

I N S T I T U T D E S T A T I S T I Q U E
B I O S T A T I S T I Q U E E T
S C I E N C E S A C T U A R I E L L E S
(I S B A)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN



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

2014/20

Efficiency and economies of scale and scope in European universities: a directional distance approach

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

EFFICIENCY AND ECONOMIES OF SCALE AND SCOPE IN EUROPEAN UNIVERSITIES: A DIRECTIONAL DISTANCE APPROACH

ANDREA BONACCORSI CINZIA DARAIO* LÉOPOLD SIMAR

May 12, 2014

Abstract:

In this paper we investigate economies of scale and scope of European universities. The proposed approach builds on the notion that university production is a multi-input multi-output process different than standard production activity. The analyses are based on an interesting database which integrates the main European universities data on inputs and outputs with bibliometric data on publications, impact and collaborations. We pursue a cross-country perspective; we include subject mix and introduce a robust modeling of production trade-offs. Finally, we test the statistical significance of scale and scope and find that size and specialization have a statistical significant impact both jointly and separately, showing an inverted u-shape effect on efficiency.

Keywords: efficiency, national academic systems, disciplinary specialization, research performance, teaching and research, nonparametric and robust frontier estimation, bootstrap.

***Bonaccorsi:** on leave, Department of Energy and Systems Engineering, University of Pisa, Italy; email a.bonaccorsi@gmail.com. **Daraio:** Department of Computer, Control and Management Engineering Antonio Ruberti (DIAG), University of Rome “La Sapienza”, Via Ariosto, 25 - 00185 Roma, Italy; email daraio@dis.uniroma1.it. Financial support from the “Progetto di Ricerca di Ateneo 2013” of the University of Rome “La Sapienza” is gratefully acknowledged. **Simar:** ISBA, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, and DIAG; email leopold.simar@uclouvain.be. Financial support from the “Inter-university Attraction Pole”, Phase VII (No. P7/06) of the Belgian Government (Belgian Science Policy) is gratefully acknowledged.

1 Introduction

This paper addresses two contested issues that are at the core of recent debates in higher education and makes the argument that, in order to address them sensibly, there is a need for the integration of existing data and for new aggregation techniques. Thus, although the ultimate issue is a policy one, the approach we suggest makes heavy use of informetrics¹. To be more precise: we argue that without an investment into new informetrics, these policy issues cannot be addressed appropriately. The two issues under discussion can be formulated as follows:

(a) how is it possible to increase the efficiency of the higher education system?

(b) should we reconsider the traditional model of universities, in which teaching and research are produced jointly by the same academic staff?

These two questions come after the higher education system, in advanced countries, has reached the point of massification (i.e. enrolment rates exceeding 50% of the relevant age cohort), while the public budget has not grown correspondingly. Universities are put under pressure to use existing resources, namely staff and funding, in the most efficient way. At the same time there is an increased pressure from the research side: the expectations of society and policy makers on the contribution of research to societal problems have grown significantly, there are new entrants in scientific arena (particularly from Asia) and the competition for funding has increased sharply. This situation creates a classical issue in public policy: we have two valuable goals (serving better mass educational needs and producing good research) between which there is tension. The trade-off between the two goals would require a grounded theory of production, which can be framed in the economic language. If we assume that universities are units of production, then these issues require investigating the existence and importance of economies of scale and scope. Do we need to increase the size of universities, in order to enhance their efficiency? Do universities benefit from having inputs (staff and funding) that can produce jointly teaching and research, or there are efficiency-enhancing specialization effects that suggest to keep these activities under separate institutions?

The paper is organized as follows.

¹In Bonaccorsi, Daraio and Simar (2013) we analyse the impact of scale and scope on the research efficiency of European universities. In this paper we extend the analysis including additional bibliometric indicators such as Normalized impact, high quality publications, Excellence rate and international collaborations. Moreover, we test the impact of scale and scope by applying state of the art approaches (Daraio and Simar, 2014).

In the next section, the relevant literature as well as the main research questions addressed in the paper are outlined. Section 3 describes the main data used in the analysis. Section 4 provides a simplified graphical illustration of university's activities and their trade-offs. Section 5 provides the methodological background, while Section 6 reports the main results and Section 7 concludes the paper.

2 Economies of scale and scope in higher education

In this section we offer a short and focused survey of the literature.

Economies of scale refer to the reduction of cost per unit of output when the size of operations increases, mainly due to the reduction of unitary fixed costs, but often due also to lower variable costs.

Economies of scope arise when the costs of production of two or more goods produced together by the same firm are lower than the costs of producing them separately by specialized firms (Baumol, Panzar and Willig, 1988).

Before entering into the details, let us remind that the issue of economies of scale and scope has been addressed according to two different approaches.

The first has worked directly with cost functions as the dual of production functions. Here the main difficulty has been the modeling of a production function which is, by definition, not only multi-input (as any production function), but also multi-output. The traditional econometric techniques used to estimate economies of scale in a monoprodukt setting were clearly inadequate. After the introduction of a full scale theory of the multi-product firm (Baumol, Panzar and Willig, 1982), several appropriate econometric techniques have been introduced (see Bonaccorsi and Daraio (2004) for an overview).

The second has adopted the approach initiated by Farrell (1957), based on the estimation of technical efficiency of the units under analysis, namely the best use of resources (inputs) to realize their outputs. In this line of research, the existence and magnitude of economies of scale and scope is derived from the difference between the efficiency scores of observed DMUs and the scores that would be obtained if the inputs (and/or outputs) were aggregated. In nonparametric efficiency analysis, traditionally based on a DEA approach (see e.g. Fare, Grosskopf and Lovell, 1994), economies of scope are computed by estimating the frontier of multiproduct firms and the frontier of firm constructed from the sum of specialized firms. This approach, however, introduces in the analysis additional assumptions (which rely on the convexity and additional assumptions on the hypothetical firm, and the sample size bias). Recent works in efficiency analysis (see e.g. Daraio and Simar, 2007) propose the conditional nonparametric analysis to investigate the impact of scale and scope, which are considered as external- environmental factors that are neither inputs nor outputs under the control

of the DMU, but might influence the performance of the units. In this paper we follow the foregoing approach, extending the efficiency methodology to robust and conditional directional distances and implementing a recently introduced test (Daraio and Simar, 2014), based on the bootstrap, to assess the significance of scale and scope impact.

In the following we proceed with the description of the literature and explicit the main research questions we address in the paper.

2.1 Are larger universities more efficient?

It is not surprising that a large literature has addressed the issue of economies of scale in higher education. Brinkman and Leslie (1986) review the first 60 years of empirical studies, most of which from United States. After almost 20 years, Cohn and Cooper (2004) have offered a comprehensive survey of findings from the cost function perspective, while Johnes (2006) has reviewed the technical efficiency literature. In general the literature has addressed the issue of increasing returns to scale in the two core production processes of universities, namely teaching and research.

Teaching is a complex process, whose technology is yet poorly understood. As several authors have noted (e.g. Hanushek, 1986; Worthington, 2001; Johnes, 2006), we really do not have a full scale theory of higher education teaching. Teaching is subject to economies of scale since expanding the size of the class of students expands the output (number of students attending a lecture) while keeping constant the input (the lecturing staff). At the same time teaching also require one-to-one interaction with students, such as examinations and tutoring, for which costs are roughly proportional to the output. The exact combination between these two opposite forces is responsible for the overall effect. As a matter of fact, the existence of economies of scale in undergraduate teaching is largely established in the literature (Cohn, Rhine and Santos, 1989; Dundar and Lewis, 1995; Glass, McKillop and Hyndman, 1995; Hashimoto and Cohn, 1997; Koshal and Koshal, 1999; Laband and Lentz, 2003).

