUCUMp100iscr), the Cumulated expenses per 1 scholar (CUMEXp1doc), and finally, the number of publications per 100 scholars (PUBp100doc).

Table 3: *Partial Ranks of Italian Universities*

University	Rank CUMEXp100iscr	Rank LAUCUMp100iscr	Rank CUMEXp1doc	Rank PUBp100doc
	1 =the worst, 45 = the best	1 = the best, 45 = the worst	1 = the worst, 45 = the best	1 = the best, 45 = the worst
Ancona	8	14	10	4
Bari	39	26	18	34
Basilicata	4	41	31	19
Bergamo	40	12	3	42
Bologna	36	7	33	14
Brescia	23	13	11	1
Cagliari	34	38	29	32
Calabria	9	30	8	26
Cassino	43	34	22	41
Catania	13	35	6	36
Chieti	32	17	19	24
Ferrara	16	27	43	3
Firenze	26	20	42	18
Genova	10	8	25	17
L'Aquila	17	36	44	9
Lecce	44	39	34	23
Messina	15	37	24	38
Milano	42	18	45	2
Milano Politecnico	37	2	12	20
Modena	7	3	39	7
Molise	33	45	4	27
Napoli Federico II	31	24	20	25
Napoli II	6	44	7	33
Napoli Navale	38	40	1	43
Napoli Orientale	14	6	14	45
Padova	21	10	17	6
Palermo	30	31	37	40
Parma	25	5	21	15
Pavia	11	4	32	5
Perugia	18	21	30	21
Pisa	22	23	38	10
Reggio Calabria	1	28	2	29
Roma III	20	43	36	37
Roma Tor Vergata	2	15	23	11
Salerno	45	32	40	30
Sassari	12	33	16	35
Siena	3	11	5	22
Teramo	41	29	9	44
Torino Politecnico	29	9	28	31
Trento	27	16	15	12
Trieste	19	22	41	16
Udine	5	25	13	13
Venezia	35	1	27	39
Verona	28	19	26	8
Viterbo Tuscia	24	42	35	28

6 Empirical results

We run a series of efficiency models of varying complexity.

We explore the efficiency of universities using separately two indicators of output, namely the research output (total number of publications - in science and social science fields- in international refereed journals, cumulated 1994-99, PUB or total number of citations - in science and social science fields- received, 1994-99, CIT) and the educational output (total number of students enrolled, average 1994-99, ISCR or cumulated total number of degrees and “diplomas” 1994-99, LAUCUM). Even if the methodology we apply here allow for multiple input multiple output analysis, in these first explorative elaborations we analyse the two outputs separately, in order to investigate the trade-off between research and teaching.[10] We present results separately for the input indicators only if results are different and deserve discussion. We study how the efficiency of each university is affected by external factors.

External factors are neither inputs nor outputs of the university production process, but they may influence the performance of universities.

As explained above (see section 4), the comparison between efficiency *conditional* to the external factors and *unconditional* efficiency informs about the sign and magnitude of the effect. This methodology is vastly superior to classical parametric regression analysis, because first it avoids the burden of functional specification of the efficient frontier and of the nature of the conditionning and second, as pointed in Florens and Simar (2004), it captures the shape of the cloud of points at its efficient boundary and not in its middle or at an average behavior (as is the case for the regression approach).

6.1 Scale and scope effects

Our initial result refers to the effect of size of universities on research efficiency. What we find is that the efficiency in the research activity model, that we call PUB MODEL (in which we use as inputs several proxies of human, financial and physical capital, and as output the total number of publication, PUB), is not affected by size effects, as measured by the number of enrolled students (ISCR). Therefore the net effect of working in a large or a small university on publication activity is not only almost zero on average, but more precisely almost zero along all the observed distribution of university size. See Appendix, Figure 3.

A similar result is observable with respect to economies of scope. Since these refer more to educational efficiency, we study how the efficiency in the teaching activity model that we call TEACH MODEL (in which we use as inputs several proxies of human, financial and physical capital, and as output the cumulated number of degrees and “diplomas”, LAUCUM), is affected by the fact that the university has many schools or faculties (number of *facoltá*,

[10]See the Conclusions for further developments of this study.

FACULT), runs many undergraduate curricula (number of different *corsi di laurea* or *corsi di diploma*, COR) or asks its professors to teach in several programmes (number of curricula per 100 professors, LOAD).

Of course these measures refer to different aspects of *specialisation*. Perhaps the most important is the number of faculties. A small number of faculties is typical of small, young universities, or of large, specialised, mainly technical universities (such as Polytechniques). Large universities usually also have a large number of faculties, but the reverse is not always true.

Given the number of faculties, the number of curricula is a decision variable for the university. Legislation in the early '90s (L.571/1993) gave Italian universities the so called autonomy, which implies the freedom to set up new curricula, within a set of ministerial guidelines and authorizations, in order to capture student demands. Further, the reform following the Bologna process encouraged universities to diversify their offering, leading to a rapid increase in the number of curricula. Given the number of professors, universities that have a strategy of diversification or extension of educational supply may ask professors to engage in more teaching activities, taking more courses per year.

All these measures are clearly *very crude approximations* for economies of scope. They try to capture first order effects, rather than going into the details of cost structures and possible synergies via resource sharing.

Furthermore, we do not have data on doctoral programmes for testing economies of scope between undergraduate and postgraduate education.

Given these limitations, data show that the number of different faculties or curricula (FACULT, COR) or number of curricula per 100 professors (LOAD) does not have any impact on the efficiency in producing graduate students (TEACH MODEL). See Appendix, Figures 4 and 5.

