Dynamics of productivity in higher education: cross-european evidence based on bootstrapped Malmquist indices

Aleksandra Parteka · Joanna Wolszczak-Derlacz

Abstract This study examines patterns of productivity change in a large set of 266 public higher education institutions (HEIs) in 7 European countries across the time period 2001–2005. We adopt consistent bootstrap estimation procedures to obtain confidence intervals for Malmquist indices of HEI productivity and their components. Consequently, we are able to assess the statistical significance of changes in HEI productivity, efficiency and technology. Our results suggest that, assessed vis-à-vis a common 'European' frontier, HEI productivity rose on average by 4 % annually. Statistically significant *changes* in productivity were registered in 90 % of observations on the institutions in our sample, but statistically significant annual *improvements* in overall productivity took place in only 56 % of cases. There are considerable national differences, with German, Italian and Swiss HEIs performing better in terms of productivity change than HEIs from the other countries examined.

Keywords Productivity · Efficiency · Higher education · Bootstrapped Malmquist index

JEL Classification C61 · D24 · I23

1 Introduction

In recent years, the higher education sector has been subject to formal quantitative research that has mainly covered such topics as the estimation of rates of return in higher education, the academic labour market, institutional behaviour, and higher education as an industry. Approaches to higher education institutions (HEIs) have been changing and, apart from recognizing their obvious role in human capital and knowledge creation, critical analysis concerning their productivity and efficiency has started to gain importance. In particular, due to changing demographic trends and competition for students, as well as a growing squeeze on public entities by financial constraints, public HEIs are under constant pressure to improve their performance. On top of this, competition between universities has been growing steadily, and European HEIs are still struggling to catch up with American institutions.

An examination of the existing literature leads us to conclude that there are several gaps in analyses of higher education productivity that need to be filled. First, major attention has thus far been focused on the analysis of productivity performance, in so far as this concerns evaluating the productivity *levels* of universities (among others: Glass et al. 1995; Johnes 2006a, b; Bonaccorsi et al. 2007). However, such efficiency scores, apart from providing a tool for comparing productivity between units (and, hence, serving as yet another university ranking system), say nothing about *changes* in productivity across time (and whether universities manage to improve their performance, stagnate or regress). Second, due to problems commonly associated with gathering comparable data for HEIs from multiple countries, such exercises (assessments of productivity changes over time) have usually been conducted with units from only one country (or, exceptionally, as in Agasisti and Johnes 2009 or Agasisti and Pérez-Esparrells 2010, two countries). In fact, Agasisti and Pérez-Esparrells (2010) state: 'Future research can extend this study. For instance, a wider comparison among universities from

A. Parteka (✉) · J. Wolszczak-Derlacz
Faculty of Management and Economics, Gdansk University of Technology, Narutowicza 11/12, 80-233 Gdańsk, Poland
e-mail: aparteka@zie.pg.gda.pl

different European countries could be useful for policy purposes' (p.102). From the policy perspective, a comparative cross-European analysis of HEI productivity is of major importance, especially in the light of the integration of European higher education systems under the Bologna process.

Finally, certain higher education studies assessing productivity changes over time (Flegg et al. 2004; Johnes 2008; Worthington and Lee 2008; Agasisti and Johnes 2009; Agasisti and Pérez-Esparrells 2010) have adopted techniques based on Malmquist indices that have not been statistically verified. In other words, these authors simply state that productivity (efficiency, technology—if the indices are decomposed) in selected HEIs has increased or decreased, but no formal tool has been applied to check whether the estimates are sensitive to random variations in the data. Traditional Malmquist methodology, based on estimations of distance measures made through data envelopment analysis (DEA), a non-stochastic procedure, does not provide any insight into the statistical significance of its results. There are, however, tools based on resampling (bootstrap) methods that allow us to correct this weakness.

Hence, the considerable limits of the existing literature stem from the facts that (i) little is known about productivity changes across universities from several countries analyzed within a common methodological framework, and (ii) methodological issues concerning the significance of the results obtained with Malmquist indices have not been appropriately addressed.

A very particular feature of our dataset is its panel dimension, which allows us to go beyond studies that only compare efficiency scores across units of higher education (usually from just one country). We have managed to gather comparable statistics concerning the inputs and outputs of 266 public HEIs from seven European countries (namely: Austria, Finland, Germany, Italy, Poland, the United Kingdom, and Switzerland) over the time period 2001–2005. Moreover, the bootstrap estimation procedure adopted (Simar and Wilson 1999) corrects the basic and possibly biased information given by Malmquist indices of productivity, providing us with confidence intervals for these indices and their components. As a result, we have a tool to verify whether changes in the productivity of European HEIs, as indicated by Malmquist indices, are significant in a statistical sense (i.e. whether the result indicates a *real* change in the productivity of a given HEI or is just an outcome of sampling noise). Thus, by focusing on the two limits described above and using an original and vast set of microdata on HEIs from several European countries in conjunction with a consistent bootstrap methodology, this study presents an important extension of the existing literature.

The remainder of the paper is organized as follows: we devote Sect. 2 to a presentation of the methodology applied in our analysis (in particular, describing ways of assessing the statistical significance of Malmquist indices) and a concise description of the studies most closely-related to our research. In Sect. 3, we first present our data and then show the statistically significant results of a cross-European assessment of productivity, efficiency and technology changes in 266 HEIs. Conclusions follow.

2 Theoretical and empirical background

2.1 Changes in productivity across time—Malmquist indices and their statistical significance

Higher education institutions are not classical firms whose aim is profit maximization; public HEIs, in particular, are by definition, non-profit organizations. Hence, we cannot assess their productivity by using the methods typically applied to the evaluation of companies producing goods or services and generating profit. Moreover, the functioning of HEIs is characterized by interplay between multiple inputs and outputs. Universities use such inputs as human resources (staff), students and financial resources and 'produce' at least two outputs, reflecting both their teaching and research missions.[1] Consequently, analysis of HEI productivity dynamics must take these features into account. Tools based on DEA have proven very useful in capturing multiple inputs and outputs at the same time and focusing on a non-parametric treatment of efficiency frontiers. We focus on changes in the productivity of European public HEIs where productivity is understood not in absolute terms, but as performance that is relative to the efficiency of technologies (represented by a frontier function). The aim is not to identify levels of productivity as previous studies have (e.g. Glass et al. 1995; Johnes 2006a, b), but to study the dynamics of productivity. Thus, below, we do not focus on the formal derivation of DEA relative productivity scores,[2] but we show the methods applied to assess *changes* in productivity in the higher education sector.

To measure productivity change between two periods of time, we adopt the output-based Malmquist index of productivity developed by Färe et al. (1992, 1994, 1997), itself drawn from the measurements of efficiency in Farrell

---

[1] Additionally, in the light of the 'triple helix' approach, a so-called 'third mission' (the engagement of universities in entrepreneurship and business-related activities) could be considered. However, this is hardly measurable (especially in a large panel of units such as that used in our case), and consequently we restrict our approach to research and teaching missions only.

[2] For details on DEA measurement, see Coelli et al. (2005).

(1957) and of productivity in Caves et al. (1982). The output-oriented model aims to maximize output while using no more than the number of inputs observed.[3] Hence, the question to be answered is: by how much can output quantities be proportionally augmented without changing input quantities? In the context of HEI efficiency, output-oriented models are usually used because the quantity and quality of inputs, such as student entrants, are assumed to be fixed exogenously and universities can hardly influence their number or characteristics, at least in the short term. We compute Malmquist indices[4] that are based on DEA scores, allowing us to measure the total factor productivity (TFP)[5] change of single HEIs between two data points:

$$TFP_{i,(t,t+1)} = m_{i,(t,t+1)} = m_i(x_{t+1}, y_{t+1}; x_t, y_t)$$
$$= \left[\frac{d_i^t(x_{t+1}, y_{t+1})}{d_i^t(x_t, y_t)} \times \frac{d_i^{t+1}(x_{t+1}, y_{t+1})}{d_i^{t+1}(x_t, y_t)} \right]^{1/2}, \quad (1)$$

where $i = 1,...,N$ denotes the DMU[6] (in our case HEI) being evaluated, x refers to inputs and y to outputs, and m is the productivity of the most recent production point defined by inputs and outputs (x_{t+1}, y_{t+1}) using period $t + 1$ technology, relative to the earlier production point (x_t, y_t) using period t technology.[7] Output distance functions are denoted as d.[8] With regard to output orientation, a value of $m_{i,(t,t+1)}$ greater than one indicates positive TFP growth in HEI i from period t to period $t + 1$, while $m_{i,(t,t+1)}$ smaller than one indicates TFP decline. For example, $m_{i,(2002,2003)} = 1.14$ would signify an improvement in TFP of HEI i between the years 2002 and 2003 of 14 %. If $m_{i,(t,t+1)}$ equals unity, then

no improvement in the TFP of HEI i was observed between the two data points.

