UNIVERSITY OF MACEDONIA OF ECONOMICS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS
INTERDEPARTMENTAL MASTER PROGRAMM IN ECONOMIC STUDIES

Meta-regression analysis of higher education productivity growth studies

Master thesis of graduate student:
Ravanos Panagiotis , A/M: 01/13
Supervisor: Professor Giannis Karagiannis, University of Macedonia, Department of Economics

THESSALONIKI FALL 2013
SUBMISSION DATE: 26/02/2014

UNIVERSITY OF MACEDONIA OF ECONOMICS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS
INTERDEPARTMENTAL MASTER PROGRAMM IN ECONOMIC STUDIES

Meta-regression analysis of higher education productivity growth studies

Ravanos Panagiotis, University of Macedonia

Master thesis submission

Interdepartmental Master Program in Economic Studies

Acknowledgments

The completion of this thesis was co-financed by the project "Scholarships SSF'', using resources from the OP" Education and Lifelong Learning ", of European Social Fund (ESF) of the NSRF 2007-2013.

Table of contents

Abstract

This study provides a meta-regression analysis (MRA) of studies measuring productivity in higher education. We consider a total of 31 published and unpublished studies (working papers, theses e.t.c), written or published until April 2013, from which we finally use 28, yielding a total of 1018 observations. We examine the impact of different study characteristics on the estimated rate of productivity change and in addition we check the robustness of our findings using three alternative estimation methods for MRA. The first uses equally weighted observations, implicitly giving more weight to studies with multiple TFP estimates, the second uses per paper weights, resulting in equal weight in cross-paper observations, and the third uses one observation per paper. The results indicate that there are actual differences between the three different weighing methods, in the MRA estimates.

Acknowledgments

The completion of this thesis was co-financed by the project "Scholarships SSF'', using resources from the OP" Education and Lifelong Learning ", of European Social Fund (ESF) of the NSRF 2007-2013.

1. Introduction

During the last years, the existing literature on efficiency and productivity analysis in the education sector has become vaster than ever. This can be interpreted as a growing interest on the performance of schools, universities, colleges, and every other form of education unit, as our society is moving to an era with need for highly educated and highly qualified employees.

Though, measuring the efficiency or the productivity of an education unit is not an easy task. Today's universities and schools differ from country to country, even in different regions of the same country, in key aspects that increase the heterogeneity between them, making their efficiency and productivity measurement more difficult. Also an education unit is much more different than a normal production unit, in aspects such as the definition of inputs and outputs, their objective function, goal e.t.c. In that sense, an educational unit may be better seen as a Decision Making Unit (DMU) instead of a pure production unit.

All these reasons make the efficiency and productivity measurement in the education sector a hard but very interesting task. The relevant literature is vast in efficiency measurement but less in the case of productivity measurement, and this holds as evidence for studies on the higher education sector, which will be the subject of this thesis.

The objective of this thesis is twofold. First, to provide a quantitative literature review of the studies measuring productivity in the higher education sector. This will be done by means of meta-regression analysis (MRA), where productivity growth is the relevant dependent variable, and selected meta-independent variables are used to account for various study characteristics. Second, we examine potential differences in the meta-regression-analysis results due to different weighting methods for constructing the dependent variable. This is an issua as long as the studies included in the analysis report more than one estimates of the relevant variable, the rate of productivity change in our case. Following Stanley (2001), we three alternate models, each one based on a different weighting method.

The rest of the paper is organized as follows: in section 2 we provide a brief theoretical framework for efficiency and productivity measurement in the higher education sector, and for the concept of meta-regression analysis; in section 3 we describe how the database was constructed, the selection of the meta-independent variables, and present and discuss some descriptive statistics, and in section 4 we present the empirical results based on the three different models. The paper ends with some concluding remarks.

2. Theoretical framework.

I. Efficiency and productivity measurement in the higher education sector

The higher education sector is part of the ''service industries''. However, universities are not pure firms with easily measured inputs and outputs. They are considered as DMUs with multiple inputs and outputs, not all of which can be measured accurately with the usually proposed proxies. The three main activities of a university are the following: teaching, research and community service. The first two activities are relatively easier to measure with inputs and outputs such as number of students (total or full time equivalent), number of research publications, total number of research grants received, etc. (see also, Worthighton (2001) and Gates and Stone (1997) for many other proposed measures for inputs and outputs in the higher education sector). The third activity is the most difficult to measure, as it accounts for the interaction of the university with third parties such as the general public, companies and firms, governments, and reflects the dual role this interaction has to both parties in terms of productivity and efficiency. Moreover, this third party activity is the one encompassing mostly quality measures, which makes its measurement even more difficult (for quality measurement issues in higher education sector, or in productivity measurement, see Coelli et al (2005), Thanassoulis, Korterlainen, Johnes and Johnes (2010), and the discussion in the next section).

To summarize, Worthighton (2001) identifies three main characteristics of the education process (and of the higher education sector as well), that complicate efficiency and productivity measurement: *First*, the organization under investigation have got multiple objectives, outputs and outcomes. *Second*, clear or unambiguous quantification is not possible for many of the higher education outputs, and *third*, there is limited knowledge of the true relationship between inputs and outputs within a higher education DMU, and full knowledge may never be acquired.

Gates and Stone (1997) provide a set of steps that ought to be followed in order to obtain measures of productivity in the higher education sector. The first and most important step is to clearly define the unit of analysis. Productivity within a university might be related to many of its departments, schools, its library, its administrative services, etc. For example in the case of a university with a very skilled and distinguished engineering department, and a less productive statistics department, the overall university performance is more or less determined by the former , a fact which will be masked if the unit of analysis is the university and not on departments. Thanassoulis (2013) in a speech in a conference about efficiency and productivity in higher education[1], argued that because of the increasing heterogeneity within a university's departments, colleges etc it is better to disaggregate, i.e. choose as unit of analysis a department, or a faculty member instead of the university as a whole. The second step is to define the objectives of the unit, which calls for developing the appropriate performance measure. The third step refers to choosing the appropriate technique to measure productivity. The most widely used technique in the case of higher education is that of Data Envelopment Analysis (DEA). DEA advantages are that (1): it is non-parametric, thus not requiring the specification of a production function (which would require many restricting assumptions regqrding functional forms), and (2) it does not need knowledge of prices for inputs and outputs, a fact that makes it ideal for the (higher) education sector, where the prices of inputs and outputs are very difficult to measure, or even do not exist at all. On the other hand, DEA does not allow for any statistics noise in the estimation, while the parametric-based Stochastic Frontier Analysis (SFA) which assumes a specific production function for the DMUs, does allow for noise.

As for the productivity measurement indexes, there is a vast bibliography going back to Farell (1957) and Malmquist (1953), which provided many alternative indexes.

We present next the most widely used index in the productivity growth studies that make up our sample of 31 papers for the meta-regression analysis, which is the Malmquist TFP index. This non-parametric index has been also applied in many other service industries apart from higer education, including healthcare (Thanassoulis and Maniadakis (2000)) and financial services (Worthinghton (2003)). Following Coelli et al (2005) and Yongmei and Wenyan (2008), the output oriented Malmquist index for

[1] Workshop ''efficiency in higher education'', 24-25 June 2013, Thessaloniki, conducted by the Department of Economics, University of Macedonia.

the i-th DMU, measuring productivity change from period t to period t+1, can be computed as follows:

$$M_i(X_i^{t+1}, Y_i^{t+1}, X_i^t, Y_i^t) = \left[\frac{D_i^t(X_i^{t+1}, Y_i^{t+1})}{D_i^{t+1}(X_i^t, Y_i^t)} \frac{D_i^{t+1}(X_i^{t+1}, Y_i^{t+1})}{D_i^{t+1}(X_i^t, Y_i^t)} \right]^{1/2}$$

where D_i^t and D_i^{t+1} are the output distance functions based on period t and t+1 respectively. The input orientated Malmquist index is computed in a similar way. The arithmetic value of Mi indicates the existence or not of productivity growth from period t to period t+1. A value of Mi > 1 indicates productivity growth (i.e. if Mi=1,032 there is a 3,2% productivity growth from period t to period t+1) while a value of Mi < 1 indicates productivity regress (i.e. if Mi=0,987 there is a 2,3% regress in productivity from period t to period t+1).

II. Meta-regression analysis theoretical framework

Before explaining the term ‘’meta-regression analysis’’, it should be wise to explain that of ‘’meta-analysis’’. Meta-analysis serves as an extensive and quantitative literature review, which purpose is to integrate and explain the various economic results found in a given domain of empirical research. In other words it is a systematic analysis of the effects that different methodologies and study specific characteristics have on a particular domain of interest. Its goal is to successfully address the question of how to obtain an overall test of the magnitude of the relationship being reviewed in a collection –or database- of studies and also the strength of the statistical relationship.