Research is an even less understood production process, for which the arguments for economies of scale are mostly linked to indivisibilities in cognitive capital (minimum scale of research teams) and above all in physical capital (scientific instrumentation). A dedicated literature has examined this issue repeatedly and has been reviewed by SPRU in the early 2000s at a request of the UK government (von Tunzelmann et al., 2003). The overall synthesis was that we do not have compelling evidence on the positive impact of the size of research organizations on scientific productivity.

It has also been noted that size may be associated to other factors, such as the pressure for visibility and the quality of the intellectual environment (Qureshi et al., 2003; Seglen and Aksnes, 2000; Bonaccorsi and Daraio, 2005). More recently, Carayol and Matt (2006) have

stressed that it is not size per se but the adoption of policies for the recruitment of high quality researchers that make a difference. Horta and Lacy (2011) have found that researchers in larger research units have indeed a larger network of scientific contacts and tend to publish more at the international level. Combining the two production processes, a summary of findings from Brinkman and Leslie (1986) is that economies of scale in higher education are pervasive, although they tend to be exhausted at a relatively small scale, in the order of 1000 full time equivalent students (FTE). Confirming the survey from Brinkman and Leslie (1986) and the results of Cohn, Rhine and Santos (1989), Johnes (2006) find economies of scale at the level of university, but claim that they are exhausted at relatively small size. These results are generally confirmed by stating that the main sources of economies of scale for universities come from undergraduate education, while research contributes little to increasing returns or even is subject to decreasing returns, with postgraduate education somewhat in the middle. Recently Brandt and Schubert (2014), by using data on Germany, show that research is subject to diminishing returns to scale at the level of research team. At the same time, universities offer an umbrella to research teams which is subject to increasing returns to scale, due to shared infrastructures, better efficiency in administrative activities and reputational effects. This might explain the dominant organizational model of universities, based on a number of semi-autonomous research teams, which however accept to operate under the administrative umbrella of universities.

2.2 Should educational and research activities be organized under the same institution?

There are several types of economies of scope. We focus here on the complementarity between teaching and research, which is at the core of the Humboldtian model of university (Schimank and Winnes, 2000). The overwhelming evidence is that it is more efficient to organize teaching and research in the same organization, asking academic staff to allocate their time budget accordingly (Johnes, 2004). Longlong et al. (2009) argued that a reform of the Chinese system that forced researchers in the traditional Academy system to teach would generate a large increase in efficiency due to pervasive economies of scope at all levels of output. Contrary to the majority of studies, however, some authors (see e.g. Izadi et al., 2002) did not find evidence of economies of scope. The evidence of full economies of scope is not incompatible with the evidence suggesting that product-specific economies of scope in research may be elusive. In other words, given the time budget of academic staff, there is evidence that research productivity diminishes with teaching loads. University departments with higher teaching loads have lower research outcomes (Worthington, 2011). For this reason, rather than putting into question the overall model, it is more interesting

to investigate the most efficient mix between research and teaching.

2.3 How generalizable are these results?

While these results deliver a rich array of implications, they mostly come from country-level studies. Therefore they are subject to serious problems of generalizability, which is a major concern for policy making if decisions must be made based on the evidence of other, poorly comparable, institutional contexts. In addition, existing studies do not offer separate analyses by disciplinary fields. The first wave of studies has been dealing with USA and Anglosaxon countries, partly due to better availability of data, partly as a consequence of major structural reforms of the university system starting in the 1980s in countries such as United Kingdom, New Zealand, Australia and Canada.

The dominance of Anglosaxon countries in the literature creates an issue of generalizability. The issue at stake is not, as it is often stated, the role of the private sector, which is instead marginal, for example, in UK or Canada. The issue is that, according to OECD, these countries have an institutional framework and labour market conditions that allow a much higher mobility of inputs, such as staff, as well mobility of students. In addition, the autonomy of universities in recruitment decisions is quite high. Placed under conditions of competition for funding, it is likely that universities in these countries enjoy more room for structural adaptation. Not surprisingly, almost all studies on UK and Australia concluded that universities operate at fairly high levels of efficiency. Among multi-country studies the generalizability is still limited, either because of a small set of countries, or because of small numbers of country observations. An example of study with a cross-country perspective is Joumady and Ris (2005), which is however based on a survey of graduates across European countries. Bonaccorsi and Daraio (2007) examined a dozen of countries based on data coming from the Aquameth project, the first research project that collected comparable data on European universities².

There are also limits in generalizability due to disciplinary differences. Dunder and Lewis (1995) argued that without a careful distinction among disciplines it is impossible to derive meaningful implications. According to them ‘the most important problem seems to be that different production technologies among academic disciplines may generate problems in analyzing departmental cost functions. For instance, results can be quite misleading if a single cost function is estimated for both chemistry and English departments because they have quite dissimilar production functions’ (Dunder and Lewis, 1995, p. 120). The impact of disciplinary specialization on university performance has been also analysed in Lopez-Illescas et al. (2011) and in Moed et al. (2011) that rightly emphasized that subject mix should be

²see Bonaccorsi and Daraio, 2007; Daraio, Bonaccorsi et al. (2011).

taken into account in the assessment of university performance.

This paper builds upon the first studies that have explicitly adopted a multi-country perspective (Daraio et al. 2011), benefiting from the construction of the Eumida dataset (Daghbashyan, Deiacio and McKelvey, 2014; Bonaccorsi, Daraio and Simar, 2014). Moving further in the direction of generalizability, this paper also introduces, although only partially, a cross-discipline perspective.³

3 Data

We exploit a large database, recently constructed by the EUMIDA Consortium (European Universities Micro Data, EUMIDA, 2010) under a European Commission tender, supported by DG EAC, DG RTD, and Eurostat.

This database is based on official statistics produced by National Statistical Authorities in all 27 EU countries (with the exception of France and Denmark) plus Norway and Switzerland. The EUMIDA project, relying on the results of the Aquameth project (Bonaccorsi and Daraio, 2007; Daraio et al. 2011) collected two data sets. Data Collection 1 (DC 1) collected a set of uniform variables on all 2457 higher education institutions that are active in graduate and postgraduate education (i.e. universities), but also in vocational training. Accordingly, all institutions delivering ISCED 5a and 6 degrees are included, and the subset of those delivering ISCED 5b degrees that have a stable organization (i.e. mission, budget, staff). Those institutions altogether constitute the perimeter of higher education institutions in Europe.

Data Collection 2 (DC 2) instead included a larger set of variables on the 850 research active institutions that are also doctorate awarding. Interestingly, the number of HEIs research active is 1364, but only 850 of these are also doctorate awarding institutions. This means that a significant portion of research active institutions is found outside the traditional perimeter of universities, that is in the domain of non-university research (particularly in countries with dual higher education systems).

Data refer to 2008, or to 2009 in some cases.

We integrate the EUMIDA data, in particular the DC 2 dataset, with the Scimago data (SIR World Report 2011, period analyzed 2005-09) which include institutions having published at least 100 scientific documents of any type, that is, articles, reviews, short reviews, letters, conference papers, etc., during the period 2005-2009 as collected by Scopus database. From Scimago data we used the following variables:

- number of publications in Scopus (PUB);

³This will be done in the modeling part below in which we use as proxy of scope (the wideness of activities carried out) the specialization index (SPEC). For more details, see the next section.

- Specialization index (SPEC) of the university that indicates the extent of thematic concentration / dispersion of an institution's scientific output; its values range between 0 to 1, indicating generalist vs. specialized institutions respectively. This indicator is computed according to the Gini Index and in our analysis it is used as a proxy of the specialization of the university. We follow previous bibliometric studies by Lopez-Illescas et al. (2011) and Moed et al. (2011) that showed the usefulness of categorizing universities in generalist versus specialist by means of the Gini index.⁴
- International Collaboration (IC), a university's output ratio produced in collaboration with foreign institutions.
- High Quality Publications (Q1), a university's ratio of publications published in the first quartile (25%) in their categories, according to the Scimago journal rank indicator.
- Normalized Impact (NI), it shows the relation between an institution's average scientific impact and the world average (that is set to one).
- Excellence Rate (EXC), it is the percentage of publications included in the 10% of the most cited papers in their respective scientific fields.