In order to distinguish between two basic mechanisms provoking TFP growth, we adopt the Malmquist decomposition proposed by Färe et al. (1992):

$$m_{i,(t,t+1)} = \underbrace{\frac{d_i^{t+1}(x_{t+1}, y_{t+1})}{d_i^t(x_t, y_t)}}_{technical\ efficiency\ change(\varepsilon_i)}$$
$$\times \underbrace{\left[\frac{d_i^t(x_{t+1}, y_{t+1})}{d_i^{t+1}(x_{t+1}, y_{t+1})} \times \frac{d_i^t(x_t, y_t)}{d_i^{t+1}(x_t, y_t)} \right]^{1/2}}_{technological\ change(\tau_i)} \quad (2)$$

where *technical efficiency change*[9] (ε) reflects changes in the relative efficiency of a unit i (e.g. universities getting closer to or further away from the efficiency frontier), while *technological change* (τ) measures the shift in the production frontier itself and reflects effects that concern the higher education system as a whole. Values of $\varepsilon_{i,(t,t+1)}$ greater (lower) than unity indicate improvements (decreases) in technical efficiency between t and $t + 1$. Similarly, values of $\tau_{i,(t,t+1)}$ greater (lower) than unity indicate technological progress (regress) between t and $t + 1$. The value of m will be equal to 1 if the net effect of changes in technical efficiency and frontier changes is null.

The problem with the approach described above is that the frontier needed for the calculation of distance functions is estimated from the data, and thus the resulting changes in m may simply be the result of sampling noise. Hence, we adopt a particular way of measuring productivity changes: we follow a bootstrap procedure to obtain bias-corrected estimates of Malmquist indices (and their components—as in Eq. 2) and their confidence intervals (Simar and Wilson, 1999). This procedure is based on bootstrap DEA analysis[10] (relying on replication of the data-generating process) and allows us to: (i) verify whether correction for the bias in non-parametric distance function estimates (and thus in Malmquist index estimates) is desirable, and (ii) check whether the changes in productivity indicated by

---

[3] In contrast, the objective of the input-oriented model is to minimize inputs while producing, at least, given output levels.

[4] The Malmquist indices and their decomposition in our paper were computed using the FEAR software package for frontier analysis with R (Wilson 2008).

[5] Färe et al. (1992) assume that production technology exhibits constant returns to scale, which implies that the Malmquist index can be interpreted as an index of total factor productivity. Allowing for variable returns to scale (convex hull or free-disposal hull) means that the solutions to programming problems can be unattainable for some observations and, in addition, that in such cases, the Malmquist index cannot be interpreted as an indicator of TFP.

[6] Decision making unit (the expression commonly used in DEA analysis). In our case, each HEI is a DMU.

[7] Here, Malmquist index m is defined as the geometric mean of two indices: the first, with period t, being the reference technology; the second, with period $t + 1$, being the reference technology. These two indices are equivalent only if the technology is Hicks output neutral (Coelli, et al. 2005, p. 291). The geometric mean is used to avoid an arbitrary choice of the technologies from period t or $t + 1$ as a reference.

[8] The values of distance functions that appear in the Malmquist index are unobserved and must be estimated from the data. Due to space limits, we do not discuss all the steps concerning the derivation of distance measures. For a concise description of the formal procedure, see Coelli et al. (2005), pp. 291–294.

[9] The 'Technical efficiency change' ε can be further decomposed into 'scale efficiency change' and change in 'pure efficiency' (Färe et al. 1994). The results are available from the authors upon request.

[10] Bootstrapping was developed by Efron (1982) and Efron and Tibshirani (1993) for cases where little or nothing is known about the underlying data generating process for a sample of observations. The data generating process can be estimated empirically by resampling the original data series to generate a set of bootstrap pseudosamples and then applying the original estimators to these pseudosamples. Bootstrapped DEA was introduced by Ferrier and Hirschberg (1997) and Simar and Wilson (1998, 2000), who demonstrated how to construct confidence intervals for DEA efficiency scores in order to overcome the main weakness of basic DEA analysis—namely, the sensitivity of the results to the sample composition.

Malmquist indices and their components are statistically significant.[11]

In line with Simar and Wilson (1999), we first compute a set of bootstrap estimates for the Malmquist index for each HEI i: $\{\hat{m}(b)_{i,(t,t+1)}\}$ for $b = 1,...B$ (where B is the total number of replications performed with pseudosamples drawn from the 'original' dataset). Then, the bootstrap bias estimate for the 'original' (non-bootstrapped) estimator is calculated as:

$$bias(\hat{m}_{i,(t,t+1)}) = 1/B \sum_{b=1}^{B} \hat{m}(b)_{i,(t,t+1)} - \hat{m}_{i,(t,t+1)}, \quad (3)$$

where the first component is the average value of the bootstrap estimates of the Malmquist index, $Avg[\hat{m}(b)] = 1/B \sum_{b=1}^{B} \hat{m}(b)_{i,(t,t+1)}$, and the second component, $\hat{m}_{i,(t,t+1)}$, denotes the 'original' (non-bootstrapped) estimates of the Malmquist index (as in Eq. 1). In the next step, for each HEI i, we compute a bias-corrected estimate of the Malmquist index, $\hat{m}\_corr_{i,(t,t+1)}$—the difference between $\hat{m}_{i,(t,t+1)}$ and $bias(\hat{m}_{i,(t,t+1)})$, which, using (3), can be expressed as:

$$\hat{m}\_corr_{i,(t,t+1)} = 2\hat{m}_{i,(t.t+1)} - 1/B \sum_{b=1}^{B} \hat{m}(b)_{i,(t,t+1)} \quad (4)$$

The choice between the 'original' estimate of the Malmquist index and its bias-corrected version is based on a comparison of the mean square errors (MSEs) of the two indices, as it is plausible that the latter may have a higher MSE (Efron and Tibshirani 1993).[12]

Finally, in order to assess whether productivity change is meaningful in the statistical sense, the (1-α) percent confidence interval is obtained with the bootstrapping procedure as:

$$\hat{m}_{i,(t+1)} + l\_\hat{m}_\alpha(b) \leq m_{i,(t,t+1)} \leq \hat{m}_{i,(t,t+1)} + u\_\hat{m}_\alpha(b). \quad (5)$$

The $l\_\hat{m}_\alpha$ and $u\_\hat{m}_\alpha$ estimated respectively define the lower and upper bootstrap estimates of the confidence interval bounds for the Malmquist index, and α (e.g. 10, 5 or 1 %) characterizes the size of the interval. Following Simar and Wilson (1999), the Malmquist index estimated is said to be significantly different from unity (and so the productivity change is statistically significant) if the interval defined in Eq. 5 does *not* include unity.

An analogous approach applies for all the components of the Malmquist index (ε and τ), so that we also obtain bias-corrected estimates of ε and τ: $\hat{\varepsilon}\_corr_{i,(t,t+1)}$ and $\hat{\tau}\_corr_{i,(t,t+1)}$, as well as confidence intervals for ε and τ, allowing us to verify their statistical significance.

2.2 Related empirical evidence in the context of higher education

So far, probably due to problems with obtaining multi-period micro-level data on the performance of single universities, few authors have applied Malmquist indices to HEIs, usually preferring to focus on institutions from one country. A multi-country setting demands the computation of an index that requires the same set of inputs and outputs for all HEIs from the sample and, additionally, the presence of the same units and variables across time; unbalanced panels with changing sets of HEIs or inputs/outputs are not allowed.

Flegg et al. (2004) apply the Malmquist approach to a sample of 45 British universities for the period 1980/1981–1992/1993. Their results show that in these years *TFP* increased by 51.5 % but that most of this rise was caused by an outward shift of the efficiency frontier (technological change) and not by the movement of universities towards the frontier (efficiency change). Johnes (2008) derives Malmquist indices for 112 English HEIs over the period 1996/1997–2004/2005 and finds an average increase in *TFP* of around 1 % per year (decomposition shows that average annual technological change was equal to approximately 6 %, but a decrease in efficiency of 5 % per year took place). Worthington and Lee (2008) analyze 35 Australian universities (1998–2003) and find an average increase in productivity growth of circa 3 %, largely due to technological progress and not technical efficiency change. All in all, the existing evidence, based on British and Australian experience, suggests a predominant role for technological change, rather than efficiency change, in provoking overall *TFP* growth in HEIs.

It should be noted, however, that the HEIs from the countries analyzed so far were characterized by high levels of efficiency (high DEA efficiency scores) at the outset.

---

[11] Bootstrap methods can also be applied in the context of the so-called 'two-stage' DEA procedure, where in the second stage estimated efficiency measures are regressed on some environmental variables. Simar and Wilson (2007) define a statistical model where truncated regression yields consistent estimates and develop a bootstrap approach as a valid inference in the second-stage regression. As demonstrated in Simar and Wilson (2011), bootstrap methods, in contrast with second-stage OLS estimates, actually provide feasible means for inference in the second stage. In the higher education context, the two-stage approach performed with the use of bootstrapped truncated regression as in Simar and Wilson (2007) is adopted by Wolszczak-Derlacz and Parteka (2011).