Meta-regression analysis (MRA) is a special form of meta-analysis that was designed to examine empirical research in economics i.e. conduct a ‘’quantitative literature review’’. It is based on econometric techniques and involves a simple regression equation between a variable of interest and a set of meta-independent variables that account for relevant study characteristics, and explain the systematic variation from other results in the empirical literature (Stanley and Jarrell, 1989). The empirical results from each study under examination, become an observation in the sample of all possible empirical results for a particular variable. As notices before, that variable in our case is higher education productivity growth results that come from both published and unpublished papers. In the independent variable side, most of the variables are dummy variables that account for study-specific characteristics.

Stanley (2001) provided a very detailed strategy of how to successfully conduct an MRA. This can be summarized in the following required steps:

1. Include all, if possible relevant studies. That is an attempt of increasing the credibility of the MRA by reducing potential biases caused by non-random selection of studies. In our case the relevant studies were published or written until April 2013.
2. Choose the relevant metric for dependent variable. The term ‘’relevant’’ refers to the most important or informative variable which answers the main questions in the particular subject. In our case the question refers to measuring productivity growth.
3. Choose the moderator or meta-independent variables. These must correspond to important study characteristics that influence the results and explain the variability of the dependent variable across studies. Some of these aspects are as follows (Stanley, Doucouliagos and Jarrell, 2008):

i) Data type and characteristics
ii) Choice of theoretical/empirical model
iii) Quality measures
iv) Geographical orientation of the study
v) Quality of publication
vi) Period covered

4. Conduct a careful econometric estimation
5. Subject the MRA results to specification testing. Issues that have to be considered here are that of heteroscedasticity, novelty or genuine effect, fashion and author effect etc. The most important issue of the above is heteroscedasticity, and this is the only one that we account for.

We further examine an issue related to econometric estimation, that of the relevant weight of different observations. There is a question about whether weighting studies individually or weighting individual studies gives rise to significant differences in the MRA results. This may be important when a study provides more than one estimates of the dependent variable, which may or may not account for alternative specifications and/or estimation techniques.

Stanley (2001) provides three alternative ways to deal with this issue. *First*, we can equally weight individual results by treating each alternative estimate as an observation, even when they come from the same study. This is referred to as the ‘’equally weighted model’’, and it implicitly gives more weight to studies reporting multiple estimates. *Second*, we can equally weight each study by treating it as a single observation regardless of how many estimates it provides. In this case we should use the arithmetic average of the reported estimates, resulting to what we referred to as the ‘’mean-value model’’. *Third*, we can equally weight each study, and use the full load of its reported results, by weighting each result with a particular weight (i.e if the study reports h results then each result will be weighted by a factor of 1/h), resulting to a ‘’paper-weighted model’’. Estimating and comparing the three different weighting methods will provide some insights on whether different weighting results to significant differences in the MRA results.

3. Construction of the database and descriptive statistics

Sources for the meta-regression analysis

Despite the vast literature that exists in the efficiency analysis of higher education institutions which, as mentioned earlier, covers many different analysis techniques (such as DEA, SFA or FDH) and several countries, including multi-country analysis, the literature of studies in productivity of higher education institutions is far more limited. The final number of studies to be considered in this paper is 28, and this includes both published and unpublished/working papers. Also there are two papers [Salleh (2008), Yi-Chung-Hsu and Jen-Chin-Chan (2005)] that are master theses or doctoral dissertations. For collection of the papers for the meta-analysis, we searched for all the papers that included TFP results based on any known TFP index such as Malmquist, Hicks-Moorsten or Luenberger, and this firstly resulted into thirty one (31) papers-studies. From these, we omitted three for reasons of robustness of the results and extreme observations (outliers), and we came up with the final database consisting of 1018 observations.

More specifically, the search domain was vast covering academic journals, university online libraries (such as the Heal-Link in Greece) and of course the entire world wide web, where essays, published or not were uploaded with free access. This search process was done using various keywords such as ‘’efficiency in higher education’’, ‘’productivity in higher education’’ and ‘’Malmquist indexes in higher education’’. In table 1 below we present a list of the academic journals that provided the published studies, including the number of papers published in each. Last but not least in order to select the studies, the following five (5) inclusion criteria were used:

1. The study includes empirical results about the productivity growth of higher education institutions, regardless of its unit of analysis.
2. The study presents empirical productivity scores for the sample it used either by unit, or the mean of them using parametric and/or non-parametric methods (DEA, SFA, and FDH) and **not** some kind of bibliometric study.
3. The TFP index presented in the paper is of the following indexes: Malmquist, Luenberger, Hicks-Moorsten, Fisher, Laspeyres or Paache.
4. The study includes as much information about the unit of its analysis as possible, number of inputs and outputs used, and quality considerations of the author (if any).
5. The study is written in any language that makes it possible to be successfully translated in English so as to extract the required pieces of information from it.

Table 1: Academic journals as sources of published studies

Scientific Journal	No. of studies
Journal of Productivity Analysis	3
Education Economics	2
Higher Education	2
International Transactions in Operational Research	2
Advances in Management and Applied Economics	1
Applied Economics	1
China Economic Review	1
Economics of Education Review	1
International Business & Economics Research Journal	1
International Business Research	1
Journal of Business Management and Applied Economics	1
Journal of Educational Administration	1
Journal of the operational Research Society	1
Managerial and Decision Economics	1
The Manchester School	1
Total published studies	23

This search resulted in 28 papers that covered the inclusion criteria and were used as the database for the final meta-regression analysis. These 28 papers cover the existing bibliography that fulfill the inclusion criteria presented above, until April 2013.

Table 3 below presents all the papers (31) used for the initial analysis, including method used, survey period, TFP index reported, and number of observations extracted from each study. The papers that were omitted are those with a (*) sign. The published papers included in the final analysis are 20, and the unpublished/working papers or theses are 8. The studies cover a wide range of 25 years of survey period, from 1984 to 2009, but the different countries studied are only 15. Table 2 presents a list of the 15 different countries used as subjects in our sample, along with the number of studies for each country. England is the country with the most (6) studies: i.e. Agasisti and Johnes (2009), Flegg, Allen, Field and Turlow (2004), Glass, McKillop and O'Rourke (1998), Johnes (2008), Paterka and Wolszczak-Derlacz (2012) and Thanassoulis, Kotrelainen, Johnes and Johnes (2010). One of the reasons may be data aviliability, as Johnes (2008) noted. Another reason may be that its universities and colleges are of the most famous, oldest and most distinguished around the globe.

Apart from England, Italy, Germany and Spain, there are only few evidence of other European countries, e.g. Sweden and Finland, mainly in cross-country studies. Eastern Europe with the exception of Poland is not represented in our sample,. Other developing countries such as Lithuania or Latvia, or even Russia are not represented either. The USA is the other country with several existing studies for higher education: Sav (2012a, 2012b, 2012c) and Foltz, Braham, Chavas and Kim (2012).

Table 2: List of countries examined in the sample

Country	No. of studies
United Kingdom	6
Italy	5
Germany	4
USA	4
Philippines	4
Spain	2
China	2
Australia	2
Switzerland	2
Austria	1
Sweden	1
Poland	1
Brazil	1
Finland	1
Malaysia	1

The work of Sav (2012b, 2012c) focuses on the decomposition of the TFP Malmquist index on its components and to the impact of the global financial crisis on the productivity of US universities and colleges. More specifically, Sav (2012c) concludes that the US universities and colleges suffered from managerial efficiency regress, which caused the mean TFP index to decrease annually by 2,3% in the period 2006-2009. He attributed this to the effects of the global financial crisis which occurred first in the USA in the end of 2007.

Two more countries are worth to be considered more extendedly, China and Philippines. Philippines are used as subject in the papers of Castano and Cabanda [2007, 2007 (2)], Fernando and Cabanda (2007) and Cuenca (2011). Philippines as a country contributes 104 observations in our sample. The most important issue about this country though, is that is has managed to organize its higher education system similar to some major industrialized countries such as the US and Australia, which unified smaller colleges and universities to bigger ones during the 25-year reform which ended in 1988 [Abbot and Doucouliagos (2001)], and the results indicate similarity in the behavior of the two countries' cumulative TFP indexes.

The connection of Philippines with China is due to their regional similarity (they are both considered to be part of the ''Asia'' region in our study) and as well as their per capita GDP income. Both countries are classified as ''upper middle income'' countries by the World Bank (2012) and together with the paper of Salleh (2012) constitute the body of 168 observations of countries with ''middle income'' classification. As will be discussed with more details in the next subsection, it is interesting to observe the differences of these country observations from those of higher income countries, in terms of mean TFP growth e.t.c.