In some sense the same pool of students is offered a more segmented structure of offer. Proliferating the number of faculties of curricula does not improve the efficiency in production of graduate students, but perhaps makes their education more finely segmented with respect to labour market requirements. Recall that data are taken from a period where the 3+2 scheme of the Bologna process was not in place. After that reform, many students were encouraged to enrol, mainly because they considered the opportunity of a short curriculum as attractive. Under these conditions, the extension and differentiation of offer may be a sensible strategy, but we have no data to support this intuition.

The impact of differentiation of offer on research efficiency is more complex. When the number of curricula per 100 professors increases to levels of 2-3 the efficiency of universities as measured by PUB is negatively affected. Beyond this level a flat region is identified. Another negative effect region is apparent at 7-8 curricula although only a few observations are available here. It seems that universities, after an initial shock, may adjust internally to heavier burdens of teaching by redistributing tasks and maintain adequate effort for research. If the burden becomes excessive, research efficiency may suffer, but again more evidence is needed to confirm this effect. See Appendix, Figure 6.

Summing up, our results are consistent *against* the bulk of literature, that finds that "being big" and "being diversified" are not necessarily good at the university level. Economies of scale and scope are not the most important drivers of efficiency in higher education institutions. They may also be detrimental to research efficiency, although counteracting effects seem to be at play. Interestingly enough, most governments still believe that concentrating research and higher education in large institutions is the single most important policy problem.

6.2 Trade-offs

We consider two types of trade-offs: between research and teaching, and between research for publication and applied research for industry.

Again, we study the overall universities, aggregating scientific disciplines and educational programmes that are extremely heterogeneous in their pattern of use of inputs and production of outputs.

We study how research efficiency (PUB MODEL) is affected by the education activities (as measured by LAUCUM), see Figure 7 in Appendix. It comes out that a good educational efficiency (in terms of high number of LAUCUM) does not deteriorate research efficiency. Moreover, we find that increasing scientific quality (as measured by the ratio CIT/PUB) improves educational efficiency (in terms of LAUCUM), see Figure 8 in Appendix.

Furthermore, we study how research efficiency (PUB) is affected by an external variable that describes the overall importance of research collaborations with industry (percentage of university budget funded by industry, average 1994-99, TRASFPRIV). Here some interesting results emerge.

Figure 9 in Appendix shows that increasing the share of university budget represented by industrial sources has a beneficial effect on research productivity. The direction of causation may be left open: on the one hand, good universities, with a recognized international standard of publication, signal their quality and attract industry interest and money; on the other hand, universities with a higher share of industry funds benefit from more resources and improve their research productivity. As a matter of fact, the overall effect is largely positive. In the right tail of the distribution of industry funding, the impact on research productivity is negative although this particular shape is mainly due to a few of observations. If this effect was confirmed by further evidence, it would imply that the trade-off actually applies, so that being exposed to industry requests to a significant extent may reduce pure publication productivity. Of course, more definite conclusions should require more data to confirm this inverse U-shaped pattern. Even more interesting, the overwhelming majority of Italian universities are located in the region of *positive effects*, meaning that collaboration with industry and international publication are not substitute but complements.

On the other hand, the share of university budget represented by industrial sources (TRASFPRIV) does not deteriorate the educational offering of universities, as it emerges from Figure 10, in Appendix.

This is a clear example of the potentiality of robust techniques. While most discussion based on regression techniques ends up in weighting contrasting average results, we believe that understanding the impact of external variables along the entire distribution is much more informative and can reconcile opposite predictions.

6.3 An example of detailed results available

In this section we report some detailed robust results available for a model in which we use as outputs PUBp100doc and LAUCUMp100iscr, as input CUMEXP, and as external factor TRASFPRIV.

The second column of Table 4 reports the values of the unconditional efficiency score of order$-m$, the third column reports the efficiency score of order$-m$ conditioned to the value of TRASFPRIV of each university, then the value of TRASFPRIV is reported, and after that, the fifth column shows the value of Q_m^z that is the ratio of conditional on unconditional order$-m$ efficiency score, than the Externality Index of order$-m$ ($EI_m = \widehat{E}(Q_m^z)$) is showed and the Individual Index or order$-m$ (II_m^z) is presented followed by the R_m value, that is given by the ratio of Q_m^z on the geometric mean of all Q_m^z. The last two columns report the number of universities that dominates each university, Np, and the number of dominating universities conditioned to the value of Z, Np_z.

Here the universities have been disposed randomly and have been coded.

To illustrate the meaning of the variables listed in Table 4, we reveal that A is the University of Basilicata, B is the University of Lecce, C is the University of Messina and D is the University of Calabria.

From this model, it appears that the University of Basilicata and the University of Lecce are efficient in the production of their two main outputs (teaching and research). In fact the number of their dominating universities (universities that do better) is zero both unconditionally and conditionally to their level of Z (TRASFPRIV), see their value of Np and of Np_z. The University of Messina, on the contrary, could increase its outputs of more than two times, but if we take into account TRASFPRIV its performance is better as its output efficiency score falls down and is more close to unity that indicates efficiency. This is important because could indicate that one of the cause of its inefficiency could be its low level of TRASFPRIV, and an increase in private funding could be a policy measure to take into account. The case of the University of Calabria again shows that if we take a more precise comparison, taking into account the percentage of TRASFPRIV owned by the university it becomes efficient. On the contrary there are some other universities in the sample, such as L, the performance of whose is not affected by the percentage of TRASFPRIV. The remaining indicators in Table 4 offer the possibility of finely characterize the profile of each university as compared with the other universities of the sample, isolating the effects of the external factor on their overall performance.