[12] Given the sample variance $s^2(b)_i$ of the bootstrap values $\{\hat{m}(b)_{i,(t,t+1)}\}$ for $b = 1,...B$, and assuming that the estimated MSE of $\hat{m}\_corr_{i,(t,t+1)}$ is $4s^2(b)_i$ (Simar and Wilson 1999: 463), it can be shown that the 'original' estimate $\hat{m}_{i,(t,t+1)}$, rather than the bias-corrected estimate $\hat{m}\_corr_{i,(t,t+1)}$, should be used if $s^2(b)_i > \frac{1}{3}bias(\hat{m}_{i,(t,t+1)})$.

There are only two published papers (that we are aware of) comparing *changes* in the productivity and efficiency of HEIs from more than one country: Agasisti and Johnes (2009) and Agasisti and Pérez-Esparrells (2010).

Agasisti and Johnes (2009) employ Malmquist indices to analyze 127 English and 57 Italian public universities over the short period 2002/2003–2004/2005. In line with the findings of the abovementioned authors, their results confirm that English HEIs did not realize gains in technical efficiency, but rather registered changes in productivity that were due to frontier shifts. On the contrary, Italian HEIs—typically less efficient at the outset than English ones—became more technically efficient with respect to the frontier. This is an important result, suggesting that HEIs from countries further away from the common 'European' higher education efficiency frontier can experience 'catching-up' effects, while those which are already highly efficient move the frontier itself up.

Agasisti and Pérez-Esparrells (2010) adopt a similar setting, counting (apart from DEA scores) Malmquist indices for 57 Italian public institutions and 46 Spanish ones, again for a relatively short time span (the academic years 2004/2005 and 2000/2001). They find that Italian universities experienced important improvements in productivity, mainly due to improvements in 'technology' (the authors argue that the change resulted from important reforms in the curriculum organization of the Italian system of higher education), while Spanish universities registered much lower improvements in overall productivity, as a result of changes in efficiency.

However, despite the great advantages of cross-country evidence, none of these papers assess the statistical significance of their results. Consequently, we cannot exclude the possibility of bias caused by sample noise.

Bootstrapped DEA techniques have been used in economic analyses of productivity levels in many different sectors, including higher education (e.g. Johnes 2006a, b). On the contrary, the application of bootstrapped Malmquist methods to the analysis of productivity *change* has in general been less frequent,[13] and it should be noted in particular that none of the papers (that we are aware of) have used a consistent bootstrap methodology for the computation of Malmquist indices in the context of the higher education sector. Hence, the 'original' estimates of the distance functions and Malmquist indices of the universities analyzed so far have not been corrected for finite-sample bias, and what remains their main weakness is that

their statistical significance is unknown. In this paper, we address these issues.

3 Empirical evidence on productivity changes in European HEIs

3.1 Data and panel composition

Our analysis draws on a university-level database containing information on the outputs and inputs of 266 public HEIs from a set of European Union (Austria, Finland, Germany, Italy, Poland and the UK) and non-EU (Switzerland) countries for which it was possible to gather comparable micro data. We draw on a balanced panel containing statistics for single European HEIs for the years 2001–2005.[14] Even though the data comes from numerous sources, particular attention has been given to ensuring the maximum level of comparability of the crucial variables across countries in accordance with the Frascati manual (OECD 2002)—for details, see the data appendix (Table 6).

Table 7 in the Appendix contains information on the number of HEIs from each country (due to space limits a detailed list of all the universities covered by our study is available upon request). To the best of our knowledge, this is the most comprehensive balanced panel micro dataset on European HEIs from several countries that has been used for *Malmquist* analysis of productivity change.[15] Moreover, so far, advanced analysis of productivity trends in universities from new EU member states has been ignored. In contrast, along with universities from six western European countries, we also included in our analysis HEIs from Poland.[16]

Our dataset only contains public HEIs, because several statistics, the crucial ones concerning funding, are often not available for private HEIs. Additionally, we decided to concentrate only on the university sector; regarding the binary higher education system, we excluded from our

---

[13] Bootstrapped Malmquist indices have been used to study productivity changes in cases of, *inter alia*, the farming sector (Balcome and Davidova 2008) and the banking (Assaf et al. 2010), insurance (Mahlberg and Url 2010) and airline industries (Assaf 2011).

[14] Data on HEIs from several countries are available for more years (e.g. 1995–2008 for Poland), but for the computation of Malmquist indices based on a frontier common for all countries we need to have the same set of units across time. For example, the necessary data on Italian HEIs are not available prior to the year 2001.

[15] The Eumida project (see Daraio et al. 2011 for details) collected data on 488 universities from 11 European countries: Finland, France, Germany, Hungary, Italy, the Netherlands, Norway, Portugal, Spain, Switzerland and the UK. However, to the best of our knowledge, its panel dataset was not balanced and, more importantly, from our point of view, the study of productivity changes using consistent bootstrapped Malmquist methodology was not performed with Eumida statistics. Its data is not publicly available.

[16] For a study on scientific productivity of Polish HEIs compared versus HEIs from other more developed European countries and based on a similar dataset as the one used in the present study, see Wolszczak-Derlacz and Parteka (2010).

sample applied science institutes/schools (such as German or Austrian *fachhohschule* and applied science HEIs in Finland and Switzerland), which were only marginally conducive to research. Moreover, we also excluded from our analysis special purpose units specializing in one discipline only (e.g. medicine, arts, sports) and distance learning universities, as these were not considered comparable with 'traditional' universities. Finally, units whose publication records (used as a measure of one of the outputs) were scant, incomplete or identified via ambiguous affiliations[17] were not taken into consideration.

The calculation of Malmquist indices required the estimation of distance functions. We first used a bootstrapped DEA method based on annual observations of 266 European HEIs, which produced two outputs from three inputs. Given the double mission of HEIs (teaching and research)[18] as outputs, we considered teaching output (measured in terms of graduates), as well as research output (quantified by means of bibliometric indicators and based on an analysis of publication records, as in, among others, Creamer 1999; Dundar and Lewis 1998). While comparison of the number of graduates (total, without distinguishing between various types of studies) across HEIs was quite straightforward,[19] a challenge was posed by the necessary cross-country comparability of research outputs. Different countries adopt specific measures of research production (such as research funds, publication records, patents and applications). However, we relied on the uniform bibliometric data from Thomson Reuters' ISI Web of Science database (a part of the ISI Web of Knowledge[20]), which lists publications from quality journals (with a positive impact factor) in the majority of scientific fields.[21] We counted all publications (scientific articles, proceedings papers, meeting abstracts, reviews, letters, notes) published in a given year, with the requirement that at least one author declared an institutional affiliation with an HEI.[22]

Concerning input measures, our dataset contained information on numbers of students, total academic staff and total real revenues. Revenues were converted from national currency units into Euro PPS[23] (using exchange rates from Eurostat), to account for cross-country differences in price level and the purchasing power of the money that HEIs dispose of.

As for data sources (Table 8 in the Appendix), the availability and coverage of university-level data differed from country to country. The most comprehensive databases concerning HEIs exist in Finland, the UK and Italy, with freely-available online platforms giving access to a broad range of statistics that are not confidential. For Swiss, Austrian and German HEIs, data was kindly provided by the staff of each country's central statistical office. In the case of Poland, unfortunately, micro-data on HEIs (even public ones) practically does not exist for research purposes. There is no on-line platform containing such data, and only a few statistics are available in paper versions of publications issued by the Polish Ministry of Science and Higher Education (MNiSW) and the Polish Central Statistical Office (GUS); part of the data used were obtained through direct contact with the statistical offices possessing them.[24]

Our benchmark Malmquist analysis is based on DEA performed with three inputs and two outputs, where DMUs are compared with respect to the common European frontier. As a robustness check, we consider alternative formulations of DEA specification: a two input-two output version of the DEA model (without students as an input)[25] and estimates based on the use of average values of inputs and outputs.[26] Finally, to check for cross country heterogeneity, we perform an additional analysis where country-specific frontiers are estimated and productivity change is estimated with respect to units from the same country.

3.2 Malmquist indices: results for European HEIs

3.2.1 Benchmark results

In benchmark estimation we considered productivity change with respect to a common frontier, thus all 266 HEIs were treated jointly, and the frontier was estimated using annual information on the whole sample of European universities. Consequently, changes in productivity were relative to the

---

[17] For example, we excluded from our analysis the University of London, which, as a confederal organization, is composed of several colleges. It was not possible to identify publication records for the University of London because we could not be sure whether the university's academic staff gave the names of their colleges or 'University of London' as their affiliation.

[18] 'Third mission' was not considered due to the methodological problems linked to its measurement and lack of relevant data.

[19] See data appendix—Table 6.

[20] www.apps.isiknowledge.com.

[21] Web of Science covers nearly 12,000 international and regional journals and book series in every area of the natural sciences, social sciences and arts and humanities. For example, in 2009, it covered over 110,000 conference proceedings. Alternative sources, such as Scopus, could have been used, but we had access to Thomson Reuters' services only.

[22] Note that papers co-authored by persons affiliated to the same institution were only counted once.

---

[23] Purchasing power standard.

[24] Detailed information is available from the authors upon request.