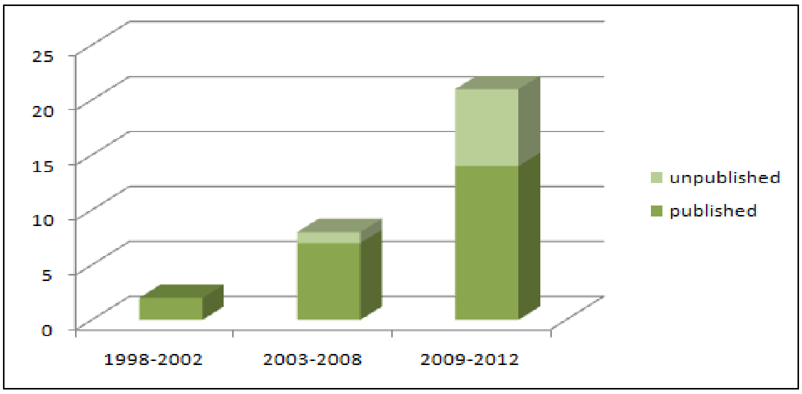

Figure 1: distribution of the sample studies by their publication year.

In the previous chart we present the distribution of the sample studies by their publication year. The last five years there was a steady increase in the number of studies dealing with productivity of higher education institutions indicating the importance of the field and the potential use of their results into policy-making issues.

Table 3: Studies used for the meta-regression analysis

Author and year	No of obs.	Country	Method	Covered period	Unit of analysis
Abbott and Dougouliagos (2001)	26	Australia	DEA	1984-1987	Institution
Agasisti and Dal Bianco (2009)	2	Italy	DEA	2001-2004	Institution
Agasisti and Perez-Esparells (2010)	2	Italy/Spain	DEA	2000-2005	Institution
Agasisti and Johnes (2009)	2	England/Italy	DEA	2002-2005	Institution
Agasisti and Pohl (2012)	122	Germany/Italy	DEA	2001-2007	Institution
Andersson, Antelius, Månsson and Sund (n.p.)	31	Sweden	DEA	2005-2008	Institution
Bolli and Farsi (2011, n.p.)	6	Switzerland	SFA	1996-2007	Department
Castano and Cabanda (2007)	59	Philippines	DEA	1999-2003	Institution
Castano and Cabanda (2007)(2)	30	Philippines	DEA	1999-2003	Institution
Cuenca (2011) (n.p.)	2	Philippines	DEA	2006-2009	Institution
Edward, Fransisco de Sousa and Ramos de Souza (2010)(n.p.)	46	Brazil	DEA	2004-2008	Institution
Fernando and Cabanda (2007)	13	Philippines	DEA	1999-2003	College-School
Flegg, Allen , Field and Turlow (2004)	1	England	DEA	1980-1993	Institution
Foltz, Braham, Chavas and Kim (2012)	2	USA	DEA	1984-1998	Institution
Garcia-Aracil and Palomares-Montero (2008) (n.p.)	172	Spain	DEA	1994-2005	Institution
Glass, McKillop and O'Rourke (1998)	1	England	DEA	1989-1992	Institution
Johnes (2008)	2	England	DEA	1996-2005	Institution
Jose, Simon and Arias (2011)*	68	Spain	DEA	2003-2007	Library
Kempkes and Pohl (2010)	72	Germany	DEA	1998-2003	Institution
Olivares and Schenker-Wicki (2012) (n.p.)	2	Germany/ Switzerland	DEA	2001-2007	Institution
Paterka and Wolszczak-Derlacz (2012)	7	Austria/Finland/ Germany/Italy/ Switzerland/ England/Poland	DEA	2001-2005	Institution
Reichmann and Sommersguter-Reichmann (2010)*	54	Germany/ Austria/ USA	DEA	1998-2004	Library
Salleh (2012)(n.p)	34	Malaysia	DEA	2006-2009	Institution
Sav (2012)	2	USA	DEA	2006-2009	Institution
Sav (2012)(2)	247	USA	DEA	2005-2009	Institution
Sav (2012)(3)	126	USA	DEA	2005-2009	Institution
Thanassoulis, Kortelainen, Johnes and Johnes (2010)	2	England	DEA	2001-2003	Institution
Worthigton and Lee (2008)	5	Australia	DEA	1998-2003	Institution
Yi-Chung Hsu and Jen-Chin Chan (2005) (n.p.)	1	China	DEA	2001-2005	Institution
Ng and Li (2009)	1	China	DEA	1998-2002	Institution
Yongmei and Wenyan(2008)*	28	China	DEA	1999-2002	Institution

Database and variable construction

In general, in a meta-regression analysis, the ''effect size'' is the most crucial variable needed (Haelermans, 2009). In our case this was the productivity growth rate. Given the relatively small number of existing studies and the number of potential dummy variables that would be used for the meta-regression analysis, it was clear that if only the mean TFP change was extracted from each study, there would be little explanatory power in the meta-regression analysis results (although, as mentioned in the introduction, this kind of regression was also run, in the context of Stanley's (2001) suggestion on different weighting methods in a meta-regression analysis).
In order to increase the explanatory power of our empirical analysis, i.e. by improving the degrees of freedom a different kind of structure was used in the construction of the workfile. Observing that many of the studies included university level growth rates it was decided to include the full amount of the reported estimates in a study. In some cases though, despite the fact that the authors reported university level growth rates, there was not distinction between the different units (universities or colleges) with respect to some meta-independent characteristics. For example, one characteristic is that of public versus private units of analysis. If the university level rates where distinguished between public and private ones, all the rates were used. If the distinction was not possible then the mean score that the paper reported was used in the sample but *twice*, to cover both cases.
We resulted in the final number of 1168 observations (1018 after the extraction of three papers in terms of robustness and outlier effects –see next chapter-) in an Excel^{TM} database, and the number of observations per paper varied from 1 to 247. Knowing that this method would unavoidably put more weight to studies with many observations, we also created an Excel^{TM} database with 31 observations-the mean score per study-which was used in order to apply Stanley's (2001) different weighting methods for the observations in the meta-regression analysis.
The set of meta-independent variables was constructed as follows:
i) Estimation method. There are two known ''general'' methods for conducting a productivity analysis. The first is based on some TFP index and the second is based on bibliometric data. We were interested –and used- only studies that had conducted the first, because comparability cannot be assured between a TFP index and a bibliometric score, which is often a per-capita score. As mentioned in the previous

chapter, the methods that our sample studies used were parametric (SFA) and non-parametric (DEA). From the 28 final studies, the 27 used DEA, which is more preferable due to certain advantages mentioned in the previous section, and 1 used SFA [Bolli and Farsi (2011)] Also some of the DEA studies used some kind of bootstrapping method to test for statistical significance of the TFP change, and these studies were distinguished from the other non-parametric ones. Three dummy variables were constructed, named ''DEA'', ''DEA bootstrapped'', and ''SFA'', to account for alternative estimation methods. To avoid collinearity issues deriving from the use of a constant term, we excluded the variable ''DEA'' from the estimation.

ii) Geographic Aspect. The 28 final studies covered up most of the world, except from the continent of Africa. So, there were constructed dummy variables in order to account for different nationality-regionality of the study: One dummy named ''Asia'' accounting for the studies with subject the countries of Australia (3), China (3) and Philippines (4)[2], one dummy named ''EU'' accounting for studies with subject European countries, one named ''USA'' accounting for 4 studies and finally, one named ''South America'' accounting for the study of Fransisco de Souza and Ramos de Souza (2010) which conducted a productivity analysis of the higher education system of Brazil. For collinearity issues we excluded the variable ''USA'' from the estimation

iii) Country income. We distinguished following World Bank (2012) classification in groups of countries: higher, middle and lower per capita income/level of development. In the sample we had 15 countries with higher income (such as USA and most of the EU countries) and middle income (such as China and Philippines). There was not any country with ''lower income'' classification, and so two relevant dummy variables were constructed. The ''higher income'' variable was excluded from the analysis for collinearity issues.

iv) Unit of analysis. The level of aggregation varies considerably from an institution(university or college) as a whole to department. Two sets of dummy variables were constructed:

I) One set of four dummy variables named respectively ''institution'', ''college/school'', ''department'' and ''library''. For collinearity issues we excluded the variable ''university from the estimation. (The last dummy variable was also

[2] (*) number of studies

excluded from the final analysis due to robustness issues, as we will explain further in the next chapter)

II) One set of two dummy variables separating aggregate universities and colleges from each other, named respectively ‘’aggregate university’’ and ‘’aggregate college’’. Both these dummies took the value of zero when the unit of analysis was other that aggregate (unit of analysis university, in the previous distinction) in order to ensure distinction and comparability.

v) Ownership status. This was also an important issue to study, as the private universities and colleges are frequently targeted for lower productivity with respect to public ones, as they unavoidably put more weight in profit maxamisation rather than productivity. One dummy was constructed, named ‘’private’’ to account for observations corresponding to private units of analysis.

vi) Publication status. A very familiar term in that context is that of ‘’publication bias’’. To explain it further, much research work has been done in order to reveal if published papers tend to overestimate real research results. In other words, there is a growing interest on whether the papers that are finally accepted for publication are only those which present results that tend to confirm economic theory and/or past published work. One dummy was created in order to try and answer to that question, in the context of higher education productivity growth studies. The dummy, named ‘’not published’’, accounted for the non-published studies that were part of our sample.