Table 1 defines and describes the inputs, outputs and conditioning factors that are used in the following analysis.

⁴See also Egghe and Rousseau (1990) for more details on disciplinary specialization indices.

Table 1: Definition of inputs, outputs and conditioning factors

Input	Definition
Output	
Conditioning factor	
Input	
NACSTA	Number of non academic staff
ACSTAF	Number of academic staff
PEREXP	Personnel expenditures (PPS)
NOPEXP	Non-personnel expenditures (PPS)
FINP	Input factor including: NACSTA, ACSTAF, PEREXP, NOPEXP
Output	
TODEG5	Total Degrees ISCED 5
TODEG6	Total Degrees ISCED 6 (Doctorate)
PUB	Number of published papers (Scimago)
IC	International collaboration (Scimago)
NI	Normalized impact (Scimago)
Q1	High quality publications (Scimago)
EXC	Excellence rate (Scimago)
FRES	Factor of volume of research including: TODEG6, PUB
FQUAL	Factor of quality of research including: IC, NI, Q1, EXC
Conditioning factors	
SIZE	It is the log of the sum of Total Students enrolled ISCED 5 and Total Students enrolled ISCED 6
SPEC	Proxy of Specialization Gini index of the scientific output (Scimago)

Source: Eumida DC2 and Scimago.

As commonly used in applied econometrics, the size is computed as the *log* of the total volume of the activity, that in our case is proxied by the sum of enrolled students at all undergraduate and post-graduate levels.

From a preliminary data analysis, we found that PUB and TODEG6 were highly correlated; that NACSTA, ACSTAF, PEREXP, NOPEXP were also highly correlated and that IC, NI, Q1, EXC were also highly correlated. We found correlations higher than 85% in all cases, and for this reason, in the analysis we used their aggregating factors, respectively

4 Production Models of European universities

In this section we present the modeling strategy of our approach. While this section introduces the main ideas of directional distances through a simple illustration, Section 5 details the methodology of directional distances and their estimation.

Figure 1 illustrates the flexibility of directional distance functions to model internal trade-offs between dimensions of the academic production. For each unit in the sample, we can assess its performance (or technical efficiency) considering also its input structure, along the research dimension (RES), considering given the teaching that it is carrying out. This corresponds for unit \mathbf{u} to move towards \mathbf{u}'' in \mathbf{u} . Alternatively, we could investigate the performance of \mathbf{u} along the teaching dimension (TEACH) keeping constant (or considering given) its research activity (this corresponds to assessing the performance of \mathbf{u} in reaching the efficient frontier from \mathbf{u} to \mathbf{u}''' in Figure 1). Finally, unit \mathbf{u} could be assessed on how it is performing in doing both teaching and research, that corresponds in Figure 1 to move towards the efficient frontier from \mathbf{u} to \mathbf{u}' .

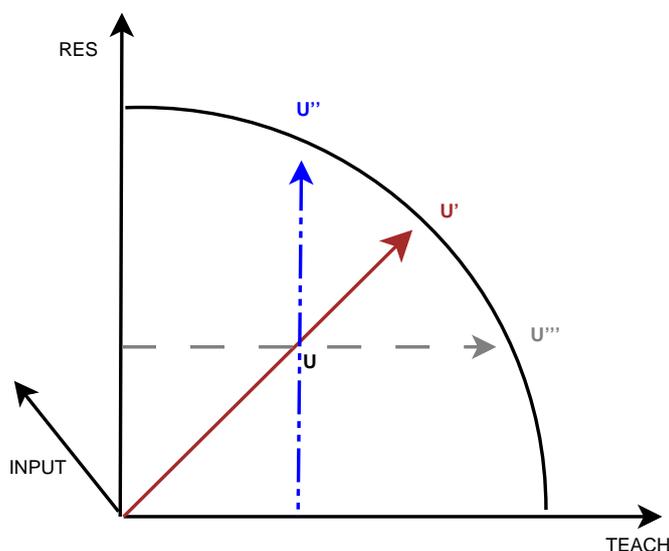


Figure 1: *An illustration of tradeoffs in the academic production.*

This is the basic illustration of the activity. Obviously, the efficiency processes described in Figure 1 may be affected by some external factors that are, at least in the short run, not

⁵The detailed results of the exploratory data analysis are not reported to save space but they are available from the authors upon request.

under the control of the units. This leads to the inclusion in the analysis of these factors whose potential impact on the performance we are interested in estimating.

In this paper we are going to evaluate the impact of SIZE and SPEC. These factors indeed might influence the probability of each unit (university) of being dominated (that is of lagging far away from the efficient boundary of the production frontier). We apply a directional output distance function, in which the direction to approach the efficient frontier is the same for each university in our sample (‘egalitarian approach’) and it is set to the European median.⁶ We think that this choice reflects the important European Research Area pillar of “cooperation and competition” because the comparison in terms of target is with respect to a median value calculated over a very skewed distribution.

We analyse the impact of scale (as proxied by the SIZE variable) and scope (as proxied by the SPEC variable) on two models of university production in Europe, namely⁷:

Humb Model Full model of academic production, in which the targets to reach the frontier are set in terms of teaching, research and quality. The following variables are used:
Input: FINP, *Outputs*: TODEG5, FRES, FQUAL; *external factors*: SIZE, SPEC.

Res Model Research model, in which teaching is considered given, and targets to reach the frontier are set in terms of research and quality. The following variables are used:
Input: FINP, *Outputs*: TODEG5 is kept constant, FRES, FQUAL; *external factors*: SIZE, SPEC.

5 Method: a flexible approach based on Directional Distances

We apply an activity analysis framework within the theory of production (see Shephard, 1970), in which producing units (hereafter “unit”), realize a set of outputs $Y \in \mathbb{R}^q$ by combining a set of inputs $X \in \mathbb{R}^p$. The technology is characterized by the attainable set T , the set of combination of (x, y) that are technically achievable

$$T = \{(x, y) \in \mathbb{R}^p \times \mathbb{R}^q | x \text{ can produce } y\}. \quad (1)$$

⁶We would like to estimate also the efficiency of the research activity itself, but this was not possible because the available inputs data refer to all the activities of universities including also teaching. We would like also to include information on the third mission activity (i.e. knowledge transfer, collaborations with industry, patents and so on), but data were not available for all the universities in our sample.

⁷Bonaccorsi, Daraio and Simar (2014) instead analyse the efficiency of a teaching model.

We know that under the free disposability assumption for the inputs and the outputs⁸, the set can be described as⁹:

$$T = \{(x, y) \in \mathbb{R}^p \times \mathbb{R}^q | H_{XY}(x, y) > 0\}, \quad (2)$$

where $H_{XY}(x, y)$ is the probability of observing a unit (X, Y) dominating the production plan (x, y) , i.e. $H_{XY}(x, y) = \text{Prob}(X \leq x, Y \geq y)$.

The efficient boundary of T is of interest and several ways have been proposed in the literature to measure the distance of the unit (x, y) to the efficient frontier. One of the most flexible approach is the directional distance introduced by Chambers et al. (1996) (see also Färe et al., 2008). Given a directional vector for the inputs $d_x \in \mathbb{R}_+^p$ and a direction for the outputs $d_y \in \mathbb{R}_+^q$, the directional distance is defined as

$$\beta(x, y; d_x, d_y) = \sup\{\beta > 0 | (x - \beta d_x, y + \beta d_y) \in T\}, \quad (3)$$

or equivalently, under the free disposability assumption (see Simar and Vanhems, 2012)

$$\beta(x, y; d_x, d_y) = \sup\{\beta > 0 | H_{XY}(x - \beta d_x, y + \beta d_y) > 0\}. \quad (4)$$

Hence, we measure the distance of unit (x, y) to the efficient frontier in an additive way and along the path defined by $(-d_x, d_y)$.