[25] Such a reduction in the number of inputs is due to possible correlation between students and other inputs. We thank an anonymous referee for pointing this out.

[26] Present outputs can be dependent not only on present inputs but also on their past values. We thank an anonymous referee for pointing this out.

Table 1 Benchmark results—trends in productivity (m), efficiency (ε) and technology (τ) in 266 European HEIs based on annual changes for 2001–2005 period, CRS

	$TFP = \hat{m}$	$\hat{\varepsilon}$	$\hat{\tau}$
Number of all indices	$N(\hat{m}) = 1,064$	$N(\hat{\varepsilon}) = 1,064$	$N(\hat{\tau}) = 1,064$
Average value of all indices	$\bar{\bar{m}} = 1.041$	$\bar{\bar{\varepsilon}} = 1.032$	$\bar{\bar{\tau}} = 1.012$
Number of statistically significant indices	$N(\hat{m}**) = 963$	$N(\hat{\varepsilon}**) = 540$	$N(\hat{\tau}**) = 284$
Average value of statistically significant indices	$\bar{\bar{m}}** = 1.045$	$\bar{\bar{\varepsilon}}** = 1.057$	$\bar{\bar{\tau}}** = 1.046$
Number of statistically significant improvements	$N(\hat{m}** > 1) = 602$	$N(\hat{\varepsilon}** > 1) = 337$	$N(\hat{\tau}** > 1) = 177$
Percentage (out of 1,064) of HEIs registering statistically significant annual improvements	$\%(\hat{m}** > 1) = 56\%$	$\%(\hat{\varepsilon}** > 1) = 32\%$	$\%(\hat{\tau}** > 1) = 17\%$

** Refer to significance at 5 % level. CRS—constant returns to scale. Results based on three input—two output model

Source: own elaboration

European efficiency frontier in public higher education (relative in the sense that they were computed with reference to other universities from the group). Later on, we take into account cross-country specificity (see Sect. 3.3).

We first calculated 'original' (not bootstrapped) estimates of Malmquist indices (and their components). Then, we applied the bootstrap method described above (maintaining the assumption of constant returns to scale and output orientation), setting the number of bootstrap replications $B = 2,000$. We compared the MSEs of bias-corrected and 'original' (non-bootstrapped) estimates of Malmquist indices, finding that in the vast majority of cases, bias correction increased MSE (for details, see Table 9 in the Appendix). Simar and Wilson (1999) obtained analogous results. Consequently, and like the aforementioned authors, we do not report bias-corrected estimates, but rely on 'original' estimates of m (TFP), ε and τ that are based on decomposition (2): $\hat{m}$, $\hat{\varepsilon}$ and $\hat{\tau}$. In Table 10 we show summary statistics of the variables used in the DEA model, while summary statistics of both the 'original' and bias-corrected estimates of the indices are reported in Table 11 in the Appendix, where it can be seen that the difference between the two is negligible (the coefficients of correlation between the 'original' and bias-corrected series are between 0.97 for τ and 0.99 for m). However, we do refer to the estimated bootstrap confidence intervals to assess whether changes in productivity, efficiency and technology are meaningful in a statistical sense. The full set of results for all HEIs is obtainable upon request; here, we present the key findings.

In Table 1, we compare all the results (N = 1,064) with the statistically significant ones (at a significance level of 5 %).[27] In particular, we show the number of cases in our

panel in which estimates of Malmquist indices were significantly different from unity, $N(\hat{m}**)$, and their average value ($\bar{m}**$), comparing them with the average value of all the indices ($\hat{m}$). Finally, we report the number of cases with statistically significant increases in TFP, $N(\hat{m}** > 1)$, and the percentage of cases in which statistically significant annual improvements in productivity were registered. The same exercise has been done with estimates of ε and τ.

The calculation of confidence intervals permits us to note that at a standard 5 % level of significance, out of the 1,064 annual estimates of TFP growth between the years 2001 and 2005, 963 were statistically different from unity. Thus, in 90 % of the HEIs in our sample statistically significant *changes* in productivity were registered. Taking into account only statistically significant estimates of m, between the years 2001 and 2005, on average, HEIs in our sample registered an increase in productivity of around 4.5 % annually (the average value of all Malmquist indices, significant and not, equals 4.1 %). Counting cases in which m was significant and greater than one, denoted in Table 2 as $\%(\hat{m}** > 1)$, we can conclude that statistically significant annual *improvements* in overall productivity took place in 56 % of cases.

Comparing statistically significant estimates ε and τ, the two basic components of m, average efficiency improved by 5.7 %, while technology shifted up by 4.6 %.[28] If we considered all the estimates, these values would be lower (3.2 % and 1.2 %, respectively). Hence, accounting for statistical significance matters for the conclusions drawn. Looking at the number of cases with significant *improvements* in efficiency and technology, it is evident that

---

[27] Results obtained with alternative levels of significance, 1 and 10 % are obtainable upon request. The choice of significance level matters for the number (and percentage of the total) of significant Malmquist and efficiency indices (clearly, a bigger α results in a wider confidence interval) but not for their average value. In the case of the estimates of technological change, the range of the average

Footnote 27 continued
values of significant indices is larger (from 3.4 % when $\alpha = 10$ to 7.3 % when $\alpha = 1$ %).

[28] Note that even though, given decomposition (2), m should be a product of ε and τ, this need not necessarily be the case when we take into account only statistically significant changes in m, ε and τ (e.g. a given HEI can register a significant change in overall productivity and efficiency, but not a significant change in technology).

Table 2 Comparison of benchmark and alternative estimates of productivity (*m*), efficiency (*ε*) and technology (*τ*) change in 266 European HEIs (based on annual changes for 2001–2005 period), CRS

DEA	Corresponding time period	Average value of all indices		
		$\bar{\bar{m}}$	$\bar{\bar{\varepsilon}}$	$\bar{\bar{\tau}}$
Three input—two output	Annual change between 2001 and 2005	1.041	1.032	1.012
Two input—two output	Annual change between 2001 and 2005	1.039	1.038	1.004
Three input—two output: moving averages	Change between 2-year moving windows	1.040	1.029	1.01

DEA	Corresponding time period	Average value of statistically significant indices		
		$\bar{\bar{m}}**$	$\bar{\bar{\varepsilon}}**$	$\bar{\bar{\tau}}**$
Three input—two output	Annual change between 2001 and 2005	1.045	1.057	1.046
Two input—two output	Annual change between 2001 and 2005	1.041	1.07	1.02
Three input—two output: moving averages	Change between 2-year moving windows	1.047	1.066	1.045

** Refer to significance at 5 % level. CRS—constant returns to scale

Source: own elaboration

efficiency change *ε* (being a 'catching up effect' towards/away from the frontier) was more common than change in technology *τ* ('frontier shift effect'). From Table 1, it emerges that of the 1,064 observations analyzed for the period 2001–2005, efficiency change was significantly higher than unity in 32 % of cases (so approximately one-third of HEIs managed to catch up towards the European efficiency frontier) while significant technological improvement took place in only 17 % of HEIs.

3.2.2 Robustness checks and extensions of the basic model

In order to check the robustness of our findings, we ask whether the way the productivity frontier was defined in the DEA estimation matters to the conclusions drawn, so we consider alternative DEA model formulations with modified sets of inputs and outputs.

Firstly, we consider a DEA model with a restricted number of two inputs (total staff, total revenues) and two outputs (teaching output—graduates, and research output—publications). Such a formulation addresses the difficulty in modelling the students-graduates productivity relationship[29] and corrects for any correlation between students and other inputs (such as teaching staff and funding).

Secondly, we perform a Malmquist analysis based on a DEA model with input and output data expressed as time averages.[30] Such an exercise permits us to correct for any random time variation in the data, as well as a possible relationship between past inputs and present outputs. We consider a DEA model with three inputs and two outputs as in the benchmark estimation, but based on moving averages of all inputs and outputs (2 year moving averages: t1 = 2001–2002, t2 = 2002–2003, t3 = 2003–2004, t4 = 2004–2005). Then, we obtain Malmquist indices based on this average data, which reflect productivity changes between periods: t1/t2, t2/t3, and t3/t4.

The results concerning *TFP* growth in European HEIs obtained with alternative DEA formulations are actually very similar to the benchmark ones (we compare them in Table 2) and the correlations between the estimates obtained with different models are fairly high.[31] The estimated annual *TFP* change indicated by the Malmquist index at most deviates from the benchmark result (4 %) by approximately 0.6 p.p.

Alternatively, as an extension to the basic analysis of annual changes in productivity, efficiency and technology, we employ a Malmquist analysis to only two periods: in this case the DEA model is estimated with 3-year averages (T1 = 2001–2003 and T2 = 2003–2005), so that the Malmquist index obtained can be interpreted as the average productivity change between T1 and T2 and fully corrects for time variation in the original annual data on inputs and

---

[29] We are aware of the fact that the basic model only partially captures the relationship between student cohorts (as inputs) and graduates (as outputs). Data is annual, so the input measuring the total number of students corresponds to students attending at any level at the university in the present year. At the same time, we do not expect the number of graduates this year to be dependent on the number of first year students this year. Unfortunately, data on students divided by year of attending the university is not available for most countries in the sample. However, one might think that the proportion of first year students to the total number of students in a given university tends to be stable, so that the basic DEA model employing the total number of students as one of the inputs and the total number of

---

Footnote 29 continued

graduates as one of the outputs approximates productivity in the teaching process well. We thank a referee for raising this point.