vii) Quality adjustment. As we are considering institutions of higher education, quality is a crucial issue to be concerned with. As Coelli et al (2005) points, in the case of decision making units which are ‘’service industries’’ in general, quality is a multi-dimensional phenomenon associated with a commodity or service produced and must be taken of well care. Quality differences in the outputs of universities and colleges do exist, and there isn’t yet some justified way to incorporate them in the outputs and inputs used. Coelli proposes two ways, either incorporating quality directly in the outputs/inputs, i.e. choosing proper outputs and inputs that include some proxy for it- or using some kind of numerical weight to outputs of different quality. All of the studies that constitute our sample and did apply some quality consideration issues in their analysis used the first way, choosing outputs or inputs that were used as quality proxies. One dummy variable was created by us to account as meta-independent variable which would capture possible differences within the productivity growth results of studies with quality considerations over those without, and was named ‘’quality’’.

viii) Type of TFP index. The Malmquist index was, as mentioned in the previous section, the most preferred among the studies that constitute our sample, mostly due to the fact that it requires only quantity data. This advantage makes it more secure in applications concerning academic units, as Karagiannis and Pantzios (2002) point. In 27 out of 28 final studies, the Malmquist TFP index was used to present the productivity growth results, while Salleh (2012) presented his results using both Malmquist and Hicks-Moorsten index, and Foltz, Braham, Chavas and Kim (2012) used the Luenberger productivity indicator. Two dummy variables were constructed to account for the observations presented via Hicks-Moorsten and Luenberger indexes, while the effect of the Malmquist index was included in the constant term of the meta-regression.

ix) Period covered by the study. We created dummy variables to account for the decades of 1980-1990, 1990-2000 and 2000-2010, in order to capture the survey periods of all the studies. As survey period is considered the period for which the inputs and outputs of each study were measured. This was an attempt to identify if the overall productivity was improving as time passed, following the improvements in literature, hardware, software and many other aspects that effect the productivity of a higher education institution.

x) Publication quality. This is an issue that needs to be considered, because ‘’worse’’ papers, tend to be published in lower quality journals. The question of interest was if the quality of the journal that a study is published affects the productivity growth scores that are reported in it. We attempted to unravel this issue, by creating dummy variables accounting for different qualities of academic journals. We used the ranking of journals by the University of Macedonia (2009) that classifies academic journals into 4 categories, A to D with A being the most known or distinguished journals, and D the less ones.

In the previous chart we present the distribution of the sample studies based on the quality of the journal they were published into, if so. We observe that the categories B and D gather the heavy load of the studies, while there is no study published in a category A journal.

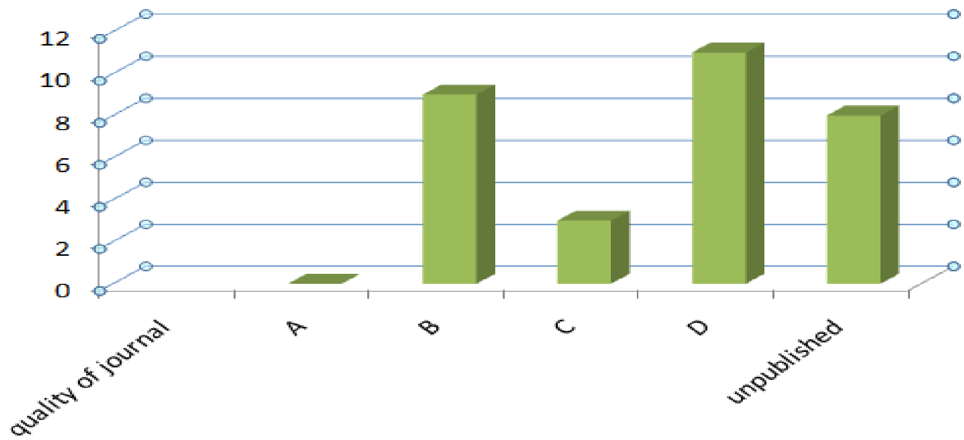

Figure 2: Sample studies based on publication quality

Descriptive Statistics

Before proceeding with the analysis of descriptive statistics, we should point at two things, regarding the exclusion of three studies from the final sample. These three studies are those of Jose, Simon and Arias (2011), Reichmann and Sommersguter-Reichmann (2010), and Yongmei and Wenyan (2008). The first two have got as analysis unit that of university library, and were excluded for issues of robustness of the results. We examined different specifications including or not the studies with libraries, and observed significant changes in the rest of the variables, in terms of sing and significance. The third study was excluded after examining the impact of outliers in the sample, because it presented the 10 most extreme values in the initial sample of 1168 observations, with the biggest ones being 2130,40 % and 911,30%. We resulted in excluding all three of them, to ensure concreteness in our final results. More details about this procedure are in the next chapter.

In the following tables we present the descriptive statistics based on the initial sample of 31 studies, as well as on the final sample of 28 studies.

At first, looking at both of the tables, we observe the large differences in the descriptive statistics between initial and final sample, which made necessary the exclusion of the three above mentioned papers. For example the initial value of the mean TFP growth of the overall sample was 16,36, indicating a productivity growth in the overall sample of 16,36%, whereas the final value in table 5 is 7,49, indicating less than half the growth of the initial value. This is attributed to the extreme values in one of the three excluded papers. The same holds for the mean TFP growth for many of the subsamples. We will analyze further the final descriptives based on table 5. Looking at the statistics for the whole sample, we observe a mean TFP growth of 7,49%.

Table 4: Descriptive Statistics Based on the Initial Sample (31 studies)

		mean	median	max	min	St. deviation	Obs.
Overall sample		16,36	0.90	2130.40	-42.20	87.04	1168
continent	EU	20.03	6.25	280.60	-42.20	38.41	517
	USA	-0.90	-2.55	94.00	-18.00	12.10	406
	South America	-5.15	-5.00	54.00	-41.00	15.51	46
	Asia	47.06	2.20	2130.40	-37.40	196.81	199
Method of estimation	DEA simple	17.37	0.90	2130.40	-42.20	91.68	1036
	DEA bootstrapped	9.15	0.95	162.20	-38.20	32.67	126
	SFA	-5.15	-8.20	8.50	-12.00	7.50	6
Income	Higher	10.59	1.10	280.60	-42.20	31.09	954
	Middle	42.09	0.00	2130.40	-41.00	190.64	214
Level of analysis	Institution	16.13	0.90	2130.40	-42.20	91.33	1027
	Department	-5.15	-8.20	8.50	-12.00	7.50	6
	School/college	-4.79	-3.40	5.60	-14.90	6.29	13
	Library	21.68	7.00	213.00	-38.20	46.92	122
Aggregate unit	University[3]	17.63	0.90	2130.40	-42.20	96.46	888
	College[4]	6.53	0.80	468.10	-16.60	45.88	139
Decade of survey period	Decade 1980-90	4.79	3.80	51.5	-11.1	10.05	30
	Decade 1990-00	26.66	3.40	2130.40	-42.20	134.31	445
	Decade 2000-10	16.68	0.80	2130.40	-42.20	88.15	1138
Public versus private units	Public unit	16.81	0.90	2130.40	-42.20	88.26	1135
	Private unit	1.15	1.10	19.50	-15.50	7.58	33
Other paper aspects	Published paper	20.29	0.00	2130.40	-38.20	100.04	874
	Unpublished paper	4.70	3.85	54.00	-42.20	13.67	294
	Quality considerations	1.11	-2.70	162.20	-17.80	19.80	325
	Not quality cons.	22.25	3.25	2130.40	-42.20	101.14	843
	Malmquist index	16.43	0.90	2130.40	-42.20	87.75	1149
	Other index[5]	12.63	10.30	42.50	-12.50	18.73	19

[3] university or college, as an aggregate unit

[4] school/college, *within* a more aggregate unit

[5] Luemberger or Hicks-Moorsten index

This mean indicates a productivity growth in the overall sample (28 papers) of 7,49%, and it is an expected outcome, since 23 out of 38 studies reported positive mean TFP growth, with the five remaining studies not reporting mean values less than -7%. So, the result of average productivity growth is not expected to change even if we used only the mean values of every study.

We observe significant differences in the mean TFP growth across countries. EU and Asia report productivity growth, while USA and South America report productivity reduction. Almost all the studies concerned with EU countries reported productivity growth, except Thanassoulis, Koretlainen, Johnes and Johnes' (2010), which reported an average of -2,1% for universities and of -11% for colleges of Higher Education. These values, 2 out of 424 concerning EU, did not put any downward pressure in the mean TFP growth.

On the other hand, we observe a mean TFP growth of -2,85% for USA based studies, indicating productivity reduction. Although, this result is fictional and does not indicate that US universities have worse performance over those of EU and Asia. Here we must point out that the Mamlquist TFP index which was mainly used for the studies has the disadvantage that each unit is compared only with the other units from the same study, i.e. each university performance is measured relevant to the performance of other universities of the same study, as Worthington and Lee (2008) point. The results about USA are further explained by the studies of Sav (2012a, 2012b, 2012c). As can be seen from table 3, the studies of Sav were conducted in the period 2005-2009 covering the period after 2007 when the global financial crisis emerged first in the US and then in the rest of the world. As Sav [2012, 2012 (3)] comments, there are dramatic changes in productivity from the academic year 2007-2008 and after, which highly contrast the pre-2007 scores. Sav reports large technological and managerial performance regress, both in public and in private US universities, and the homogeneity of his results is indicated by the standard deviation of 4,89 which is the lowest one in that category of descriptive statistics.