This way of measuring the distance is very flexible and generalizes the “oriented” radial measures initiated by Farrell (1957). Indeed by choosing $d_x = 0$ and $d_y = y$ (or $d_x = x$ and $d_y = 0$), we recover the traditional output (reps. input) radial distances. The flexibility is that we might have some elements of the vector d_x and/or of the vector d_y be set to zero, for focusing on the distances to the frontier along certain particular paths (for instance if some inputs or outputs are non-discretionary, not under the control of the manager, etc.).

Consistent nonparametric estimators of Equation (4) have been proposed in Simar and Vanhems (2012); Daraio and Simar (2014) analyse in details the case when some directions are set to zero, as well as statistical issues in this context.

For a discussion about the choice of a direction, see Färe et al. (2008). The direction can be different for each unit (like in the radial cases) or it can be the same for all the units. Färe et al. (2008) argue that a common direction would be a kind of egalitarian evaluation reflecting some social welfare function. Researches often select in the latter case $d_x = \mathbb{E}(X)$ and $d_y = \mathbb{E}(Y)$, where in practice empirical averages are chosen.

In this paper we select the same direction for all the units, setting a reference with respect to the European standard. The reference is made with respect to the median value of each output calculated at European level over the analysed sample.

⁸The free disposability we used in this paper is the assumption that if $(x, y) \in T$ then $(\tilde{x}, \tilde{y}) \in T$ for all $\tilde{x} \geq x$ and all $\tilde{y} \leq y$. It is a minimal assumption generally made on production processes.

⁹See Daraio and Simar (2007) for further details and illustrations.

Quantile frontiers for evaluating the performance of firms by using oriented radial measures (input or output) have been extended to directional distance in Simar and Vanhems (2012) and this extension is quite natural after the representation given in (4). In place of looking to the support of the distribution H_{XY} we benchmark the unit against a point which leaves on average $\alpha \times 100\%$ of points above the frontier. This benchmark is the α -quantile frontier. Formally the α -order directional distance is defined as

$$\beta_\alpha(x, y; d_x, d_y) = \sup\{\beta > 0 \mid H_{XY}(x - \beta d_x, y + \beta d_y) > 1 - \alpha\}. \quad (5)$$

Here a value $\beta_\alpha(x, y; d_x, d_y) = 0$ indicates a point (x, y) on the α -quantile frontier, a positive value is a point below the quantile frontier and a negative value is a point above the quantile frontier. We see clearly that when $\alpha \rightarrow 1$ we recover the full frontier definition.

A nonparametric estimator is provided in Simar and Vanhems (2012). Daraio and Simar (2013) show in details how to compute this estimator when some directions are set to zero, as well as statistical issues in this context.

The projection of any $(x, y) \in T$ on the estimated α -quantile frontier is given by the points $(\hat{x}_\alpha^\partial, \hat{y}_\alpha^\partial)$ defined as

$$\hat{x}_\alpha^\partial = x - \hat{\beta}_\alpha(x, y; d_x, d_y)d_x, \quad \text{and} \quad \hat{y}_\alpha^\partial = y + \hat{\beta}_\alpha(x, y; d_x, d_y)d_y. \quad (6)$$

Since the resulting estimator will not envelop all the data points, the resulting frontier is more robust to outliers and extreme data points than its full version above.

This is the approach we implemented in our empirical analysis.

5.1 Second stage regression: impact of scale and scope on efficiency

Badin et al. (2012) propose a general methodology to investigate the impact of external-environmental factors on the efficiency scores of units. Daraio and Simar (2014) extend the methodology to conditional directional distances to investigate the effect of z on the mean of the conditional directional distances.

This method contributes to the literature on the so called two-stage approach, where estimated unconditional efficiency scores (input or output oriented) are regressed in a second stage against the Z variables. However we know from the literature (see Badin et al. 2014 for a detailed explanation and more references) that this is valid only under a ‘separability’ assumption where it is assumed that the frontier of the attainable set is not changing with the values of z .

As indicated in Badin et al. (2012), the use of the estimated conditional efficiency scores for this second stage regression, does not requires this restrictive assumption. Hence, the

flexible second stage regression can be written as the following location-scale nonparametric regression model:

$$\beta_\alpha(X, Y; d_x, d_y | Z = z) = \mu(z) + \sigma(z)\varepsilon, \quad (7)$$

where ε and Z are independent with $\mathbb{E}(\varepsilon) = 0$ and $\mathbb{V}(\varepsilon) = 1$. This model permits to detect the location $\mu(z) = \mathbb{E}(\beta_\alpha(X, Y; d_x, d_y | Z = z))$ and the scale effect $\sigma^2(z) = \mathbb{V}(\beta_\alpha(X, Y; d_x, d_y | Z = z))$.

These two functions can be estimated non parametrically from a sample of observations $\{Z_i, \widehat{\beta}_\alpha(X_i, Y_i; d_x, d_y | Z_i)\}$, $i = 1, \dots, n$ by using local constant smoothing techniques to guarantee positive estimates of both functions, as suggested by Daraio and Simar (2014). The analysis of the estimated $\widehat{\mu}(z)$ as a function of z will enlighten the potential effect of Z on the average efficiency, with the help of $\widehat{\sigma}(z)$ which may indicate the presence of heteroskedasticity.

5.2 Testing the significance of scale and scope

Here we apply the approach of Daraio and Simar (2014). The test statistics is based on the partial derivatives of the mean function ($\mu(z) = \mathbb{E}(\beta_\alpha(X, Y; d_x, d_y | Z = z))$) that are:

$$\eta_j(z) = \partial\mu(z)/\partial z_j, \quad \text{for } j = 1, \dots, r.$$

The null hypothesis (H_0) to test is that the first r_1 components of Z do not affect $\mu(z)$ against the alternative hypothesis that some components of Z affect $\mu(z)$. The constructed test statistics, following Daraio and Simar (2014) to which the reader is referred for further details, is τ .

We will reject the null in favor of the alternative when τ is too big. Both the p-value of H_0 and critical values of any size are determined by the bootstrap.

5.3 Analyzing the gaps

It may be useful for policy makers to measure, in original units of the inputs and of the outputs, the estimated distance of an observation to the frontier. This allows us to appreciate the efforts to be achieved in increasing the outputs and reducing the inputs to reach the efficient frontier. For the full frontier these measures are given by what we call the ‘‘gaps’’ to efficiency. They are directly given by:

$$G_x = \widehat{\beta}(x, y; d_x, d_y)d_x, \quad \text{and} \quad G_y = \widehat{\beta}(x, y; d_x, d_y)d_y. \quad (8)$$

For the partial frontiers, the gaps appear as being the difference between (x, y) and the projections on the α -quantile frontier given in (6). They are particularly useful to detect

outliers in the direction given by (d_x, d_y) . This will be the case in the input direction if $G_{\alpha,x} = \widehat{\beta}_\alpha(x, y; d_x, d_y)d_x$ has some elements with large negative value: the point (x, y) is well below the estimated α -frontier in the input direction, and/or a very large negative value in some elements of the vector $G_{\alpha,y} = \widehat{\beta}_\alpha(x, y; d_x, d_y)d_y$ warns a point being well above the quantile frontier in the output direction.

As explained in Section 4, in the empirical analysis that follows in the next sections we pursue an output orientation approach aiming at estimating the efforts needed to increase the outputs of the units, given the level of their inputs used, and hence we estimate the robust gaps $G_{\alpha,y}$ in terms of percentage values of the analysed outputs.

6 Results

In this section we summarize the main results of the analysis carried out.

6.1 Impact of scale and scope on efficiency

In this subsection we report the results of the impact of scale and scope analysis obtained for the HUMB model.

Figure 2 illustrates the results of the nonparametric regression of the estimated $\mu(z)$ in function of SIZE and SPEC.