[30] We thank a referee for this suggestion.

[31] The correlation coefficient between different estimates of *m* ranges between 0.62 and 0.97.

Table 3 Results—trends in productivity (*m*), efficiency (ε) and technology (τ) in 266 European HEIs, change between T1:2001–2003 and T2:2003–2005 (based on 3-year averages); CRS

	$TFP = \hat{m}$	$\hat{\varepsilon}$	$\hat{\tau}$
Number of all indices	$N(\hat{m})=266$	$N(\hat{\varepsilon})=266$	$N(\hat{\tau})=266$
Average value of all indices	$\bar{\bar{m}} = 1.089$	$\bar{\bar{\varepsilon}} = 1.06$	$\bar{\bar{\tau}} = 1.03$
Number of statistically significant indices	$N(\hat{m}**) = 245$	$N(\hat{\varepsilon}**) = 136$	$N(\hat{\tau}**) = 71$
Average value of statistically significant indices	$\bar{\bar{m}}** = 1.096$	$\bar{\bar{\varepsilon}}** = 1.094$	$\bar{\bar{\tau}}** = 1.100$
Number of statistically significant improvements	$N(\hat{m}** > 1) = 193$	$N(\hat{\varepsilon}** > 1) = 91$	$N(\hat{\tau}** > 1) = 71$
Percentage (out of 266) of HEIs registering statistically significant improvements between T1 and T2	$\%(\hat{m}** > 1) = 72\%$	$\%(\hat{\varepsilon}** > 1) = 34\%$	$\%(\hat{\tau}** > 1) = 27\%$

** Refer to significance at 5 % level. Results based on three input—two output model. CRS—constant returns to scale

Source: own elaboration

outputs. Crucial results based on such averaged data are reported in Table 3 and can be compared with the evidence on annual changes reported in Table 1.

On average, productivity in European HEIs rose by approximately 9 % between the initial period T1 and final period T2 ($\bar{\bar{m}} = 8.9\%$ and $\bar{\bar{m}}** = 9.6\%$)—note that this result is actually in line with the evidence on annual change reported in Table 1 (where $\bar{\bar{m}} = 4\%$; $\bar{\bar{m}}** = 4.5\%$) because the input and output values for T1 and T2 are in fact averaged data around 2002 and 2004. Consequently, the estimates of *TFP* growth obtained with 3-year averages should be approximately twice as large as those obtained with annual data, and this indeed is the case. The only difference is that when we consider a longer time horizon, the proportion of HEIs registering statistically significant improvements in productivity is larger than in the case of annual changes in productivity (72 % versus 56 %, respectively).

3.3 Malmquist indices: accounting for cross-country heterogeneity

Our dataset has the important property of panel dimension. Thus, we can check for country-specific trends in productivity, efficiency and technology change. In Table 4, we report the average (by country) values of *m*, ε and τ (all and only those which are statistically significant) and the percentage of cases with statistically significant annual improvements in productivity, efficiency and technology. In most cases (with the exception of technology change in Poland) accounting for statistical significance only negligibly alters the average values of the indices estimated, so in the interpretation of the results we limit ourselves to the significant ones.

The average statistically meaningful *TFP* change indicated by the Malmquist index ranges from 0.98 (*TFP* decline of 2 % annually) in Austrian HEIs to 1.09 (*TFP* growth of 9 % annually) in Switzerland, where the average efficiency change was also the highest (rising by 19 %

annually). Only Austrian HEIs registered a decline in average efficiency: by 4 % ($\bar{\bar{\varepsilon}}** = 0.96$).

In all of the countries examined, $\%(\hat{m}** > 1) > \%(\hat{\varepsilon}** > 1) > \%(\hat{\tau}** > 1)$, so that the percentage of cases with HEIs registering statistically significant improvements in *TFP* was larger than the percentage of cases with statistically significant efficiency growth, which, in turn, was higher than the percentage of cases with statistically significant positive changes in technology. For instance, 69 % of 216 annual observations on Italian HEIs (54 university units observed across four time periods) registered statistically significant *TFP* growth; 45 % were characterized by statistically significant improvements in efficiency, but only 9 % showed statistically significant improvements in technology.[32] Among the seven European countries analyzed, Italy had the highest percentage of public HEIs with significant *TFP* growth and significant efficiency improvement.

The time dimension can also be important when analyzing the productivity changes of HEIs in European countries. We have isolated only HEIs with Malmquist indices statistically significantly different from unity (at the 5 % level)—that is, either higher (statistically significant productivity increases) or lower (statistically significant productivity decreases). Of

---

[32] Comparing our results to those of other authors, we can only compare evidence concerning Italy and the UK (no comparable studies have been performed for other European countries). First of all, our results concerning Italian HEIs are in line with those of Agasisti and Johnes (2009), who also find a dominating efficiency change effect driving *TFP* growth in Italian universities in the period 2002/2003 and 2004/2005. On the other hand, (Agasisti and Pérez-Esparrells 2010), in contradiction to our findings, conclude that productivity growth in Italian HEIs was due to major technological change (frontier shift). However, they do not count annual changes, as we do, but overall change between 2000/2001 and 2004/2005. Secondly, our estimate concerning British HEIs (2 % annual rise in *TFP*) is higher than the one obtained by Johnes (2008), who finds a 1 % increase in *TFP* per year over the period 1996/1997–2004–2005/, but in line with these authors and with Agasisti and Johnes (2009) and Flegg et al. (2004), we also find a dominating role of frontier shifts in causing productivity gains in UK universities.

Table 4 Results by country (1): annual changes in productivity (*m*), efficiency (*ε*) and technology (*τ*)—mean values (all and statistically significant) and percentage of cases with statistically significant improvements; CRS

Country	$TFP = \hat{m}$	$\hat{\varepsilon}$	$\hat{\tau}$
AUT	$\bar{\bar{m}} = 0.99$	$\bar{\bar{\varepsilon}} = 0.97$	$\bar{\bar{\tau}} = 1.02$
n = 9, 4 time periods	$\bar{\bar{m}} ** = 0.98$	$\bar{\bar{\varepsilon}} ** = 0.96$	$\bar{\bar{\tau}} ** = 1.03$
N = 36	$\%(\hat{m}** > 1) = 50\%$	$\%(\hat{\varepsilon}** > 1) = 28\%$	$\%(\hat{\tau}** > 1) = 6\%$
FIN	$\bar{\bar{m}} = 1.01$	$\bar{\bar{\varepsilon}} = 1.00$	$\bar{\bar{\tau}} = 1.01$
n = 15, 4 time periods	$\bar{\bar{m}} ** = 1.01$	$\bar{\bar{\varepsilon}} ** = 1.01$	$\bar{\bar{\tau}} ** = 1.04$
N = 60	$\%(\hat{m}** > 1) = 47\%$	$\%(\hat{\varepsilon}** > 1) = 25\%$	$\%(\hat{\tau}** > 1) = 20\%$
GER	$\bar{\bar{m}} = 1.06$	$\bar{\bar{\varepsilon}} = 1.05$	$\bar{\bar{\tau}} = 1.01$
n = 67, 4 time periods	$\bar{\bar{m}} ** = 1.06$	$\bar{\bar{\varepsilon}} ** = 1.1$	$\bar{\bar{\tau}} ** = 1.04$
N = 268	$\%(\hat{m}** > 1) = 62\%$	$\%(\hat{\varepsilon}** > 1) = 32\%$	$\%(\hat{\tau}** > 1) = 16\%$
ITA	$\bar{\bar{m}} = 1.07$	$\bar{\bar{\varepsilon}} = 1.07$	$\bar{\bar{\tau}} = 1.00$
n = 54, 4 time periods	$\bar{\bar{m}} ** = 1.08$	$\bar{\bar{\varepsilon}} ** = 1.10$	$\bar{\bar{\tau}} ** = 1.01$
N = 216	$\%(\hat{m}** > 1) = 69\%$	$\%(\hat{\varepsilon}** > 1) = 45\%$	$\%(\hat{\tau}** > 1) = 9\%$
POL	$\bar{\bar{m}} = 1.03$	$\bar{\bar{\varepsilon}} = 1.01$	$\bar{\bar{\tau}} = 1.03$
n = 31, 4 time periods	$\bar{\bar{m}} ** = 1.03$	$\bar{\bar{\varepsilon}} ** = 1.02$	$\bar{\bar{\tau}} ** = 1.13$
N = 124	$\%(\hat{m}** > 1) = 57\%$	$\%(\hat{\varepsilon}** > 1) = 29\%$	$\%(\hat{\tau}** > 1) = 25\%$
SWT	$\bar{\bar{m}} = 1.08$	$\bar{\bar{\varepsilon}} = 1.09$	$\bar{\bar{\tau}} = 0.99$
n = 11, 4 time periods	$\bar{\bar{m}} ** = 1.09$	$\bar{\bar{\varepsilon}} ** = 1.19$	$\bar{\bar{\tau}} ** = 0.97$
N = 44	$\%(\hat{m}** > 1) = 55\%$	$\%(\hat{\varepsilon}** > 1) = 36\%$	$\%(\hat{\tau}** > 1) = 5\%$
UK	$\bar{\bar{m}} = 1.02$	$\bar{\bar{\varepsilon}} = 1.00$	$\bar{\bar{\tau}} = 1.02$
n = 79, 4 time periods	$\bar{\bar{m}} ** = 1.02$	$\bar{\bar{\varepsilon}} ** = 1.01$	$\bar{\bar{\tau}} ** = 1.03$
N = 316	$\%(\hat{m}** > 1) = 46\%$	$\%(\hat{\varepsilon}** > 1) = 24\%$	$\%(\hat{\tau}** > 1) = 21\%$