Regarding Asia, which has also got positive mean TFP growth (2,99%) we have to make two comments. *First*, that these results are relative, as mentioned earlier in the text, and concern the performance of Asian universities and colleges with respect to others of the same study. *Second*, the studies regarding Asian universities and colleges where constructed to measure productivity improvements after large educational reforms such as the one that took place in Australia before 1990 [Abbot and

Doucouliagos, (2008)]. That could be a reason that we observe Asian universities surpassing those of other continents, such as USA.

Moving to descriptive statistics by method of analysis, the studies using DEA with a bootstrapping method result generally in lower TFP growth indexes with respect to studies using simple DEA.

Table 5: Descriptive Statistics Based on the Final Sample (28 studies)

		mean	median	max	min	St. deviation	Obs.
Overall sample		7.49	0.65	280.60	-42.20	26.00	1018
continent	EU	19.87	6.30	280.60	-42.20	35.26	424
	USA	-2.85	-2.70	20.60	-16.60	4.89	377
	South America	-5.15	-5.00	54.00	-41.00	15.51	46
	Asia	2.99	1.20	47.80	-37.40	12.29	171
Method of estimation	DEA simple	7.74	0.50	280.60	-42.20	26.57	954
	DEA bootstrapped	4.65	1.50	47.80	-37.40	15.10	58
	SFA	-5.15	-8.20	8.50	-12.00	7.50	6
Income	Higher	8.97	0.95	280.60	-42.20	27.74	832
	Middle	0.85	-0.60	54.00	-41.00	14.35	186
Level of analysis	Institution	7.72	0.80	280.60	-42.20	26.18	999
	Department	-5.15	-8.20	8.50	-12.00	7.50	6
	School/college	-4.79	-3.40	5.60	-14.90	6.29	13
Aggregate unit	University	8.88	0.85	280.60	-42.20	27.88	864
	College	0.30	0.70	19.50	-16.60	5.86	135
Decade of survey period	Decade 1980-90	4.79	3.80	51.5	-11.1	10.05	30
	Decade 1990-00	3.44	2.50	51.50	-42.20	9.11	363
	Decade 2000-10	7.57	0.45	280.60	-42.20	26.33	988
Public versus private unit	Public unit	7.70	0.60	280.60	-42.20	26.37	985
	Private unit	1.15	1.10	19.50	-15.50	7.58	33
Other paper aspects	Published paper	8.62	-0.60	280.60	-16.60	29.51	724
	Unpublished paper	4.70	3.85	54.00	-42.20	13.67	294
	Quality considerations	-2,77	-2.90	51.50	-16.60	6.19	291
	Not quality cons.	11.60	2.90	280.60	-42.20	29.54	727
	Malmquist index	7.39	0.60	280.60	-42.20	26.12	999
	Other index	12.63	10.30	42.50	-12.50	18.73	19

Those with simple DEA resulted to a mean productivity growth of 7,74%, while those with a bootstrapping method to a mean of 4.65%. This is also an expected outcome, since studies using bootstrapping method to account for the statistical significance of the results report insignificant productivity growth results as ‘’zeros’’, i.e. a Malmquist productivity growth of 1.023 (2,3% growth) that is proved insignificant, will be reported by the authors as 1.000 (0% growth), while simple DEA studies will report it as 1.023, resulting generally into higher productivity growth. Though, this difference is not large enough, as it is less than three presentence points.

With regard to the other method, Stochastic Frontier Analysis, we cannot make empirical comparisons with DEA. The reason is that the SFA is used only in one of the 28 studies with 6 observations at departmental level of analysis [Bolli and Farsi (2011)].

Next we examine the statistics for the aspect of analysis level. There are two different categories here. The first separates between different levels within a university/college: aggregate level, ‘school’ level, and departmental level (the university library level was excluded from the final analysis).

With regard to aggregate institutions, departments and ‘’schools’’, we observe that institutions do have a positive productivity growth of 7,72% whereas department and ‘’school’’ levels show a regress of 5,15% and 4,79% respectively. This result can be interpreted as follows: Institutions (universities or colleges) are aggregate units consisting of many schools or departments, which may face positive or negative productivity growth regarding their field of expertiese. If we conduct a productivity growth study at the aggregate level, these data show that we are most probably to face positive growth relative to that we would face if we conducted the same study using departmental or ‘’school’’ data. Although, these results vary a lot compared to the ones of ‘’schools’’ departments, as we can observe from the maximum value of 280,60%, the minimum of -42,20% and the relatively large standard deviation, almost four times that of departmental and school data.

One of the most interesting comparisons is that of aggregate units with each other, i.e. universities with colleges. In table 5 we observe that both categories face a mean productivity growth, but the 8,88% of aggregate universities is more than 8 times than the 0,30% of aggregate colleges. How can we interpret this result? Thanassoulis, Kotrelainen, Johnes and Johnes (2010) offer the explanation of a ‘boundary shift’. In their study the authors reported that, with respect to post-1992 higher education institutions, pre-1992 ones and Guild Higher education colleges experienced a negative boundary shift, i.e. their ‘’mix of outputs’’ moved to a less productive position as academic years moved on. Also, Abbot and Doucouliagos (2001) in their study concerning Australian colleges of higher education they report in their sample during 1984-1987 there was no technical and scale efficiency change in mean, and also that in 10 of the colleges scale efficiency fell during that period.

Concerning studies from countries with different per capita income classification, both per capita income categories achieve productivity growth, with ''higher income country'' studies (such as USA, or the European countries) achieving almost 9 times higher growth than ''middle income country'' studies. What must be kept in mind though, is the high diversity of this category (relative to that of middle income), which is interpreted by minimum and maximum values. As a result, ''middle income country'' studies' estimates appear on average to be more ''concentrated'' around their mean.

We move to the descriptive statistics based on different decades. What should be pointed out here is that some of the observations, due to their sample period, exceed the length of one decade e.g the decade of 90's and enter the next decade too e.g. the 00's. In such cases, the study appears to be conducted in both decades. Moving to the statistics, we observe that the decade of 00's is the ''richest'' one in terms of productivity growth, with a mean productivity growth of 7,57 %. This result leads us to the proposition that the tertiary education did achieve increasing productivity growth as years proceeded. Observing all the decades together, we see a regress of almost 1,5% form 1980's to 1990's, and then a positive growth of almost 4,2% from 1990's to 2000's.

Proceeding to the aspect of ownership status, we observe a higher mean productivity growth of 7,7% in public units relative to a mean 1,15% in private ones. First reason for this result is the difference between the two units' goals. A public higher education institution generally aims in maximizing the ''welfare'' of its participants and the public, while a private one aims in profit maximizing, leaving the goal of productivity (i.e higher degrees for the students, more published researches, better input utilization e.t.c.) in second place. The study of Castano and Cabanda (2007) offers other explanations such as under-utilization of inputs and incorrect selection of input combinations.

Moving on to other paper aspects of the analysis, we differentiate between published and unpublished studies, between those which incorporate quality considerations and those which did not, and between those which used the Malmquist TFP index, and those which used another one.

With regard to published versus unpublished paper results, we observe a mean TFP growth of 8,62% in published ones versus 4,70% in unpublished ones. Since none of the studies makes any kind of comment upon this issue, we can attribute this to an earlier referred concept, that of ‘’publication bias’’, and argue that papers reporting ‘’good results’’ i.e. higher TFP growth for universities and colleges, tend to be more easily published with respect to studies reporting low or even negative productivity results, unless there is some strong argument supporting the results. For example, in Sav’s studies (2012a, 2012b, 2012c) the impact of global financial crisis is used to explain low productivity results and managerial regress in US higher education system.

Next issue to be concerned is that of quality considerations, where we observe that studies without quality adjustment in their input or output choices report a mean of 11,60% productivity growth, versus the negative -2,77% of those studies which argue that their input or output specification strongly incorporates quality. As the data show, it seems that including quality measurement in a TFP growth study, will result in lower presentences. This opens an area of discussion that is far from the objectives of this study, that of proper quality measurement. Thanassoulis, Kortelainen, Johnes and Johnes (2010), Johnes (2008) argue that quality incorporation faces practical difficulties and needs additional data which are often unable to get access to. A common quality consideration effort would be to break down the student enrollment numbers by degree class awarded. This requires comparability between degree classifications which is not obvious in every country higher education system.

In table 6 which follows there is a comparison of the descriptive statistics for the four countries with the most observations (USA, Spain, Germany and Philippines) and UK, which is appears in the sample of most different studies (six). These countries’ observations account for 78,8% of the final sample of 1018 observations. What is important to mention about this table is the concreteness of the results, indicated by the very low standard deviations (the highest being 32,24 of Germany). Germany ranks first among these countries with a mean indicating 23,32% productivity growth, while USA and Philippines experience regress.