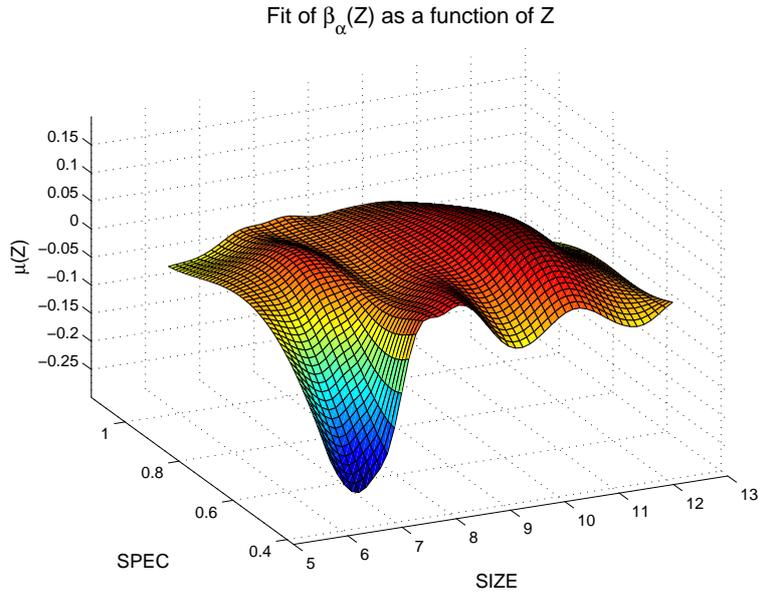


Figure 2: *Nonparametric regression of the estimated $\mu(z)$ versus $Z_1 = SIZE$ and $Z_2 = SPEC$. Note that $SIZE$ is expressed in log.*

In Figure 3 the nonparametric regression of conditional efficiency vs $SIZE$ is reported. The partial impact of $SIZE$ is represented by fixing the $SPEC$ value at its median level. We observe an inverse U-shaped impact of $SIZE$, given that $SPEC$ is fixed at its median. To read the plot we have to remind that the smaller the level of β_α greater the efficiency of the unit is. It appears that there is a lot of uncertainty when Z_1 is smaller than 8 (as pointed out by the enormous bootstrap error bounds) because there are few and heterogeneous small universities in our sample.

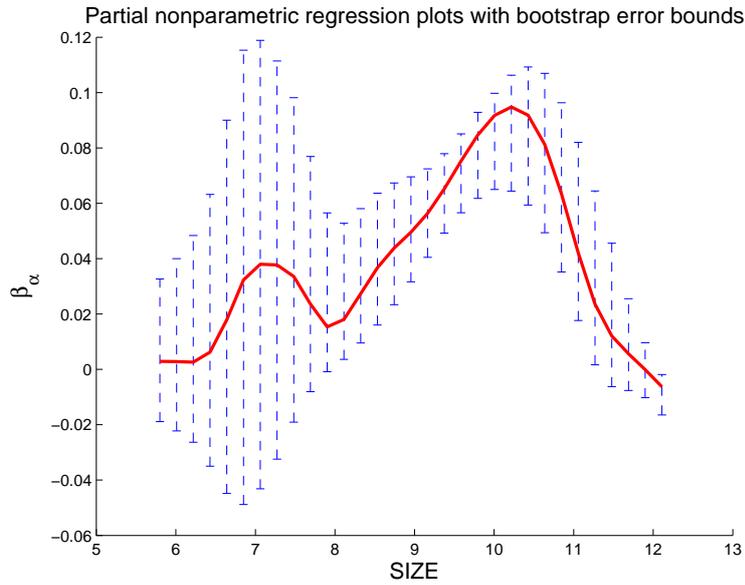


Figure 3: *Partial nonparametric regression of conditional efficiency as a function of SIZE. Note that SIZE is expressed in log.*

In Figure 4 the nonparametric regression of conditional efficiency vs SPEC is reported. The partial impact of SPEC is represented by fixing the SIZE value at its median level. We observe that SPEC has an even clearer inverse U-shaped impact with respect to SIZE; again to read the plot we have to remind that the smaller the level of β_α greater the efficiency of the unit is.

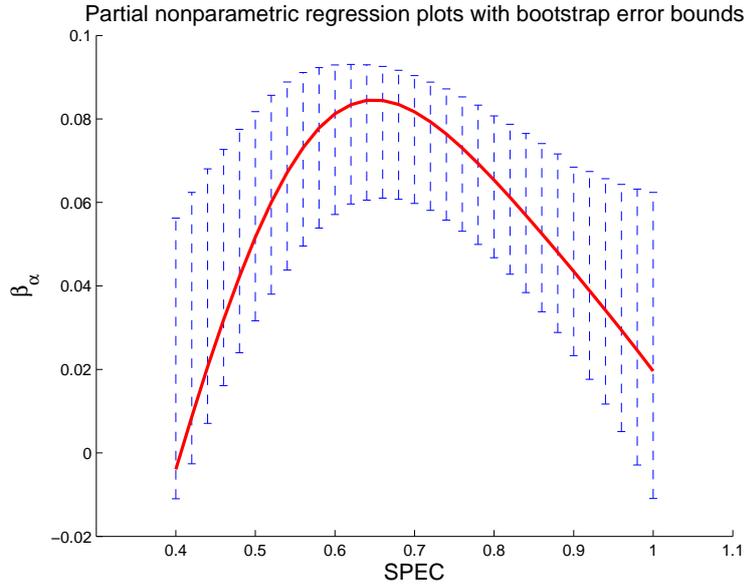


Figure 4: *Partial nonparametric regression of conditional efficiency as a function of disciplinary specialization (SPEC).*

Here below we report the results of the testing of scale and scope carried out for the full HUMB MODEL, with $B = 1000$ bootstrap loops. The test has been implemented by following the approach of Daraio and Simar (2014), analysing the impact of both external factors Z together and also each factor separately.

The obtained results confirm that SIZE and SPEC have a statistically significant impact both together and in isolation. Indeed we have:

scale and scope $Z = (Z_1, Z_2)$, p-value of $\tau = 0.036$, observed value of test statistics $\tau = 0.0245$, Bootstrap 95% percentile: 0.0228;

size $Z = Z_1$, p-value of $\tau = 0.000$, observed value of test statistics $\tau = 0.001619$, Bootstrap 95% percentile: 0.00079;

spec $Z = Z_2$, p-value of $\tau = 0.031$, observed value of test statistics $\tau = 0.02291$, Bootstrap 95% percentile: 0.02031.

6.2 Efficiency results and analysis of gaps

In this section we summarize the obtained results grouped by countries, together with the European average computed over the analysed sample to facilitate the interpretation. We

remind again that to read the results, the smaller the level of the efficiency greater the efficiency of the unit or group of units is.

Table 2 reports in the columns: Country, number of observations ($\#$ obs), number of dominating units ($\#$ dom), empirical estimates of the probability of being dominated (\widehat{H}_{XY}) and robust directional measure of efficiency conditioned to SIZE and SPEC, our Z variables ($\widehat{\beta}_{\alpha,XY|Z}$).

The last line of the Table shows the average at European level. An outline of the efficiency analysis results on the Humboldtian model could be obtained by comparing the average performance at national level with the European average. Countries that are performing much better than the European standard are UK, Sweden and Switzerland, followed by Belgium, Austria, Ireland and Netherlands, the others follow, this appears by inspecting the conditional efficiency scores ($\widehat{\beta}_{\alpha,XY|Z}$).

Table 2: Efficiency Results for Humb Model: averages by country.