** Refers to significance at 5 % level. CRS—constant returns to scale. Results based on three input—two output model. n—number of HEIs by country, N—total number of cases analyzed (by country), N = 4n

Source: own elaboration

Table 5 Results by country (2)—changes in productivity (*m*), efficiency (*ε*) and technology (*τ*) based on 3 year averages (T1 = 2001–2003 and T2 = 2003–2005) and alternative frontier definition (E-European frontier; C-country specific frontier): mean values of all indices; CRS

Country	$TFP = \bar{\bar{m}}$		$\bar{\bar{\varepsilon}}$		$\bar{\bar{\tau}}$	
	E	C	E	C	E	C
FIN (n = 15)	1.028	1.019	1.00	0.976	1.027	1.044
GER (n = 67)	1.112	1.106	1.081	0.984	1.03	1.127
ITA (n = 54)	1.175	1.169	1.157	0.996	1.016	1.174
POL (n = 31)	1.056	1.055	1.077	1.038	0.98	1.016
UK (n = 79)	1.038	1.033	0.968	0.958	1.075	1.078

CRS—constant returns to scale, n—number of HEIs by country. Results based on three input—two output model

Source: own elaboration

these, we have calculated the average *TFP* across the HEIs in given countries in each time period. Figure 1 shows the average significant change in *TFP* in public European HEIs, by country and year (detailed data are reported in Table 12 in the Appendix). It turns out that if we take into account

exclusively those universities that really (in a statistical sense) registered a change in productivity, on average German, Italian and Swiss HEIs performed better (having constant rises in *TFP*) than HEIs in other countries.

Due to space limits, we are not able to report all the Malmquist indices for every HEI and year analyzed. However, we count European universities that registered constant statistically significant improvements in *TFP* (thus having Malmquist indices significantly larger than unity in all of the time periods). Of the 266 universities in our sample, this was the case in only 28 units. Among these we find: two HEIs from Finland, eight HEIs from Germany, fourteen HEIs from Italy, one from Poland, two from Switzerland and one from the UK.[33]

Finally, in order to check whether the definition of the frontier matters for country-specific conclusions, we consider two alternative applications of the DEA model: the first based on a pooled sample (266 HEIs) and thus reflecting the general 'European frontier'; the second based on separate DEA models for each country (where each HEI

---

[33] List of units available upon request.

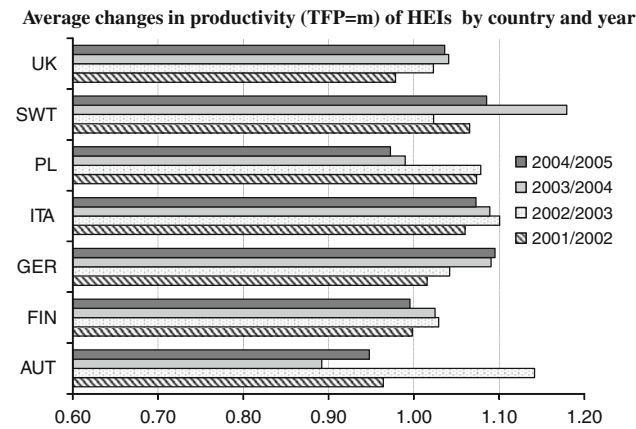**Average changes in productivity (TFP=m) of HEIs by country and year**

Fig. 1 Average statistically significant changes in total factor productivity (*m*) of HEIs by country and year. Note: Results based on three input—two output model; only $\hat{m}$ statistically different from the unity taken into account for the calculation of averages; 5 % level of statistical significance. *Source*: Own elaboration

was evaluated with respect to the units from the same country, e.g. comparing the performance of Italian HEIs with other Italian HEIs etc.). This exercise could be done for five of the seven countries in our sample—in the cases of Austria and Switzerland the number of decision-making units is not sufficient to estimate the frontier and assure a reasonable level of discrimination.[34] A comparison between the two approaches to frontier definition can be particularly informative when comparing efficiency change and analyzing whether universities were getting closer to (or further away from) the overall 'European' efficiency frontier or their national frontiers (influenced by country-specific educational policies etc.).[35] Table 5 summarizes the results by country, based on 3-year averaged data and corresponding to the two alternative definitions of the frontier (*E*—European frontier and *C*—country-specific frontier). The values reported correspond to mean *TFP*, efficiency and technology change in HEIs from single countries in the periods T1:2001–2003 and T2: 2003–2005.

In turns out that frontier definition is less important for the measurement of general productivity change indicated by the Malmquist index (which remains the main interest of our analysis) than for its components. The correlation between the series of $\bar{\hat{m}}$ obtained with the European frontier (E) and that using the country-specific frontier (C) equals 0.99 and their average values are very similar. Italian and German HEIs registered the biggest *TFP* change in T1 (2001–2003) and T2 (2003–2005) (by 17 and 11 %, respectively). However, the channels through which productivity changes are materialized differ depending on the frontier formulation. This observation leads us to the issue of frontier definition in DEA/Malmquist studies performed for units from different countries.[36]

As far as the common European frontier is concerned (E), HEIs from all countries obtained an average rise in productivity ($\bar{\hat{m}} > 1$), which in the cases of German, Italian and Polish HEIs was mainly due to an increase in their relative efficiency (movement towards the European frontier—catching up effect), while in the case of Finish and British universities productivity growth was achieved more through technology change.

Using the country-specific frontier model (C), again, universities from all countries on average registered improvements in their productivity. It is notable, however, that this time a rise in *TFP* was mostly due to shifts of their country-specific frontiers (as indicated by $\hat{\tau}$, the technology change estimate in country-specific frontier setting). Only in the case of Polish HEIs do we obtain a value of $\bar{\hat{\varepsilon}}$ greater than 1 in the country-specific approach, meaning that units from Poland were not only catching up with the European frontier but also with the national one. On the contrary, German HEIs on average caught up with European ones ($\bar{\hat{\varepsilon}} = 1.081$ in the European frontier setting) but moved back from the German efficiency frontier ($\bar{\hat{\varepsilon}} = 0.98$ in the country-specific frontier setting). This means that the German higher education frontier was rising more quickly than the overall European one. Similar patterns emerge when analyzing the Italian case. Consequently, the choice of benchmark against which we assess the efficiency performance of universities makes a difference. This is an important result and will be included amongst the guidelines for future research which we propose together with our conclusions.

---

[34] As a rule of thumb, we use the criterion that the number of DMUs should be at least 2*i**o,* where *i**o is the product of the numbers of inputs and outputs (Boussofiane et al. 1991; Dyson et al. 2001). In our case, the use of a DEA model with three inputs and two outputs requires a set of at least 12 units for each country.

[35] As an alternative, another possible method for comparing country frontiers is the meta-frontier approach, where the metafrontier function is defined as: "an overarching function that encompasses the deterministic components of the stochastic frontier production functions for the units that operate under the different technologies involved" (Battese et al. 2008). Such a methodology has been mainly applied in the context of regional variation in the data. In our case, such a European metafrontier would envelope country-specific productivity frontiers for HEIs from separate countries in the sample. More on the application of the metafrontier method for the decomposition of the Malmquist index can be found in Oh and Lee (2010). We thank a referee for drawing our attention to this issue.

4 Conclusions and suggestions for future research

Despite increasing pressure on public universities to constantly optimize results using limited resources, *changes* in

---

[36] We thank a referee for pointing out this issue.

university productivity have only been marginally analyzed, usually with respect to HEIs from just one or at most two countries. Cross-country multi-period analysis of productivity trends is demanding, as it requires the collection of micro data for the same units and for multiple time periods. In the case of universities from several European countries, it has proven to be a quite challenging, albeit feasible, piece of research.

Our paper contributes to the existing literature by presenting productivity changes (along with efficiency and technology trends) in 266 public HEIs from seven European countries for the years 2001–2005 (analyzed mainly annually, but also in terms of time averages). Moreover, we have proposed the application of important methodological improvements, providing consistent estimates of Malmquist indices, along with their confidence intervals, based on a bootstrap method. Consequently, our conclusions are based on statistically significant results that do not suffer from sample noise and, hence, are statistically robust.