Before ending this section, we would like to refer to other two important aspects of productivity analysis that was not able to be captured in this study, and so are to be left for further examination in future ones. The first is that of educational reforms,

Table 6 – Descriptive Statistics for the 4 countries with the most observations plus United Kingdom.

country	**U.S.A.**	**Spain**	**Germany**	**Philippines**	**United Kingdom**
Mean	-2.85	6.60	23.32	-0,09	5,91
Median	-2,70	5.60	6.50	-0,40	1,20
Maximum	26.60	45.10	168.40	29,90	51,50
Minimum	-16.60	-42,20	-16.30	-15,50	-11,10
St. Deviation	4.89	10.04	32.24	6,86	19,25
Obs.	377	173	143	104	8

which was analyzed in earlier section. The second is that of index orientation. The Malmquist TFP index, which was used in 999 observations (98,1% of the total) can be input or output oriented. Although, only some of the studies referred to their choice of orientation, which rendered it impossible to capture this aspect in our meta-regression analysis, and be able to see the possible differences in the resulting TFP growth index of different orientation choices.

4. Estimation Results

There were several issues to be addressed before the estimations were done. At first was the degree-of-freedom problem, which was overrun with the inclusion of as many observations as possible, in the main meta-regression (equally weighted model). The second one was heteroscedasticity. Heteroscedasticity tests [White (1980), Engle (1982)] rejected the homoscedasticity hypothesis, and so heteroscedasticity consistent standard errors were needed to be used. Third, there were collinearity problems among some meta-independent variables that made it necessary for them to be excluded. The dummy variable ‘’department’’ – for observations at a departmental level of analysis – was perfectly collinear with the dummy variable ‘’SFA’’, as they both considered data from the same study. The former was removed.

Another issue we should consider before presenting the results is the distinction of Stanley (2001) between three different weighting methods for the dependent variable which was analyzed in previous section. To revise, Stanley in his paper [Stanley 2001)] proposes three different weighting methods for the dependent variable observations in meta-regression analysis estimation, which we will follow in order to examine any possible differences in the meta-regression estimates due to different weighting method: *First* all observation are treated as equally weighted even if they come from the same study. This is the equally weighted model in tables that follow in this chapter. *Second*, each study is weighted equally by weighting all the observations from the same paper with a particular weight, different for each study. This is the paper-weighted model. *Third*, we equally weight each study by using only one observation per study, and this results in the ''mean-value'' model.

The meta-regression model which was used for the estimations of all the above mentioned regressions in any of the different samples we will consider later, can be written as follows [Karagiannis (2013)]:

$$TFP_i = a_0 + \sum_{i=1}^{j} \beta_j D_{ij} + \mu_i$$

where D_{ij} is a dummy variable for the meta-independent variable j (j=1,2,…,J) and the estimate point I (i=1,2,..., I), β_j gives the impact of the study-characteristic j, and μ_i is the error term. A positive (negative) coefficient β_j can be interpreted as that particular characteristic j is inflating (deflating) the higher education TFP growth estimates.

The last issues we have to consider are (as mentioned in earlier chapter) those of outliers and robustness of the results. We ran the estimations using four different samples, resulting in tables 7, 8, 9 and 10 below.

Table 7 presents the results with the use of the full load of papers, i.e. all 31 of them. For issues of robustness we decided to redo the estimations excluding the papers of Jose, Simon and Arias (2011) and Reichmann and Sommersguter-Reichmann (2010), which focused on university libraries as their unit of analysis. The reason for this was the differences in inputs and outputs. This resulted in table 8. From comparison of the two tables there are significant differences in many coefficients' signs and especially values, thus we decided to extract these papers from the final sample.

For example, the coefficient of ''bootstrapped DEA'' changes sign between the two tables, while the ''South America'' coefficient changes from insignificant to significant in a 5% level.

In order to examine the impact of outliers in the results we decided to redo the estimations extracting Yongmei and Wenyan (2008), which had the most extreme values in the sample. We resulted in table 9 below, which we compared with both tables 7 and 8 and observed large changes in almost all the coefficients' values and signs. For example, the coefficients of ''bootstrapped DEA'' and ''quality'' suffer from a sign change, those of ''college/school'' and ''Asia'' from a significance

Table 7 - estimation results based on the initial sample (31 papers)

category	variable	Equally weighted OLS	Paper weighted OLS	Mean Value OLS
Constant term	(Simple DEA, university,Malmquist index, region: US, decade:00-10)	-198,25 ** (-2.907)	0,33** (3,188)	-130,56 (-0,722)
Method of estimation	Bootstrapped DEA	-0,29 (-0,018)	-25,62 *** (-2,232)	-14,06 (-0,477)
	SFA	10,09 (1.099)	-41,93** (-2,088)	-3,79 (0,117)
Level of analysis (LOA)	Library	97.01 *** (4.717)	6,37 (0.679)	88,88 (1,009)
	College/school	12,69 (0,789)	-2,86 (-0,239)	-26,74 (-0, 593)
	Overall college	-24.49 ** (-2.786)	-14,35 ** (-2,671)	-
income	middle	45,37** (2,116)	3,14 (0,153)	50,68 (1,273)
	lower	-	-	-
Continent-Region (REG)	Asia	62,91 *** (2,742)	2,81 (0,326)	26,35 (0,589)
	EU	122.14*** (4.837)	26,85*** (2,467)	74,19 (1,017)
	South America	45,29 (1,379)	-38,80 *** (-3,466)	-8,86 (-0,231)
Quality adjustment		-84.39 *** (-2.850)	-9,67 (-0,944)	-55,16 (-0,747)
Quality of paper	B degree	-91,46 *** (-4,376)	-13,18 (-1,143)	-26,16 (-0,779)
	C degree	-74.76 *** (-3,350)	-3,07 (-0,171)	-30,49 (-0,606)
Decade of survey period	1980-1990	63,93 ** (2.088)	24,82 *** (2,487)	33,48 (0,808)
	1990-2000	18,60 (1,054)	2,670 (0,204)	3,07 (0,106)
Type of index (TFPI)	Luenberger	107,75 *** (4,256)	-7,26 (-0,394)	60,45 (0,750
	Hicks-Moorsten	-22,69 (-1,478)	-18,18*** (-2,474)	-
Private university		30,48 ** (1,978)	9,88 ** (1,967)	-
Other paper aspects	Number of inputs & outputs	24.22 *** (2.895)	-0,289 (-0,257)	16,43 (0,734)
	Unpublished paper	-42,92*** (-3,417)	16,92 (1,100)	-15,41 (-0,455)
N		1168	1168	31

change, while those of ''South America'' and ''decade 1980'' experience both changes. These issues led to the decision of excluding all three papers, in order to

Table 8 - estimation results after excluding the library papers (29 papers)

category	variable	Equally weighted OLS	Paper weighted OLS	Mean Value OLS
Constant term	(Simple DEA, university,Malmquist index, region: US, decade:00-10)	-361,65*** (-3,625)	0,33 * (3,167)	-153,92 (-0,783)
Method of estimation	Bootstrapped DEA	15,82 (0,83)	-25,81 ** (-2,229)	-10,06 (-0,362)
	SFA	46,66 ** (3.020)	-41,97* (-2,086)	4,04 (0,124)
Level of analysis (LOA)	College/school	18.78 (0,897)	-3,03 (-0,252)	-24,69 (-0,558)
	Overall college	-13,31** (-1,969)	-14,27** (-2,624)	-
income	middle	15,73** (0,447)	3,73 (0,179)	37,98 (1,184)
	lower	-	-	-
Continent-Region (REG)	Asia	103,45 *** (2,638)	2,14 (0,229)	56,00 (0,848)
	EU	197.18*** (4.934)	27,07** (2,435)	92,49 (1,039)
	South America	121.12** (2.315)	-39,38 *** (-3,400)	23,36 (0,403)
Quality adjustment		-170,98*** (-3,705)	-10,19 (-0,970)	-50.98 (-0,722)
Quality of paper	B degree	-67,26*** (-2,844)	-13,10 (-1,129)	-26,47 (-0,769)
	C degree	-119,88*** (-4.055)	-2,33 (-0,126)	-48,91 (-0,822)
Decade of survey period	1980-1990	92.73** (2.101)	25,04 ** (2,488)	23,97 (0,599)
	1990-2000	19.56 (1,075)	3,04 (0,227)	-1,69 (-0,056)
Type of index (TFPI)	Luenberger	161.27*** (3,785)	-7,28 (-0,391)	83,21 (0,850)
	Hicks-Moorsten	-7,61 (-0,377)	-18,08** (-2,449)	-
Private university		32,60 * (1,572)	10,07* (1,953)	-
Other paper aspects	Number of inputs & outputs	44.43 *** (3.623)	-0,314 (0,275)	17,99 (0,768)
	Unpublished paper	-46,93*** (-3,368)	17,11 (1,102)	-21,01 (-0,599)
N		1046	1046	29

reach our final results, which are presented in table 10 below, and which we will examine furtherly, and compare with the descriptive statistics of table 5.