Country	#obs	#dom	\widehat{H}_{XY}	$\widehat{\beta}_{\alpha,XY Z}$
AT	14	4.21	0.0105	0.040465
BE	4	2.75	0.0069	0.061991
CH	11	1.18	0.0029	0.008743
CZ	14	3.00	0.0075	0.042476
DE	71	10.55	0.0263	0.153871
ES	47	6.15	0.0153	0.097338
FI	12	1.75	0.0044	0.012852
HU	6	27.50	0.0686	0.209560
IE	10	2.40	0.0060	0.033448
IT	60	4.23	0.0106	0.064099
NL	13	3.46	0.0086	0.048254
NO	8	6.25	0.0156	0.115628
RO	14	1.86	0.0046	0.024394
SE	17	1.71	0.0043	0.014079
SK	4	2.25	0.0056	0.013651
UK	89	1.80	0.0045	0.027551
EU	400	4.96	0.0124	0.068804

Note: only countries with at least 3 observations are reported in the table.

The last line reports the average over the whole analyzed sample.

Table 3 reports the estimated gaps in percentage of the outputs produced by the units to reach the robustly estimated efficient frontier.

Table 3: Gaps in percentages for Humb Model: averages by country.

Country	#obs	#DEG5	#DEG6	#PUB	IC	Q1	NI	EXC
AT	14	0.13	0.11	0.03	0.03	0.04	0.04	0.04
BE	4	0.04	0.03	0.02	0.05	0.06	0.06	0.05
CH	11	0.02	0.00	0.00	0.01	0.01	0.01	0.01
CZ	14	0.06	0.06	0.11	0.06	0.08	0.07	0.13
DE	71	0.27	0.08	0.11	0.16	0.15	0.15	0.15
ES	47	0.12	0.17	0.14	0.11	0.10	0.12	0.13
FI	12	0.03	0.02	0.03	0.01	0.02	0.01	0.02
HU	6	0.25	0.30	0.26	0.21	0.23	0.32	0.26
IE	10	0.04	0.31	0.22	0.03	0.05	0.04	0.05
IT	60	0.14	0.23	0.08	0.08	0.06	0.07	0.07
NL	13	0.05	0.03	0.02	0.04	0.05	0.04	0.06
NO	8	0.20	1.98	0.24	0.11	0.12	0.11	0.15
RO	14	0.02	0.25	0.19	0.03	0.09	0.05	0.13
SE	17	0.03	0.01	0.01	0.01	0.01	0.01	0.01
SK	4	0.03	0.03	0.04	0.01	0.02	0.02	0.02
UK	89	0.02	0.06	0.06	0.03	0.03	0.03	0.04
EU	400	0.11	0.15	0.09	0.07	0.08	0.08	0.09

Note: only countries with at least 3 observations are reported in the table.
The last line reports the average over the whole analyzed sample.

Table 4 reports in the columns: Country, number of observations ($\#$ obs), number of dominating units ($\#$ dom), empirical estimates of the probability of being dominated (\hat{H}_{XY}) and robust directional measure of efficiency conditioned to SIZE and SPEC, our Z variables ($\hat{\beta}_{\alpha,XY|Z}$).

The last line of the Table shows the average at European level. An outline of the efficiency analysis results on the Research model could be obtained by comparing the average performance at national level with the European average. Overall, the results are similar to the ones obtained in the Humboldtian Model: Countries that are performing much better than the European standard are UK, Sweden and Switzerland, followed by Belgium, Austria, Ireland and Netherlands, the others follow, this appears by inspecting the conditional efficiency scores ($\hat{\beta}_{\alpha,XY|Z}$).

Table 4: Efficiency Results for Res Model: averages by country.

Country	#obs	#dom	\widehat{H}_{XY}	$\widehat{\beta}_{\alpha,XY Z}$
AT	14	4.21	0.0105	0.051877
BE	4	2.75	0.0069	0.061991
CH	11	1.18	0.0029	0.008743
CZ	13	3.15	0.0079	0.094881
DE	71	10.55	0.0263	0.163797
ES	47	6.15	0.0153	0.123091
FI	11	1.82	0.0045	0.027432
HU	6	27.50	0.0686	0.371831
IE	10	2.40	0.0060	0.048087
IT	60	4.23	0.0106	0.089005
NL	13	3.46	0.0086	0.059462
NO	8	6.25	0.0156	0.143286
RO	9	2.33	0.0058	0.037946
SE	16	1.75	0.0044	0.021866
SK	3	2.67	0.0067	0.043278
UK	86	1.83	0.0046	0.039575
EU	387	5.10	0.0127	0.090002

Note: only countries with at least 3 observations are reported in the table.
The last line reports the average over the whole analyzed sample.

Table 5 reports the estimated gaps in percentage of the outputs produced by the units to reach the robustly estimated efficient frontier.

Table 5: Gaps in percentages for Res Model: averages by country.

Country	#obs	#DEG5	#DEG6	#PUB	IC	Q1	NI	EXC
AT	14	0.00	0.17	0.04	0.04	0.05	0.05	0.06
BE	4	0.00	0.03	0.02	0.05	0.06	0.06	0.05
CH	11	0.00	0.00	0.00	0.01	0.01	0.01	0.01
CZ	13	0.00	0.10	0.21	0.12	0.18	0.15	0.27
DE	71	0.00	0.09	0.13	0.17	0.16	0.16	0.16
ES	47	0.00	0.22	0.18	0.14	0.13	0.15	0.18
FI	11	0.00	0.04	0.04	0.03	0.03	0.03	0.04
HU	6	0.00	0.58	0.62	0.38	0.41	0.57	0.49
IE	10	0.00	0.49	0.36	0.04	0.08	0.06	0.09
IT	60	0.00	0.42	0.14	0.12	0.08	0.10	0.10
NL	13	0.00	0.04	0.02	0.05	0.06	0.05	0.07
NO	8	0.00	4.01	0.40	0.14	0.15	0.14	0.20
RO	9	0.00	0.36	0.29	0.05	0.14	0.08	0.21
SE	16	0.00	0.04	0.04	0.02	0.02	0.02	0.02
SK	3	0.00	0.08	0.14	0.03	0.07	0.07	0.08
UK	86	0.00	0.08	0.08	0.05	0.04	0.04	0.05
EU	387	0.00	0.26	0.14	0.10	0.10	0.10	0.13

Note: only countries with at least 3 observations are reported in the table.

The last line reports the average over the whole analyzed sample.

Summing up, the inspection of average efficiency values per country shows large differences due to the national context. The interpretation of these differences will require a dedicated effort. A preliminary conjecture could be as follows. In order to make the best use of their inputs, universities should be put in the position to move in the multidimensional strategic space. This space includes inputs and outputs. Efficient universities are those that adjust their mix of inputs in order to achieve the best possible mix of outputs. It is clear that universities do not have full discretionary power over inputs and outputs, as our analysis has clearly recognised. However, national contexts may provide more or less strategic autonomy, that is, may support universities in their strategic positioning or may, on the contrary, create legal and administrative constraints. Supporting the autonomy of universities in strategic positioning is generally associated to two conditions. As for education, it requires that universities are in the position to match appropriately the profile of students to the teaching offering. While this may have different implications in different fields, there is a well known general problem that cuts across fields of education and countries, namely the role of professional education, also called vocational training. Some countries allocate vocational training to separate institutions, while others add to the general mission of univer-

sities. In the latter case universities have, in general, larger student loads and lower teaching efficiency, given the mismatch between the educational needs of students and the rigidity of the university offering. As for research, efficiency requires that government research funding is allocated according to criteria that gives a premium to research quality. This follows the adoption of evaluation exercises, or formula-based funding criteria based on research quality. Universities that are placed in an institutional context based on research quality funding develop over time strategies to improve their positioning in research. The confirmation of these conjectures is left to further empirical evidences.

7 Conclusions

In this paper we analysed the issue of scale and scope in European universities by applying state of the art directional distances techniques on an original database of European universities built by integrating input/output university data with bibliometric indicators.

Moreover, we improved over previous studies adopting a cross-country perspective, applying robust nonparametric estimators and testing for the significance of scale and scope effects by using the bootstrap.

Here we offer a suggested interpretation of the obtained results.