Table 4: *Order−m efficiency scores of the Model with outputs: PUBp100doc and LAUCUMp100iscr, input CUMEXP, and Z is TRASFPRIV.*

University	$\widehat{\lambda}_m(x,y)$	$\widehat{\lambda}_m(x,y \mid z)$	Z	Q^z_m	$\widehat{E}(Q^z_m)$	II^z_m	R^z_m	Np	Np_z
A	1	0.99999	1.73	0.99999	0.96078	1.0408	1.0956	0	0
B	0.99998	1	0.84	1	0.88719	1.1272	1.0956	0	0
C	2.1181	1.4664	0.12	0.69233	0.81396	0.85058	0.75853	21	3
D	1.3012	1	1.1	0.76852	0.93174	0.82482	0.84201	6	0
E	1.0372	1.0376	3.59	1.0004	0.96771	1.0338	1.0961	1	1
F	1.4899	1.3574	2.37	0.91104	0.91803	0.99238	0.99815	15	4
G	1	0.99999	0.35	0.99999	0.82245	1.2159	1.0956	0	0
J	1.1695	1.1822	1.13	1.0109	0.93583	1.0802	1.1075	2	1
K	0.99982	0.99999	3.87	1.0002	0.96798	1.0333	1.0958	0	0
L	1.8689	1.8765	1.21	1.0041	0.9453	1.0622	1.1001	15	4
M	1	1	1.27	1	0.95099	1.0515	1.0956	0	0
N	1.8453	1.8645	1.69	1.0104	0.96181	1.0505	1.107	25	6
O	1.0567	1.0569	3.34	1.0002	0.96562	1.0358	1.0958	2	1
P	1.1867	1	1.88	0.84269	0.95465	0.88272	0.92327	1	0
Q	1.4004	1.4152	1.76	1.0106	0.95984	1.0529	1.1072	10	2
R	1.1852	1.1421	6.27	0.9636	0.95755	1.0063	1.0557	3	2
S	1.366	1.3665	3.54	1.0004	0.96757	1.0339	1.0961	1	1
T	1.0556	1.0677	3.04	1.0115	0.95274	1.0616	1.1082	1	1
U	0.99355	1	3.74	1.0065	0.96782	1.04	1.1027	0	0
V	0.99429	1	1.47	1.0057	0.96178	1.0457	1.1019	0	0
W	1	0.99999	1.37	0.99999	0.95788	1.044	1.0956	0	0
X	1.4446	1.0073	0.44	0.69727	0.8288	0.84131	0.76395	19	1
Y	2.3148	1.4164	0.67	0.61191	0.85699	0.71403	0.67042	17	5
Z	1.0904	1.0904	0.94	1	0.90571	1.1041	1.0956	2	1
AA	0.99999	1	0.17	1	0.81518	1.2267	1.0956	0	0
BB	1.0041	1.0056	3.69	1.0015	0.96781	1.0348	1.0972	1	1
CC	1.7419	1.1302	0.26	0.64881	0.81811	0.79306	0.71085	19	2
DD	1.0976	1	3.24	0.91105	0.9632	0.94587	0.99817	1	0
EE	0.99721	1	4.22	1.0028	0.97284	1.0308	1.0987	0	0
FF	1.4136	1.4348	1.95	1.015	0.95066	1.0677	1.112	6	2
GG	1.1842	1.1248	1.83	0.9498	0.95708	0.9924	1.0406	5	2
JJ	1.2832	1	0.02	0.77932	0.81205	0.9597	0.85384	4	0
KK	2.238	1.9598	2.16	0.87571	0.93511	0.93648	0.95945	13	3
LL	1.2059	1.21	5.69	1.0034	0.99302	1.0104	1.0993	4	3
MM	1.4111	1.0064	0.32	0.71319	0.82081	0.86889	0.78139	8	1
NN	1.728	1.4163	3.86	0.81959	0.96795	0.84672	0.89796	10	4
OO	1.2031	1.0947	3.42	0.90994	0.96675	0.94124	0.99695	4	2
PP	1	0.99999	0.34	0.99999	0.82188	1.2167	1.0956	0	0
QQ	1.1878	1	6.99	0.8419	0.86137	0.9774	0.9224	1	0
RR	1.0469	1.0469	4.55	1	0.98504	1.0152	1.0957	1	1
SS	1.3511	0.99999	0.77	0.74013	0.87422	0.84661	0.8109	5	0
TT	1.3065	1.0561	2.56	0.80837	0.91142	0.88693	0.88566	5	1
UU	0.99855	1	1.5	1.0015	0.96242	1.0406	1.0972	0	0
VV	1.1427	0.99999	1.63	0.87511	0.9628	0.90892	0.95879	1	0
WW	1	1	1.12	1	0.9345	1.0701	1.0956	0	0
Mean	1.27690	1.15180	2.17820	0.92100	0.92344	0.99650	1.0091	5.0889	1.2222

7 Conclusions

This paper has explored two large debated issues in the economics and policy of science and education.

With respect to the problem of economies of scale and scope we add evidence to the notion that they are not significant factors in explaining research and education productivity.

With respect to the issue of trade-offs we find that beyond a threshold quality of publication increasing scientific quality improves educational efficiency. On the other hand, a good educational efficiency does not deteriorate research efficiency.

With respect to the trade off between academic publication activities and applied industrial research, we come to the conclusion that trade offs effects (complementarity/rivalry relations) have a *local* characterization. More evidence is needed to confirm the inverted U-shaped conjecture that appears from our first exploration of data.

From a methodological point of view, our results clearly show the advantage of robust nonparametric techniques over standard regression techniques and regression-based frontier estimation techniques both from the methodological rigor of the analysis and for the richness of information for interpretation and policy consideration.