These robust results indicate that, of the 1,064 annual estimates of *TFP* growth in the European HEIs analysed, 963 (90 %) were statistically different from unity (at a standard 5 % level of significance) so the majority of HEIs registered statistically significant *changes* in productivity. Between the years 2001 and 2005, HEIs from our sample registered an average increase in productivity of around 4.5 % per year and efficiency change predominated over technology improvements. However, the methodology adopted permits us to state that only approximately half of the cases were characterized by statistically significant *annual improvements* in overall productivity. In the other cases, either the *TFP* of HEIs declined or their Malmquist indices were not significantly different from unity (no improvement—no regress).

Our study has benefited from the advantage of being based on panel data, with information on the productivity performance of universities from several European countries. Consequently, we have thoroughly analyzed cross-country variation in productivity changes that are typical for universities from different systems of higher education. The average *TFP* index ranged from 0.98 (*TFP* decline of 2 %, annually) in Austria to 1.09 (*TFP* growth of 9 %, annually) in Switzerland. There is also much inter-country variation in the proportion of universities that registered statistically significant improvements in productivity. For instance, two-thirds of Italian HEIs registered statistically significant *TFP* improvements (the best score across the seven countries), while this was typical for less than half (46 %) of British universities.

With regard to the time dimension, we have been able to check which universities registered constant *TFP* growth in every time period across the years 2001–2004. On average, German, Italian and Swiss HEIs, whose *TFP* rose consistently, performed better than HEIs from the other countries. Looking at single university units, in our basic analysis evaluating HEIs vis-à-vis a common European frontier of productivity, we found that only 28 European universities (out of 266) registered statistically significant improvements in productivity in *all* of the years between 2001 and 2005.

We have extended our analysis by comparing the results obtained with alternative datasets, by changing the set of inputs and outputs in the DEA estimation and by employing alternative definitions of the productivity frontier ('European' and 'country specific'). Our basic finding of approximately 4 % annual productivity growth in European HEIs is robust to changes in the formulation of the DEA model for the Malmquist index calculations. Frontier definition is not so important for the measurement of general productivity change (the Malmquist index remains fairly stable) but proves to be relevant when comparing efficiency and technology developments. A joint treatment of universities with respect to a common productivity frontier is appropriate if the researcher is interested in comparing HEIs as units competing jointly within the European system of higher education, as we were. Assessing HEIs against other units from the same country tells more about the movement of national frontiers of higher education. Consequently, through alternative frontier measurement we demonstrate that, depending on the research question formulated at the outset, the need to take into account the heterogeneity of higher education systems across countries should be considered.

Appendix

See Tables 6, 7, 8, 9, 10, 11 and 12.

Table 6 Data appendix—detailed description of variables used for the calculation of Malmquist indices

Variable	Country	Remarks
Number of students[a,b]	Finland	Total number of registered students (all types of undergraduate and postgraduate studies with foreigners), headcounts
	Switzerland	Total number of registered students (all types of undergraduate and postgraduate studies with foreigners) headcounts, referring to beginning of the academic year
	Germany	Total number of registered students (all types of undergraduate and postgraduate studies), with foreigners, referring to the winter semester; headcounts
	Austria	Total number of registered students (all types of undergraduate and postgraduate studies) referring to the winter semester. with foreigners, headcounts
	UK	Total number of registered students (all types of undergraduate and postgraduate studies) with foreigners, headcounts, full time and part time
	Italy	Total number of registered students (all types of undergraduate and postgraduate studies) with foreigners
	Poland	Total number of registered students (all types of undergraduate and postgraduate studies without foreigners (separate data concerning foreign students available only since 2006 when percentage of foreign students in total ranged between 0.02 and 2.6 %) headcounts, full time and part time
Total academic staff[a,c]	Finland	Professors, associate professors, senior assistants, assistants, lecturers, teachers and research personnel, full time equivalent,
	Switzerland	Professors, adjuncts and lectures, full time equivalent, referring to the last day of each year
	Germany	Professors, lecturers, scientific assistants, scientific and artistic employees, teaching personnel, full time employment
	Austria	Professors, assistants and other academic staff, full time equivalent
	UK	Teachers, teachers and researchers, researchers, full time equivalent.
	Italy	Professors (1st and 2nd category) researchers, registered at the end of the year, who in December received at least 95 % of the salary typical for the post full time employment
	Poland	Professors, docents, adjuncts, assistants, senior lecturers, lecturers and specialist librarians, full time employment
Total revenues[a]	Finland	Originally reported in euro, yearly
	Switzerland	Originally reported in Swiss frank, yearly
	Germany	Originally reported in euro, yearly
	Austria	Originally reported in euro, yearly
	UK	Originally reported in pounds, yearly
	Italy	Originally reported in euro, yearly
	Poland	Originally reported in PLN, yearly
Number of publications	Finland	According to Thomson Reuter's ISI Web of Science (set of journals, conference proceedings etc. common to all countries). HEIs for which identification of the publication record was impossible, were excluded from the sample
	Switzerland	
	Germany	
	Austria	
	UK	
	Italy	
	Poland	
Number of graduates[1]	Finland	Total number of graduates (all types of studies), all
	Switzerland	Total number of graduates (all types of studies), all
	Germany	Total number of graduates (all types of studies), all
	Austria	Total number of graduates (all types of studies), all
	UK	Total number of graduates (all types of studies), all
	Italy	Total number of graduates (all types of studies), all
	Poland	Total number of graduates (all types of studies) without foreigners (separate data concerning foreign students available only since 2008 when percentage of foreign students in total ranged between 0 and 2.25 %)

[a] If not stated differently, data reported originally for respective academic year (thus value in our dataset matched with the year 2002 refers to academic year 2001/2002, and so on)

[b] According to the UOE manual (2004, p.22) as students we consider any individual participating in tertiary education service in the reference period

[c] In line with the UOE manual (2004, p.34) as academic staff we consider: "personnel whose primary assignment is instruction, research or public service; personnel who hold an academic rank with such titles as professor, associate professor,assistant professor, instructor, lecturer, or the equivalent of any of these academic ranks; personnel with other titles if their principal activity is instruction or research."

Source: own elaboration

Table 7 Panel composition—number of HEIs from every country covered by the analysis

Country	Number of HEIs
Poland (PL)	31
Austria (AUT)	9
Finland (FIN)	15
Germany (GER)	67
Italy (ITA)	54
Great Britain (UK)	79
Switzerland (SWT)	11
Total	266

Source: own elaboration

Table 8 Data sources on HEIs covered by our study

Country	Source	Online platform
Finland	Finnish Ministry of Education	https://kotaplus.csc.fi/online/Haku.do
Switzerland	Swiss Federal Statistical Office	www.statistique.admin.ch
Germany	Federal Statistical Office (Destatis)	www.destatis.de
Austria	Austrian Federal Ministry of Science and Research	http://www.bmwf.gv.at/unidata
UK	Higher Education Statistics Agency	http://www.heidi.ac.uk/
Italy	Ministry of Science and Education (MIUR)	www.nuclei.cnvsu.it; www.dalia.cineca.it
Poland	Ministry of Science and Higher Education	www.nauka.gov.pl
	Central Statistical Office (GUS)	www.stat.gov.pl

Source: own elaboration

Table 9 Comparison between mean square error of bias corrected and 'original' estimates of *Malmquist* index and its components

Number of obs. in which	
$MSE(\hat{m}\_corr_{i,(t,t+1)}) > MSE(\hat{m}_{i,(t,t+1)})$	1,013
Number of obs. in which $MSE(\hat{\varepsilon}\_corr_{i,(t,t+1)}) > MSE(\hat{\varepsilon}_{i,(t,t+1)})$	1,056
Number of obs. in which $MSE(\hat{\tau}\_corr_{i,(t,t+1)}) > MSE(\hat{\tau}_{i,(t,t+1)}) > MSE(\hat{\tau}_{i,(t,t+1)})$	1,062
Total number of observations	1,064

Source: own elaboration

Based on three input—two output model

Table 10 Summary statistics of input and output variables used in DEA model (three input—two output model)

Variable (annual)	Mean	Min	Max	SD	N
Total number of academic staff	1268.98	113	5,567	947.905	1330
Total number of students	20569.68	1,584	136,839	15591.09	1,330
Total real revenues (euro PPS)	1.60E + 08	1.13E + 07	9.15E + 08	1.32E + 08	1,330
Number of publications	799.2128	3	6,336	894.3691	1,330
Total number of graduates	3401.953	129	21,044	2667.351	1,330