As we can see from the table 10 below, the resulting coefficient sings do not differ so much from the results discussed earlier in the descriptive statistics' chapter. The

Table 9 - estimation results after excluding the outliers' paper (30 papers)

category	variable	Equally weighted OLS	Paper weighted OLS	Mean Value OLS
Constant term	(Simple DEA, university,Malmquist index, region: US, decade:00-10)	63.24 *** (4.908)	0,01 (0,636)	37,60 (1,341)
Method of estimation	Bootstrapped DEA	-34,40*** (-4,921)	-27,09 ** (-2,461)	-6,78 (-0,725)
	SFA	-20,11*** (-5,165)	-43,23** (-2,378)	-14,78 (-1,224)
Level of analysis (LOA)	Library	26,55 *** (3,288)	23,47*** (3,171)	12,74 (0,801)
	College/school	23,57*** (3,047)	7,981 (0,742)	1,27 (0,131)
	Overall college	-7,13 *** (-4,342)	-12,33** (-2,379)	-
income	middle	20,08 (1,191)	-0,38 (-0,021)	21,97 (0,782)
	lower	-	-	-
Continent-Region (REG)	Asia	11,54 (0,704)	3,88 (0,465)	-12,27 (-0,528)
	EU	27,71*** (4.590)	29,38*** (2,721)	13,71 (1,071)
	South America	-31,39* (-1,725)	-22,38 ** (-2,329)	-32,51 (-1,284)
Quality adjustment		28,20 *** (4,712)	-8,07 (-0,834)	5,83 (0,637)
Quality of paper	B degree	-36,89*** (-4,926)	-12,47 (-1,131)	-16,97 (-1,210)
	C degree	3,61 (0,715)	4,12 (0,255)	10,24 (0,408)
Decade of survey period	1980-1990	-27,59 (-1,381)	27,09 *** (3.069)	12,80 (0,684)
	1990-2000	-39,28*** (-6,598)	0,43 (0,036)	-10,47 (-1,104)
Type of index (TFPI)	Luenberger	53,68** (2,263)	-6,33 * (-0,359)	1,57 (0,075)
	Hicks-Moorsten	-15,94 (-1,489)	-6,21 (-1.012)	-
Private university		29,06 *** (3,742)	10,92** (2,199)	-
Other paper aspects	Number of inputs & outputs	-7,84*** (-5,078)	-0,74 (-0,625)	-4,32 (-1,366)
	Unpublished paper	-10,03** (-1,972)	21,40 (1,508)	-6,73 (-0,455)
N		1140	1140	30

existing differences which will be mentioned below can be interpreted to the fact that the MRA examines the conditional inference of the meta-independent variables on the meta-dependent one, while the descriptive statistics examined the sole effect of one meta-independent variable on the dependent one.

In table 10 we present the final results (after excluding 3 papers from our initial sample) of the estimation of the main (equally weighted) meta-regression analysis model, and of the two additional models proposed due to Stanley's (2001) distinction about different weighting methods for the dependent variable observations. The values in parentheses are the t-statistics of the coefficients, and the stars indicate statistical significance at 10% level (one star), 5% level (two stars) and 1% level (three stars). Regarding the equally weighted model, with respect to a study conducted using simple DEA analysis (in a university level of analysis, in US at the decade of 2000-2010, as is presented by our constant term) the use of different estimation method (SFA or DEA with a bootstrapping method) puts a significant and downward pressure to the mean TFP growth result, as both these coefficients are statistically significant and negative. This is a result in accordance with the previously presented descriptive statistics of table 5.

The use of a college/school as unit of analysis instead of an institution inflates significantly the estimated TFP results, but the opposite holds for the use of an overall college as unit of analysis in an MRA study, which appears to deflate the resulting TFP growth significantly. While the latter is in agreement with the descriptive statistics in section 3, the former is not.

To continue, differences in per capita incomedo affect productivity growth in higher education. The latter is inflated if the study is conducted in the European or Asian continent with respect to the continent of the US. The continent dummy of South America does not have a significant effect on the resulting TFP growth as indicated by its t-statistic. Quality adjustment by the author on the input or output side in order to capture some quality aspects in higher education productivity appears to deflate the resulting TFP growth, reinforcing the previous same finding in the descriptive statistics' section, although the result is insignificant, as indicated by the statistic in table 10.

Also, the publication of the study in a B or C degree classification journal, instead of a D degree[6], appears to have a negative and significant effect on the TFP

Table 10 - estimation results using the final sample (28 papers)

category	variable	**Equally weighted OLS**	**Paper weighted OLS**	**Mean Value OLS**
Constant term	(Simple DEA, institution,Malmquist index, region: US, decade:00-10)	-1,64 (-0,159)	0,01 (0,528)	29,53 (0,973)
Method of estimation	Bootstrapped DEA	-21,89*** (-4,498)	-27,38 *** (-3,360)	-5,64 (-0,577)
	SFA	-6,92* (-1,945)	-43,28** (-2,379)	-14,59 (-1,193)
Level of analysis (LOA)	College/school	15,96** (2,499)	7,82 (0,726)	1,69 (0,169)
	Overall college	-5,25 *** (-3,768)	-12,17 ** (-2,314)	-
income	middle	7.76 (0,643)	1,50 (0,079)	18,39 (0,641)
	lower	-	-	-
Continent-Region (REG)	Asia	29,99 *** (2,633)	1,899 (0,215)	-3,20 (-0,115)
	EU	54.57*** (9,262)	29,66** (2,689)	19,55 (1,421)
	South America	4,06 (0.309)	-24,20** (-2,404)	-22,78 (-0,753)
Quality adjustment		-5.05 (-1.017)	-8,71 (-0,877)	6,66 (0,648)
Quality of paper	B degree	-27,26 *** (-4,174)	-12,38 (-1,117)	-17,12 (-1,209)
	C degree	-19,90 *** (-4,562)	6,42 (0,392)	4,49 (0,177)
Decade of survey period	1980-1990	-16,84 (-1,164)	27,44 *** (3,103)	10,11 (0,484)
	1990-2000	-31,01*** (-6,995)	0,85 (0,069)	-11,80 (-1,318)
Type of index (TFPI)	Luenberger	70,31 *** (4,181)	-6,41 (-0,358)	8,73 (0,372)
	Hicks-Moorsten	-7,02 (-0,835)	-6,05 (-0,980)	-
Private university		21,70 *** (3,397)	11,15** (2,189)	-
Other paper aspects	Number of inputs & outputs	0,23 (0,184)	-0,77 (-0.643)	-3,72 (-1,070)
	Unpublished paper	-16,74*** (-4,127)	21,66 (1,511)	-8,45 (-0,599)
N		1018	1018	28

[6] According to the University of Macedonia's Journal classification of 2009, as mentioned in section 3.

growth results. The same holds for the decade dummies. For studies conducted in the decades of 90's and 80's the resulting TFP growth ends up being deflated with respect to the decade of 2000-2010. As we also observe, the inflation is stronger regarding the decade of 90's than 80's, which strongly agrees with the descriptive statistics of table 5. The use of Luenbegrer productivity indicator instead of Malmquist TFP index appears to strongly and significantly inflate the results, while the use of the Hicks-Moorsten index does not appear to have a significant effect on TFP growth with respect to Malmquist index. Private institutions as data of analysis appear to inflate the TFP growth with respect to public ones, and this is a point where the results disagree with the descriptives.

Examining the other paper aspects, *first* the coefficient of ''unpublished paper'' dummy is negative and statistically significant. This indicated that a study which doesn't get published will be more likely to express lower TFP growth results, with respect to others which do get published. This result strongly agrees with the descriptive statistics' section, and reinforces our assumption of the existence of publication bias issues in higher education productivity studies. *Second*, the number of inputs and outputs used doesn't have a significant effect on the TFP growth.

We now turn our attention on Stanley's (2001) distinction and compare with the above analyzed model.

There exist major differences in the paper weighted model with respect to the main one analyzed above (columns 1 and 2 of table 10 above):

a) The effects of different decades change a lot between the two models. The effect of 1990's decade is now positive and significant, instead of negative and insignificant, while the effect of 2000's decade is positive and significant instead of negative and insignificant.

b) The continent coefficient of ''Asia'' along with that of ''college/school'', are now insignificant-instead of significant- but still remain positive.

c) The continent coefficient of ''South America'' suffers big change, as it is now negative and significant, instead of positive and insignificant in the main model.

d) The ''unpublished paper'' effect is now insignificant (and positive), while in the main model this coefficient was deflating the resulting TFP growth statistically significantly. This means that in terms of the paper weighted model the publication bias issues appear to be insignificant, while in terms of the equally weighted model we observed significance of the same coefficient.