In this paper we presented a first preliminary investigation on the Italian system of universities at aggregate level. We are working on the development of a model with multiple outputs that jointly capture the effects of teaching and publication activities using the parametrization of robust nonparametric frontiers, as proposed by Florens and Simar (2004). This will allow the estimation of elasticities of substitution between outputs in a robust way, avoiding the drawback of regression-based frontier estimation techniques which capture the shape of the cloud of points (and so elasticities of substitution of the production function) in its middle (due to the averaging of the regression) and not, as it should, at its efficient boundary. These developments will permit to confirm the nature of local effects that we have identified in this research.

After that, we are going to carry out an analysis at more disaggregated level, investigating more finely on the effects of scale, scope and trade-offs at the level of schools/faculties. We also plan to enrich the analysis by using information on the patenting activity of Italian professors and data on territorial agglomeration.

A promising line of development of this work would be the extension of the analysis at European level, developing a quantitative approach to characterize and compare the profiles of European universities.

References

[1] Abbott M., Doucouliagos C. (2003) The efficiency of Australian universities: A data envelopment analysis approach. *Economics of Education Review*, 22, 89-97.

[2] Adams, J.D. and Griliches, Z. (2000), "Research Productivity in a System of Universities", in Encaoua, D. et al.(eds.), *The Economics and Econometrics of Innovation*, 105-140, Kluwer Academic Publishers, Netherlands.

[3] Balconi M., Breschi S., Lissoni, F. (2004) Networks of inventors and the role of academia: An exploration of Italian patent data, *Research Policy*, 33(1), 127-145.

[4] Bonaccorsi A. (2003), eds., *Il sistema della ricerca pubblica in Italia*, F.Angeli, Milan.

[5] Bonaccorsi A. and Daraio, C. (2003), "A robust nonparametric approach to the analysis of scientific productivity", *Research Evaluation*, 12 (1), 47-69.

[6] Bonaccorsi A. and Daraio, C. (2004a), "Econometric approaches to the analysis of productivity of R&D systems. Production functions and production frontiers", in H.F. Moed, W. Glanzel and U. Schmoch (edited by), *Handbook of Quantitative Science and Technology Research*, Kluwer Academic Publishers, 51-74.

[7] Bonaccorsi A. and Daraio, C. (2004b), "Exploring size and agglomeration effects on public research productivity", *forthcoming* in *Scientometrics*.

[8] Breschi S., Lissoni F., and Montobbio F. (2004), "Open science and university patenting: A bibliometric study of the Italian case", paper presented at the Workshop "The empirical economic analysis of the academic sphere" March 17, 2004, BETA, Univ. Louis Pasteur, Strasbourg (France).

[9] Brinkman P.T. (1981) "Factors affecting instructional costs at major research universities". *Journal of Higher Education*, 52, 265-279.

[10] Brinkman, P.T., Leslie L.L. (1986) "Economies of scale in higher education: Sixty years of research". *The Review of Higher Education*, 10 (1), 1-28.

[11] Calderini M., Franzoni C. (2004) "Who is patenting in public research? And to what detriment? An empirical analysis of the relationship between scientific careers and patent applications", Paper presented to the 4th workshop on Economic Transformation in Europe, Sophia Antipolis, January 29-30, 2004.

[12] Cazals, C., Florens, J.-P. and Simar, L. (2002), "Nonparametric frontier estimation: a robust approach", *Journal of Econometrics*, 106, 1-25.

[13] Campbell E.G., Blumenthal D. (2000), "Data withholding in academic medicine: characteristics of faculty denied access to research...", *Research Policy*, Vol 29(2), 303-313.

[14] Charnes, A., Cooper, W.W., and Rhodes, E. (1978), "Measuring the Efficiency of Decision Making Units", *European Journal of Operational Research*, 2, 429-444.

[15] Cohn E., Rhine S.L.W., Santos M.C. (1989) "Institutions of higher education as multiproduct firms: Economies of scale and scope", *Review of Economics and Statistics*, 71 (May), 284-290.

[16] Cole, S., Cole, J. and Simon, G. (1981), "Change and consensus in peer review", *Science*, vol. 214.

[17] Daraio, C. (2003), *Comparative Efficiency and Productivity Analysis based on Nonparametric and Robust Nonparametric Methods. Methodology and Applications*, Ph.D. Dissertation, Sant'Anna School of Advanced Studies, Pisa.

[18] Daraio C. and Simar, L. (2003), "Introducing environmental variables in nonparametric frontier estimation: a probabilistic approach", Discussion Paper no. 0313, Institut de Statistique, UCL, Belgium, *forthcoming* on *The Journal of Productivity Analysis*.

[19] David P. (2000) "The digital technology boomerang: new intellectual property rights threaten global open science" *forthcoming* in the World Bank Conference Volume: ABCDE-2000.

[20] Debreu, G. (1951), "The Coefficient of Resource Utilization", *Econometrica*, 19, 273-292.

[21] De Groot H., McMahon W.W., Volkwein J.F. (1991) "The cost structure of American research universities". *Review of Economics and Statistics*, 424-451.

[22] Deprins, D., Simar L. and Tulkens H. (1984), "Measuring labor-efficiency in post offices", in Marchand, M., Pestieau, P. and Tulkens, H. (eds.) *The Performance of public enterprises - Concepts and Measurement*, Amsterdam, North-Holland, 243-267.

[23] Farrell, M.J. (1957), "The measurement of the Productive Efficiency", *Journal of the Royal Statistical Society*, Series A, CXX, Part 3, 253-290.