The large literature on economies of scale in higher education has proved elusive for decades. One problem, admitted by many authors, has been the limited generalizability of findings, which derive from single-country datasets. Yet a more tricky issue has been the lack of theoretical foundations for the expected results. Why should we expect variable, or constant, economies of scale in higher education?

Teaching is subject to increasing returns to scale from the increase in size of classes and to constant returns to scale in tutorial and examination activities. The overall effect depends on the mix between classroom and tutorial activity. This mix is variable across fields of education. In addition, since the increase in the size of classes deteriorates the quality of teaching, the overall effects also depends on the degree to which universities are hard pressed to meet the quality standards of education, as reflected for example in international guidelines. The pressure on quality is likely to be felt more in systems in which there is more mobility among students and thus more competition among universities. The returns to scale are likely to be different in the case of quality-adjusted output of education, or quality-unadjusted output.

On the other hand, research enjoys increasing returns to scale due to cognitive indivisibilities and experimental infrastructure but only up to a small size of research teams. The minimum efficient scale is exhausted at a small number of researchers. Beyond this threshold, decreasing returns to scale are expected, due to the loss of autonomy and increasing

complexity and bureaucratization in the management of research teams. Yet increasing returns could appear again not at the level of research teams but of the entire university, due to the efficiency in administrative activities and to visibility and reputation effects of the umbrella organization.

From this discussion, it is clear that the existence and magnitude of economies of scale depends on the combination between teaching and research and is influenced by the subject mix, the institutional context, the competition among universities, and the internal governance of universities. A similar situation is found for the theory of economies of scope. There are several possible definitions of economies of scope in university production, one of which refers to the joint production of teaching and research by the same academic staff. Here the overall findings from the literature show that the joint production is more efficient than the separate production, but also that, keeping the time budget fixed, there are also trade-offs (in particular, the increase in teaching load diminished research production).

First, universities are multi-input multi-output units, so that the appropriate theoretical framework is the theory of multi-product firms, not the conventional production function (Bonaccorsi and Daraio, 2003, 2004). What is needed is an estimation of the internal trade-offs between inputs and between outputs.

Second, due to the specific production technology, the substitution effects (between inputs and between outputs), as well as the complementarity effects, have intrinsically local nature. The question whether ‘the universities’ exhibit increasing or decreasing returns is meaningless. Universities exhibit local effects, depending on the volume of activity, as well as on the combination between different activities in the neighborhood of the point of estimation. Consequently an appropriate framework for the theory of university production is not the Cobb-Douglas production function, in which substitution effects depend on a fixed functional form, but the nonparametric and robust efficiency framework adopted here. In this framework there is not a functional form dictating the shape of substitution effects.

Third, the theory of university production is clearly still in its infant stage and need to be developed further. In this paper we have tried to make some step towards a better analytical framework and have offered new empirical results from a novel dataset, using state of the art techniques in efficiency analysis.

8 Appendix

Technical Details on gaps calculation when a factorization is used

It is well known that nonparametric efficiency analysis gain in precision when working in space with lower dimensions (this is the usual “curse of dimensionality” of nonparametric techniques, see e.g. Daraio and Simar, 2007, for a discussion). In the application reported in this paper, the original data are transformed before entering into the analysis, to reduce the dimension of the problem (by using input and/or output factors as defined in Daraio and Simar, 2007). In this case of course, once the gaps have been computed for the variables used in the analysis, the researcher is willing to evaluate the corresponding gaps in the original inputs and outputs. This can be done by transforming back the gaps in the factors into the original units. We briefly explain how to achieve this.

Suppose we are able to reduce the dimension among a selection of inputs, because they are highly correlated. Denote the corresponding matrix of selected inputs by \tilde{X} that has n rows (the observations) and \tilde{p} columns (we could follow exactly the same procedure for a subset of highly correlated outputs \tilde{Y}). In this method (see e.g. Daraio and Simar (2007); Härdle and Simar, 2012), the highly correlated \tilde{p} columns can be replaced, without much loss of information by a single new variable through a linear combination. The best linear combination is given by the eigenvalue of the matrix $\tilde{X}'\tilde{X}$ corresponding to its highest eigenvalues. We call this unique linear combination the “input-factor” $F_{\tilde{X}}$. The ratios of the largest eigenvalue over the sum of all the \tilde{p} eigenvalues allows us to appreciate the loss of information due to the reduction of dimension. In practice, this ratio should be large, say above 0.85, meaning that less than 15% of the total information shared by the \tilde{p} original inputs is conserved in this unique input-factor $F_{\tilde{X}}$. Note also that if the columns of \tilde{X} are in different units, we scale them by their standard deviations to obtain unit free variables more adapted to linear combinations. The formal steps of this dimension-reduction are summarized by the next few steps:

- [1] If needed, scale the columns of \tilde{X} : $\tilde{X}_s = \tilde{X} \text{diag}(1./s_{\tilde{x}})$, where $\text{diag}(\cdot)$ is a diagonal matrix, $./$ is the Hadamard element-wise division between the vector of ones and the vector $s_{\tilde{x}}$ which is the vector of the empirical standard deviations of the \tilde{p} columns of \tilde{X} .
- [2] The input factor is given by $F_{\tilde{X}} = \tilde{X}_s a_1$ where $a_1 \in \mathbb{R}^{\tilde{p}}$ is the eigenvector of $\tilde{X}_s' \tilde{X}_s$ corresponding to its largest eigenvalue λ_1 .
- [3] The percentage of inertia of this factor (percentage of information contained in the

factor) is given by $\lambda_1/(\lambda_1 + \dots + \lambda_{\tilde{p}})$. This percentage should be high enough to validate the procedure (say, above 80–85%).

So, in the analysis, the factor $F_{\tilde{X}}$ will act as a single observed input and will be combined with other inputs (or other input factors) and outputs (or other output factors) along the lines of the techniques developed above. The gaps obtained at the end are thus in the units of the factors $F_{\tilde{X}}$ used and not in the units of the original variable \tilde{X} . We know that the value of the input factor variable on the efficient frontier is $\widehat{F}_{\tilde{X}}^{\partial} = F_{\tilde{X}} + G_F$. It is easy to check that the coordinates of $F_{\tilde{X}}$ in the original units of \tilde{X}_s is given by $F_{\tilde{X}}a'_1$. For the same reason, the coordinates of the frontier point are $\widehat{F}_{\tilde{X}}^{\partial}a'_1$, so the measure of the gaps in the units of \tilde{X}_s are given by $G_{\tilde{X}_s} = G_F a'_1$. Of course we have also to rescale back this solution, if step [1] above has been used. Finally, an estimate of the gaps in the units of the original \tilde{p} input variables, for the n observations is given by:

$$G_{\tilde{X}} = G_{\tilde{X}_s} \text{diag}(s_{\tilde{x}}) = G_F a'_1 \text{diag}(s_{\tilde{x}}). \quad (9)$$

References

- [1] Badin, L., Daraio, C. and L. Simar (2012), How to measure the impact of environmental factors in a nonparametric production model, *European Journal of Operational Research*, 223, 818–833.
- [2] Badin L., Daraio C., Simar L. (2014), Explaining Inefficiency in Nonparametric Production Models: the State of the Art, *Annals of Operations Research*, 214, 5–30.
- [3] Baumol, W., Panzar J., Willig R. (1988), *Contestable Markets and The Theory of Industry Structure*.
- [4] Bonaccorsi A., Daraio C. (2003), A robust nonparametric approach to the analysis of scientific productivity, *Research Evaluation*, 12 (1), 47–69.
- [5] Bonaccorsi A., Daraio C. (2004), 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.
- [6] Bonaccorsi A., Daraio C. (2005), Exploring size and agglomeration effects on public research productivity, *Scientometrics*, Vol. 63, No. 1, 87-120.
- [7] Bonaccorsi A., Daraio C. (2007), eds., *Universities and Strategic Knowledge Creation. Specialization and Performance in Europe*, Edward Elgar Publisher, Cheltenham (UK).