Source: own elaboration

Table 11 Summary statistics of the estimates

Variable	Mean	Min	Max	SD	N	n	T
$\hat{m}_{i,(t,t+1)}$	1.006	0.607	1.435	0.107	1,052	266	4
$\hat{\varepsilon}_{i,(t,t+1)}$	0.973	0.548	1.485	0.133	1,059	266	4
$\hat{\tau}_{i,(t,t+1)}$	1.044	0.838	1.424	0.098	1,063	266	4
$Avg[\hat{m}(b)_{i,(t,t+1)}]$	1.004	0.625	1.413	0.104	1,046	266	4
$Avg[\hat{\varepsilon}(b)_{i,(t,t+1)}]$	0.982	0.506	1.501	0.130	1,058	266	4
$Avg[\hat{\tau}(b)_{i,(t,t+1)}]$	1.034	0.889	1.357	0.083	1,064	266	4
$bias(\hat{m}_{i,(t,t+1)})$	0.000	−0.046	0.042	0.011	1,034	266	4
$bias(\hat{\varepsilon}_{i,(t,t+1)})$	0.008	−0.096	0.104	0.028	1,058	266	4
$bias(\hat{\tau}_{i,(t,t+1)})$	−0.010	−0.096	0.073	0.026	1,061	266	4
$\hat{m}\_corr_{i,(t,t+1)}$	1.006	0.584	1.438	0.107	1,054	266	4
$\hat{\varepsilon}\_corr_{i,(t,t+1)}$	0.964	0.531	1.495	0.140	1,058	266	4
$\hat{\tau}\_corr_{i,(t,t+1)}$	1.053	0.725	1.491	0.118	1,063	266	4
$l_\alpha(\hat{m}),\ \alpha = 0.05$	0.992	0.571	1.401	0.105	1,054	266	4
$u_\alpha(\hat{m}),\ \alpha = 0.05$	1.021	0.661	1.454	0.108	1,050	266	4
$l_\alpha(\hat{\varepsilon}),\ \alpha = 0.05$	0.894	0.469	1.408	0.131	1,061	266	4
$u_\alpha(\hat{\varepsilon}),\ \alpha = 0.05$	1.053	0.581	1.672	0.160	1,059	266	4
$l_\alpha(\hat{\tau}),\ \alpha = 0.05$	0.974	0.672	1.290	0.109	1,064	266	4
$u_\alpha(\hat{\tau}),\ \alpha = 0.05$	1.134	0.890	1.601	0.122	1,057	266	4

Without outliers (method of Hadi 1994). Results based on three input—two output model

Source: own elaboration

Table 12 Average significant annual change in *TFP* in public HEIs by country and time period (corresponding to Fig. 1)

Time period	$TFP = \hat{m}$						
	AUT	FIN	GER	ITA	PL	SWT	UK
2001/2002	0.964	0.998	1.016	1.060	1.074	1.066	0.979
2002/2003	1.142	1.029	1.042	1.101	1.078	1.023	1.023
2003/2004	0.892	1.025	1.091	1.089	0.990	1.179	1.041
2004/2005	0.948	0.995	1.095	1.073	0.973	1.085	1.036

Results based on three input—two output model, only *m* statistically different than unity taken into account for the calculation of averages, 5 % level of statistical significance

Source: own elaboration

References

Agasisti T, Johnes G (2009) Beyond frontiers: comparing the efficiency of higher education decision-making units across countries. Educ Econ 17(1):59–79

Agasisti T, Pérez-Esparrells C (2010) Comparing efficiency in a cross-country perspective: the case of Italian and Spanish state universities. High Educ 59(1):85–103

Assaf A (2011) Bootstrapped Malmquist indices of Australian airports. Serv Ind J 31(5):829–846

Assaf A, Barros C, Matousek R (2010) Productivity and efficiency analysis of Shinkin banks: evidence from bootstrap and Bayesian approaches. J Banking Financ 35(2):331–342

Balcome K, Davidova S (2008) The use of bootstrapped Malmquist indices to reassess productivity change findings: an application to a sample of Polish farms. Appl Econ 40(16):2055–2061

Battese G, Rao DSP, O'Donnell CJ (2008) A Metafrontier production function for estimation of technical efficiencies and technology gaps for firms operating under different technologies. J Prod Anal 21(1):91–103

Bonaccorsi A, Daraio C, Simar L (2007). Efficiency and productivity in European universities: exploring trade-offs in the strategic profile. In: Bonaccorsi A, Daraio C (eds) Universities and strategic knowledge creation: specialization and performance in Europe, Chapter 5. Edward Elgar Publishing, Cheltenham/Northampton, pp 144–206

Boussofiane A, Dyson RG, Thanassoulis E (1991) Applied data envelopment analysis. Eur J Oper Res 52(1):1–15

Caves DW, Christensen LR, Diewert WE (1982) The economic theory of index numbers and the measurement of input, output and productivity. Econometrica 5:1393–1414

Coelli TJ, Rao DP, O'Donnell CJ, Battese GE (2005) An introduction to efficiency and productivity analysis, 2nd edn. Springer, Berlin

Creamer EG (1999) Knowledge production, publication productivity and intimate academic partnerships. J High Educ 70(3):261–277

Daraio C, Bonaccorsi A, Geuna A, Lepori B, Bach L, Bogetoft P, Cardoso MF, Castro-Martinez E, Crespi G, de Lucio IF, Fried H, Garcia-Aracil A, Inzelt A, Jongbloed B, Kempkes G, Llerena P, Matt M, Olivares M, Pohl C, Raty T, Rosa MJ, Sarrico CS, Simar L, Slipersaeter S, Teixeira P, Vanden Eeckaut P (2011) The European university landscape: a micro characterization based on evidence from the Aquameth project. Res Policy 40(1):148–164

Dundar H, Lewis DR (1998) Determinants of research productivity in higher education. Res High Educ 39(6):607–631

Dyson RG, Allen R, Camanho AS, Podinovski VV, Sarrico CS, Shale EA (2001) Pitfalls and protocols in DEA. Eur J Oper Res 132(2):245–259

Efron B (1982) The Jacknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia

Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman and Hall, London

Färe R, Grosskopf S, Lindgren B, Roos P (1992) Productivity changes in Swedish pharmacies 1980–1989: a nonparametric approach. J Prod Anal 3:85–101

Färe R, Grosskopf S, Norris M, Zhang Z (1994) Productivity growth, technical progress and efficiency changes in industrialized countries. Am Econ Rev 84:66–83

Färe R, Griffell-Tajte E, Grosskopf S, Knox Lovell CA (1997) Biased technical change and Malmquist productivity index. Scand J Econ 99:119–127

Farrell MJ (1957) The measurement of productive efficiency. J R Stat Soc A 120:253–281

Ferrier GD, Hirschberg JG (1997) Bootstrapping confidence intervals for linear programming efficiency scores: with an illustration using Italian bank data. J Prod Anal 8:19–33

Flegg AT, Allen DO, Field K, Thurlow TW (2004) Measuring the efficiency of British universities: a multi-period data envelopment analysis. Educ Econ 12(3):231–249

Glass JC, McKillop DG, Hyndman NS (1995) Efficiency in the provision of university teaching and research: an empirical analysis of UK universities. J Appl Econom 10:61–72

Hadi AS (1994) A modification of a method for the detection of outliers in multivariate samples. J R Stat Soc Ser B 56:393–396

Johnes J (2006a) Data envelopment analysis and its application to the measuring of efficiency in higher education. Econ Educ Rev 25:273–288

Johnes J (2006b) Measuring efficiency: a comparison of multilevel modelling and data envelopment analysis in the context of higher education. Bull Econ Res 58:75–104

Johnes J (2008) Efficiency and productivity change in the English higher education sector from 1996/97 to 2004/2005. Manch Sch 76(6):653–674

Mahlberg B, Url T (2010) Single market effects on productivity in the german insurance industry. J Banking Financ 34(7):1540–1548

OECD (2002) Frascati manual. Proposed Standard Practice for Surveys on Research and Experimental Development. OECD, Paris

Oh D-h, Lee J-d (2010) A metafrontier approach for measuring Malmquist productivity index. Empir Econ 38(1):47–64

Simar L, Wilson P (1998) Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models. Manag Sci 44(11):49–61

Simar L, Wilson P (1999) Estimating and bootstrapping Malmquist indices. Eur J Oper Res 115:459–471

Simar L, Wilson P (2000) A general methodology for bootstrapping in non-parametric frontier models. J Appl Stat 27(6):779–802

Simar L, Wilson P (2007) Estimation and inference in two stage, semi-parametric models of productive efficiency. J Econom 136:31–64

Simar L, Wilson P (2011) Two-stage DEA: caveat emptor. J Prod Anal 36:205–218

Unesco-UIS/OECD/Eurostat (2004) 2004 data collection on education systems. In: UOE data collection manual. OECD, Paris

Wilson P (2008) FEAR: a software package for frontier efficiency analysis with R. Soc Econ Plan Sci 42(4):247–254

Wolszczak-Derlacz J, Parteka A (2010). Scientific productivity of public higher education institutions in Poland—a comparative bibliometric analysis. EY—Better Governance, Warsaw. Downloadable at: http://webapp01.ey.com.pl/EYP/WEB/eycom_down load.nsf/resources/Scientific_productivity.pdf/$FILE/Scientific_productivity.pdf

Wolszczak-Derlacz J, Parteka A (2011) Efficiency of European public higher education institutions: a two-stage multicountry approach. Scientometrics 89:887–917

Worthington AC, Lee BL (2008) Efficiency, technology and productivity change in Australian universities, 1998–2003. Econ Educ Rev 27:285–298