[24] Florens, J.P. and Simar, L. (2004), "Parametric Approximations of Nonparametric Frontier", Discussion Paper no. 0222, Institut de Statistique, UCL, Belgium, forthcoming in *Journal of Econometrics*.

[25] Gautier A., Wauthy X., (2004), “Teaching versus research: a multi-tasking approach to multi-department universities, paper presented to the 21st JMA, Lille, 27-28 May 2004.

[26] Geuna A. (1999) “The changing rationale for university research funding: Are there negative unintended consequences?” *Journal of Economic Issues*, Vol. 35, 607-32. Also available as SPRU Electronic Working Paper Series No. 33.

[27] Geuna A., Nesta L. (2004) “University patenting and its effects on academic research. The emerging European evidence”, Forthcoming, *Research Policy*.

[28] Gittelman M., Kogut B. (2003), “Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns”, *Management Science*, Vol 49 (4), 366-383.

[29] Graves P., Marchand J.R., Thompson R. (1982), “Economics Departmental Rankings: Research Incentives, Constraints and Efficiency”, *American Economic Review*, Vol. 72(5), 1131-1142.

[30] Halsey A.H. (1980), *Higher education in Britain - A study of University and Polytecnhnic teachers*, Final report on SSRC Grant.

[31] Henderson R., Jaffe A.B., Trajtenberg M. (1998) “ Universities as a source of commercial technology: A detailed analysis of university patenting” 1965-1988, *Review of Economics and Statistics*, Vol. 80, 119-127.

[32] Hicks D., Hamilton K. (1999) “Does university-industry collaboration adversely affect university research?”, *Issues in Science and Technology online*, 74-75, Summer 1999, http://www.nap.edu/issues/15.4 .

[33] Johnston R. (1994), “Effects of resource concentration on research performance”, *Higher Education*, 28(1), 25-37.

[34] Lloyd P., Morgan M., Williams R. (1993) Amalgamations of universities: are there economies of size and scope? *Applied Economics*, 25, 1081-1092.

[35] Martin B.R. (2001) “The changing social contract for science and the evolution of the university”, in Geuna A., Salter A.J., Steinmuller W.E. (eds.) *Science and innovation: Rethinking the rationales for funding and governance*, Edward Elgar, Cheltenham.

[36] Mowery D.C., Nelson R.R., Sampat B.N., Ziedonis A.A. (1991) “The growth of patenting and licensing by US universities: An assessment of the effects of the Bayh-Dole Act of 1980”, *Research Policy*, Vol. 30, 99-119.

[37] Mowery D.C., Sampat B. (2001) "University patents and patent policy debates in the USA, 1925-1980", *Industrial and Corporate Change*, 10(3), 781-814.

[38] Mowery D.C., Ziedonis A.A. (2002) "Academic patent quality and quantity before and after the Bayh-Dole Act in the United States", *Research Policy*, Vol. 31, 399-418.

[39] Narin, F., Hamilton, K.S. (1996), "Bibliometric performance measures", *Scientometrics*, Vol. 36, 293-310.

[40] Nelson R. (2004) "The market economy and the scientific commons", *Research Policy*, 33(3), 455-471.

[41] Nelson R., Hevert K.T. (1992) Effect of class size on economies of scale and marginal costs in higher education. *Applied Economics*, 24, 473-482.

[42] Ramsden P. (1994), "Describing and explaining research productivity", *Higher Education*, 28, 207-226.

[43] Reskin B.F. (1977), "Scientific productivity and the reward structure of science", *American Sociological Review*, vol. 42, p. 491-504.

[44] Ruggiero J. (2004), "Performance evaluation in education. Modeling educational production", in W.W.W. Cooper, L.M. Seiford and J. Zhu eds., *Handbook on Data Envelopment Analysis*, Kluwer Academic Publishers, pp. 323-348.

[45] Shephard, R.W. (1970). *Theory of Cost and Production Function.* Princeton, NJ: Princeton University Press.

[46] Simar, L. and Wilson, P.W. (2003), "Estimation and Inference in Two-stage, Semiparametric Models of Production Processes", Discussion Paper no. 0307, Institut de Statistique, UCL, Belgium.

[47] Thursby J.G., and Kemp S. (2002), "Growth and productive efficiency of university intellectual property licensing", *Research Policy*, 31(1), 109-124.

[48] Verry D.W., Layard P.R. (1975) Cost functions for university teaching and research, *Economic Journal*, 85, 55-74.

[49] Verry D.W., Davis B. (1976) *University costs and outputs.* Amsterdam: Elsevier.

[50] von Tunzelmann N., M. Ranga, B. Martin and A. Geuna (2003), "The Effects of Size on Research Performance: A SPRU Review", Report prepared for the Office of Science and Technology, Department of Trade and Industry.

[51] Zhang, Y., R. Bartels (1998), "The Effect of sample size on mean efficiency in DEA with application to electricity distribution in Australia, Sweden and New Zealand", *The Journal of Productivity Analysis*, 9, 187-204.