- [8] Bonaccorsi A., Daraio C., Simar L. (2013), What is the impact of scale and specialization on the research efficiency of European universities?, in J. Gorraiz, E. Schiebel, C. Gumpenberger, M. Horlesberger and H. Moed eds., Proceedings of ISSI 2013, Vienna, Vol. 2, pp. 1817-1829 (ISBN: 978-3-200-03135-7).
- [9] Bonaccorsi A., Daraio C., Simar L. (2014), Scale and research specialization in European universities: a directional distance approach to teaching efficiency, in Bonaccorsi A. (ed.), *Knowledge, Diversity and Performance in European Higher Education*, Cheltenham, Edward Elgar, pp. 292-312.
- [10] Brandt T. and Schubert T. (2014), Is the University Model an Organizational Necessity? Scale and Agglomeration Effects in Science, in Bonaccorsi A. (ed.), *Knowledge, Diversity and Performance in European Higher Education*, Cheltenham, Edward Elgar, pp. 233-266.
- [11] 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.
- [12] Carayol, N., Matt, M. (2006), Individual and collective determinants of academic scientists' productivity, *Information Economics and Policy*, 18(1), 55–72.
- [13] Chambers, R. G., Y. Chung, and R. Färe (1996), Benefit and distance functions, *Journal of Economic Theory*, 70, 407–419.
- [14] Cohn, E., Cooper, S. T. (2004), Multi-product cost functions for universities: economies of scale and scope, in *International Handbook on the Economics of Education*, edited by Johnes G. and Johnes J., Edward Elgar Publishing, pp. 579-612.
- [15] Cohn, E., Rhine, S. L. W., Santos, M. C. (1989). Institutions of higher education as multi-product firms: Economies of scale and scope. *The Review of Economics and Statistics*, 71(2), 284-290.
- [16] Daghbashyan Z., Deiacio E., McKelvey M. (2014) How and why does cost efficiency of universities differ across European countries? An explorative attempt using new microdata. In Bonaccorsi A. (ed.) *Knowledge, Diversity and Performance in European Higher Education*. Cheltenham, Edward Elgar.
- [17] Daraio C., Bonaccorsi A. et al. (2011), The European University landscape: A micro characterization based on evidence from the Aquameth project, *Research Policy*, 40, 148-164.

- [18] Daraio, C. and L. Simar (2007), *Advanced Robust and Nonparametric Methods in Efficiency Analysis. Methodology and applications*, Springer, New York.
- [19] Daraio C. and Simar, L. (2014), Directional Distances and their Robust versions: Computational and Testing Issues, *European Journal of Operational Research*, 237, 358-369.
- [20] Dunder, H., Lewis, D. R. (1995). Departmental productivity in American Universities: Economies of scale and scope. *Economics of Education Review*, 14(2), 119-144.
- [21] Egghe, L., Rousseau, R. (1990), *Introduction to informetrics: Quantitative methods in library, documentation and information science*.
- [22] EUMIDA (2010), Feasibility Study for Creating a European University Data Collection, Contract No. RTD/C/C4/2009/0233402, European Commission.
- [23] Farrell, M.J. (1957), The measurement of productive efficiency, *Journal of the Royal Statistical Society*, A(120), 253–281.
- [24] Färe, R., S. Grosskopf and C. A. K. Lovell. (1994), *Production Frontiers*, Cambridge University Press.
- [25] Färe, R., S. Grosskopf and D. Margaritis (2008), Efficiency and Productivity: Malmquist and More, in *The Measurement of Productive Efficiency*, 2nd Edition, Harold Fried, C.A. Knox Lovell and Shelton Schmidt, editors, Oxford University Press.
- [26] Glass, J.C., Mckillop, D.G. and Hyndman, N.S. (1995a), The achievement of scale efficiency in UK universities: a multiple-input multiple-output analysis, *Education Economics*, 3, 249-263.
- [27] Glass, J.C., Mckillop, D.G. and Hyndman, N.S. (1995b), Efficiency in the provision of university teaching and research: an empirical analysis of UK universities, *Journal of Applied Econometrics*, 10, 61-72.
- [28] Hanushek, E. A. (1986), The Economics of Schooling: Production and E, *Journal of economic literature*, 24(3), 1141–1177.
- [29] Härdle, W. and L. Simar (2012), *Applied Multivariate Statistical Analysis*, Third Edition, 458p., Springer-Verlag, Berlin.
- [30] Hashimoto, K., Cohn, E. (1997). Economies of scale and scope in Japanese private universities. *Education Economics*, 5, 107-115.

- [31] Horta, H., Lacy, T. A. (2011), How does size matter for science? Exploring the effects of research unit size on academics' scientific productivity and information exchange behaviors, *Science and Public Policy*, 38(6), 449–460.
- [32] Izadi, H., Johnes, G., Oskrochi, R., Crouchley, R. (2002). Stochastic frontier estimation of a CES cost function: The case of higher education in Britain. *Economics of Education Review*, 21, 63-71.
- [33] Johnes, J (2004), Efficiency measurement, in *International Handbook on the Economics of Education*, edited by Johnes G. and Johnes J., Edward Elgar Publishing, pp. 613–742.
- [34] Johnes, J (2006), Data envelopment analysis and its application to the measurement of efficiency in higher education, *Economics of Education Review*, 25(3), 273–288.
- [35] Joumady, O., Ris, C. (2005), Performance in European higher education: A non-parametric production frontier approach, *Education Economics*, 13(2), 189–205
- [36] Koshal, R. K., Koshal, M. (1999). Economies of scale and scope in higher education: A case of comprehensive universities. *Economics of Education Review*, 18, 269-277.
- [37] Laband, D. N., Lentz, B. F. (2003). New estimates of economies of scale and scope in higher education. *Southern Economic Journal*, 70, 172-183.
- [38] Li, Q. and J.S. Racine (2007), *Nonparametric Econometrics: Theory and Practice*, Princeton University Press.
- [39] Longlong, H., Fengliang, L., Weifang, M. (2009). Multi-product total cost functions for higher education: The case of Chinese research universities. *Economics of Education Review*, 28, 505-511.
- [40] Lopez-Illescas C., de Moya-Anegon F., Moed H. F., (2011), A ranking of universities should account for differences in their disciplinary specialization, *Scientometrics*, 88, 563-574.
- [41] Moed H.F., de Moya-Anegon F., Lopez-Illescas C., Visser M. (2011), Is concentration of university research associated with better research performance?, *Journal of Informetrics*, 5 (4), 649-658.
- [42] Qureshi, S., Hlupic, V., de Vreede, G. J., Briggs, R. O., Nunamaker, J. (2003), Managing knowledge in a collaborative context: how may intellectual resources be harnessed towards joint effect?, *Knowledge and business process management*, 174–198.

- [43] Shephard, R.W. (1970), *Theory of Cost and Production Function*. Princeton University Press, Princeton, New-Jersey.
- [44] Schimank U., Winnes M. (2000) Beyond Humboldt? The relationship between teaching and research in European university systems. *Science and Public Policy*, vol. 27, pp. 398-408.
- [45] Seglen, P. O., Aksnes, D. W. (2000), Scientific productivity and group size: A bibliometric analysis of Norwegian microbiological research, *Scientometrics*, 49(1), 125–143.
- [46] Simar, L. and Vanhems A. (2012), Probabilistic Characterization of Directional Distances and their Robust Versions, *Journal of Econometrics*, 166, 342-354.
- [47] SIR (Scimago Institutional Ranking) World Report (2011).
- [48] 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.
- [49] Worthington, A.C., (2001), An empirical survey of frontier efficiency measurement techniques in education, *Education Economics*, 9 (3), 245-268.
- [50] Worthington, A. C., Higgs, H. (2011). Economies of scale and scope in Australian higher education. *Higher Education*, 61(4), 387-414.