Appendix

Scale and scope effects

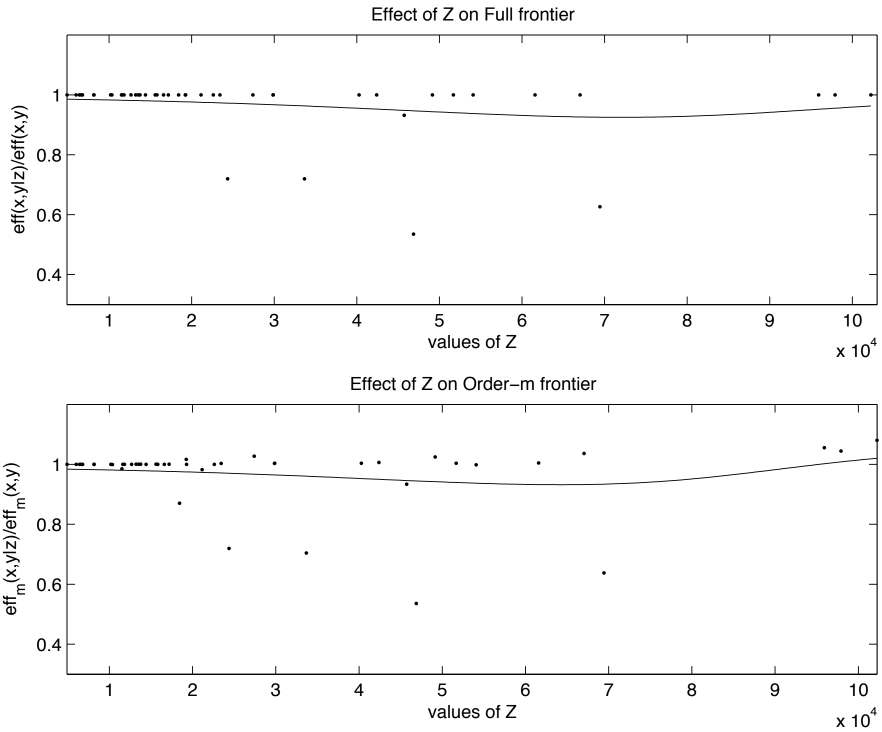

Figure 3: *PUB MODEL Z ISCR.*

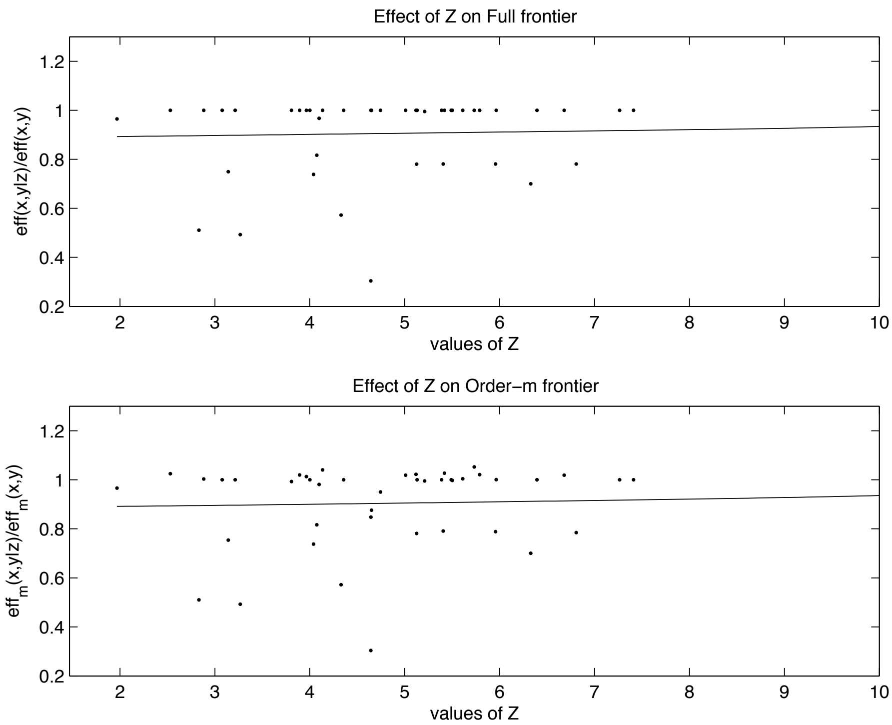

Figure 4: *TEACH MODEL Z FACULT.*

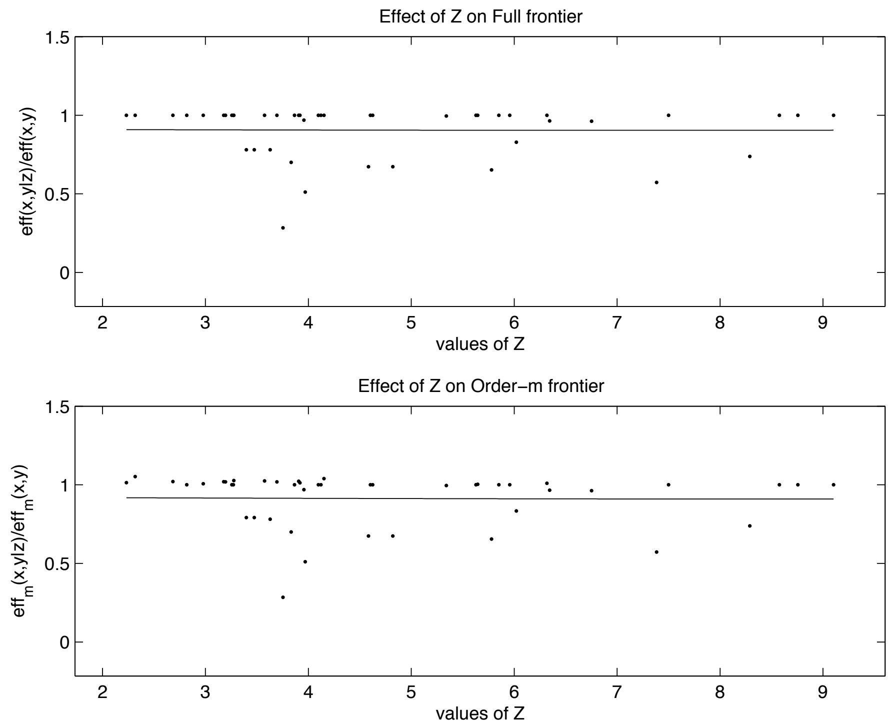

Figure 5: *TEACH MODEL Z LOAD.*

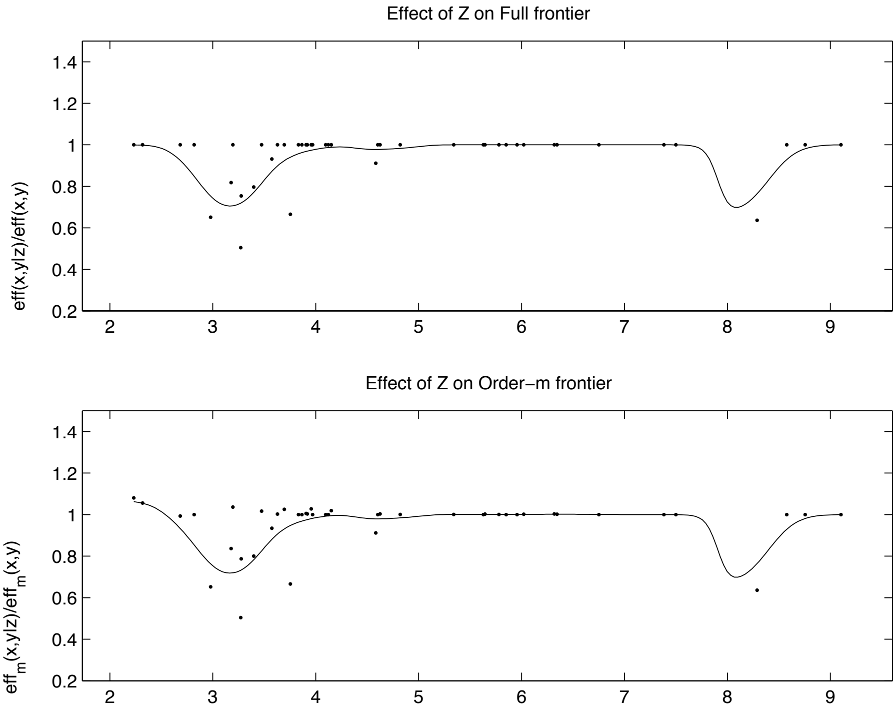

Figure 6: *PUB MODEL Z LOAD.*

Trade-off research vs teaching

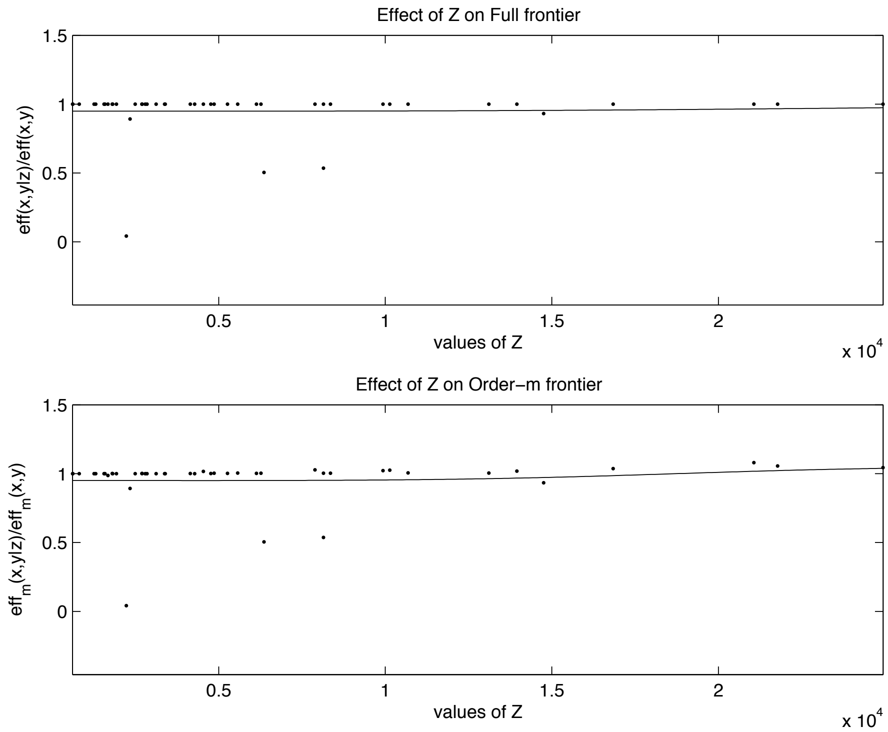

Figure 7: *PUB MODEL Z LAUCUM.*

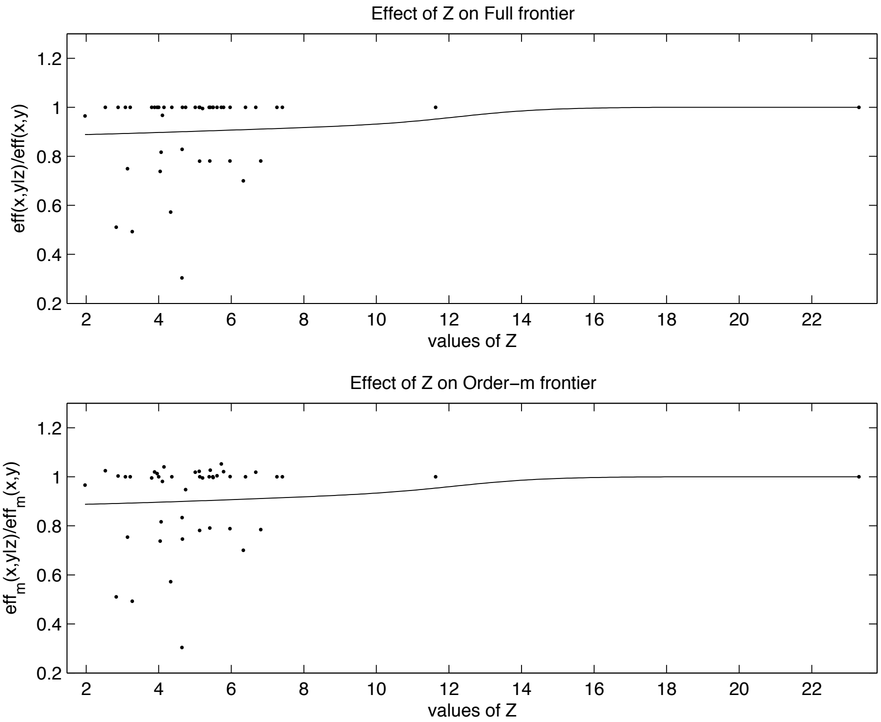

Figure 8: *TEACH MODEL Z CIT/PUB.*

Trade off scientific research vs applied/industry research

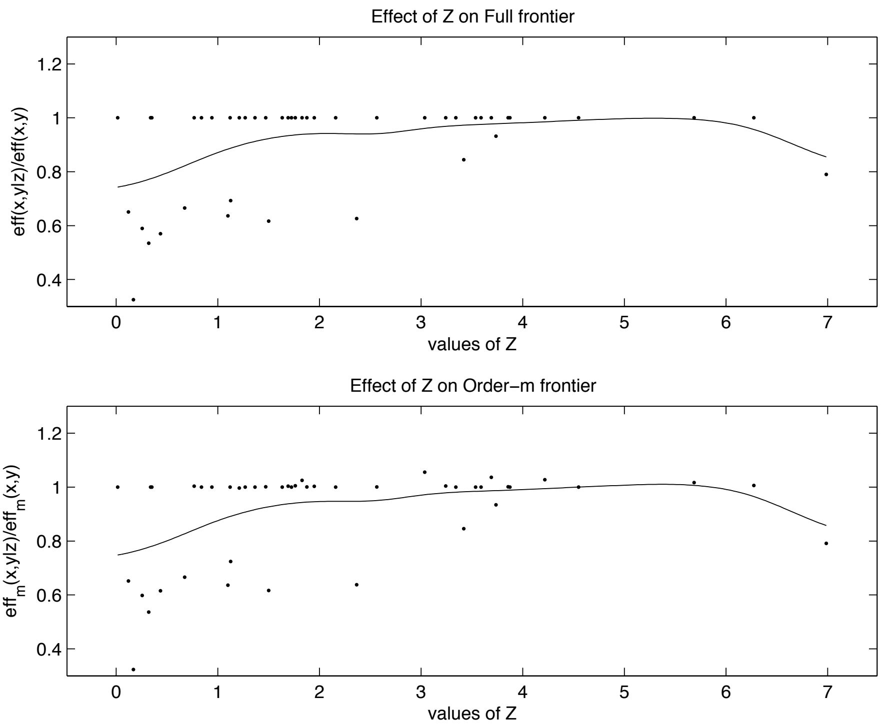

Figure 9: *PUB MODEL Z TRASFPRIV.*

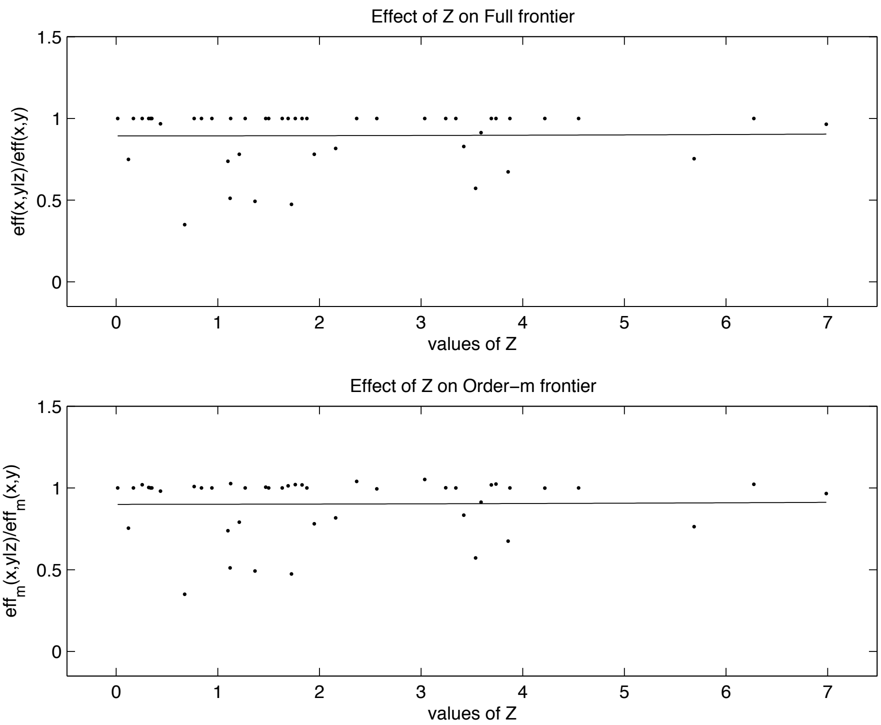

Figure 10: *TEACH MODEL Z TRASFPRIV.*