Concerning the comparison with the third model Stanley (2001) proposed, the ‘’mean value’’ as we named it, in the latter all the coefficients are rendered insignificant, including the constant term. This resets the proposition in earlier sections about the possible degree-of-freedom problems, and reinforces our choice of using as many as possible observations from each study, and does not allow us to make any other comment regarding coefficients. Regarding signs, this model differs with the main MRA model in the following coefficient signs: ‘’Asia’’, ‘’quality’’, ‘’C degree’’ classification of journal, ‘’decade 1980’s’’ and ‘’South America’’ coefficients.

Concluding on Stanley’s (2001) distinction on three weighting methods for the dependent variable observations for the MRA results, there exist major differences the three proposed models, as between the first two of them (columns 1 and 2 of table 10 above) 11 out of the 19 coefficients suffered from a sign change, a significance change, a value change or all three of them. Regarding the third model we observed a large significance change, as all coefficients are rendered statistically insignificant, accompanied with some sign and value changes, regarding the main MRA model as well as the ‘’paper weighted’’ model too.

Before ending this section we ought to discuss the comparison between the descriptive statistics and the estimation results derived from the main MRA model. As mentioned before, the existing differences can be interpreted to the fact that the MRA examines the combinatorial effect of all the meta-independent variables on the meta-dependent one, while the descriptive statistics examined the sole effect of one meta-independent variable on the dependent one.

By looking at tables 5 and 10 we observe more similarities than differences. Specifically, the results remain almost the same for the categories of different continent, income, aggregate unit of analysis, estimation method, decades of survey, published versus unpublished paper, quality considerations, and level of analysis, except that of school/college, as we observe that the coefficients maintain the same relationships with each other in both tables. For example, the EU-continent-based studies appear to have larger TFP growth than US-based ones in table 5, and this is also confirmed in table 10 where the ‘’EU’’ coefficient is positive and statistically significant. On the other side we observe differences in the categories of diferrent TFP index (the case of Hicks-Moorsten index) and that of pubic versus private unit based studies, where in table 5 public units appear to have larger TFP growth from private ones, while in table 10 the coefficient of ‘’private university’’ is positive and statistically significant in a 5% level, indicating the opposite result.

5. Summary and conclusions

The main objective of this study was to provide a meta-regression analysis in order to explain the variation of total factor productivity growth for studies focusing on the productivity of higher education institutions. After an extended search in many different potential sources of studies, including the internet and online databases, the papers which fulfilled the inclusion-in-study criteria was 31, considering both published papers as well as working papers and theses, covering the existing literature until April 2013. From those, three papers where extracted for robustness and outlier effect issues, which resulted to the final number of 28 studies which yielded 1018 observations.

The second objective of this study was to examine empirically the theoretical proposition of Stanley (2001) about three different weighting methods for the dependent variable observations in an MRA analysis. This study contributes to the existing literature not only by being the first to examine the above issue empirically, but also by being the first conducting a meta-regression analysis on studies focusing on higher education productivity growth, as all the existing meta-regression studies were focusing on higher education efficiency analysis.

The econometric results discussed in the previous chapter suggest that European and Asian universities are doing better in terms of productivity than US universities. Although what we ought to refer to is that in case of US universities there is a large effect from the global financial crisis which started in 2007 from US, and according to Sav (2012b) it resulted in US universities having so low productivity performance. The effect of different estimation methods is inconclusive. Also, adopting quality considerations appear to lead in lower TFP growth, indicating that there exist quality issues in higher education institutions, or that there is need for better proxies for quality capturing in such studies. One more important outcome of this study is that there must be publication bias issues in the field of higher education productivity literature.

The examination of Stanley’s (2001) proposition indicated that the different weighting of the dependent variable observations leads to significant changes in the MRA results, as in the case of equal paper weight (by using the full load of observations) we observed many differences with the main MRA model, both major and minor. In the case of equal paper weight by using the mean observation from each study, all coefficients were rendered insignificant.

References

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., & Battese, G. E. (2005). An Introduction to Efficiency and Productivity Analysis (2nd ed.). New York: Springer

Farrell, M. (1957). The measurement of productivity efficiency. J Roy Stat Soc, 120, 253-290.

Gates, S. and A. Stone (1997). Understanding productivity in higher education, RAND Instititute on Education and Training.

Karagiannis, G. Meta-regression Analysis of Agricultural Productivity Studies, May 2013.

Stanley, T.D. (2001). Wheat from chaff: meta-analysis as quantitative literature review, Journal of Economic Perspectives, 15, 131150.

Stanley, T.D. and S.B. Jarell. Meta-regression analysis: A quantitative method of literature surveys, Journal of Economics Surveys, 1989, 3, 54-67.

Worthignton, A.C. (2001). An empirical survey of frontier efficiency measurement techniques in education, Education Economics, 3, 2001.

Καραγιάννης Γ. και Χρ. Ι. Πάντζιος, Αξιολόγηση της Αποδοτικής Λειτουργίας των Ερευνητικών Μονάδων του ΕΘΙΑΓΕ, Εκδόσεις ΕΘΙΑΓΕ, 2002.

List of studies included in the meta-regression analysis (the papers with * where extracted from the final analysis)

Abbot, M. and C. Dougouliagos. Total factor productivity and efficiency in Australian colleges of advanced education. *Journal of Educational Administration*, 2001, 39, 384–393.

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

Agasisti, T. and C. Perez-Esparrells, Comparing efficiency in a cross-country perspective: the case of Italian and Spanish state universities, *Higher Education*, 2010, 59, 85-103.

Agasisti, T. and G. Johnes. Beyond frontiers: comparing the efficiency of higher education decision-making units across more than one country, *Education Economics, 2009,* 17, 59–79.

Agasisti, T. and C. Pohl. Comparing German and Italian Public Universities: Convergence or Divergence in the Higher Education Landscape?, *Managerial and Decision Economics 2012*, 33, 71-85.

Andersson, C., J. Antelius, J. Månsson and K. Sund, Technical Efficiency and Productivity at Higher Education Institutions – Some problems and some solutions , Swedish National Audit Office, Linnaeus University, Sweden.

Bolli, T. and M. Farsi , The dynamics of labor productivity in Swiss universities, *KOF Working Papers 278,2011*, Swiss federal institute of technology, Zurich.

Castano M. N. and E.C. Cabanda, Sources Of Efficiency And Productivity Growth In The Philippine State Universities And Colleges: A Non-Parametric Approach, *International Business & Economics Research Journal* , 2007, 6, 79-90.

Castano, M. N. and E.C. Cabanda, Performance evaluation of the efficiency of Philippine Private Higher Educational Institutions: application of frontier approaches, *International Transactions in Operational Research*, 2007, 14, 431-444.

Cuenca, J. S., Efficiency of State Universities and Colleges in the Philippines: a Data Envelopment Analysis, Phillipine Institute for development studies, Discussion Paper series No. 2011-14, 2011.

Edward C. , R. Francisco de Sousa and H. Ramos de Souza The productive efficiency of federal institutions of Brazilian higher education, , Qualidade do Gasto Público, Finanças Públicas – XV Prêmio Tesouro Nacional – 2010.

Fernando, B. I. S. and E. C. Cabanda, Measuring efficiency and productive performance of colleges at the university of Santo Tomas: a nonparametric approach, *International Transactions in Operational Research,* 2007,14, 217-229.

Flegg A. T. , D. O. Allen , K. Field and T. W. Thurlow, Measuring the efficiency of British universities: a multi-period data envelopment analysis, *Education Economics*, 2004, 12, 231-249.

Foltz, J., B. Barham, J. Chavas and K. Kim, Efficiency and technological change at US research universities, *Journal of Productivity analysis*, 2012, 37, 171–186.

García-Aracil, A and D. Palomares-Montero, Evaluation of Spanish Universities: Efficiency, Technology and Productivity Change, Paper presented in the Prime-Latin America Conference at Mexico City, September 24-26 2008.

Glass J.C., D. Mckillop and G. O' Rourke, A Cost Indirect Evaluation of Productivity Change in UK Universities, *Journal of Productivity Analysis*, 1998,10, 153-175.

Haelermans, C., A meta-regression analysis of education efficiency scores, TIER working paper series, 2009.

Johnes, J., Efficiency and productivity change in the English higher education sector from 1996/97 to 2004/5, *The Manchester School*, 2008, 76, 653–674.

Jose S. , C. Simon and A. Arias, Changes in productivity of Spanish university libraries, *Omega,* 2011, 39, 578–588. **(*)**

Kempkes G. and C. Pohl, The efficiency of German universities – some evidence from nonparametric and parametric methods, *Applied Economics*, 2010, 42, 2063–2079.

Olivares M, and A. Schenker-Wicki, The Dynamics of Productivity in the Swiss and German University Sector: A Non-Parametric Analysis That Accounts for Heterogeneous Production, UZH Business Working Paper Series, Working Paper No. 309, 2012.

Parteka, A. and J. Wolszczak-Derlacz, Dynamics of productivity in higher education: cross-european evidence based on bootstrapped Malmquist indices, *Journal of Productivity analysis,* 2012, 38.

Reichmann G. and M. Sommersguter-Reichmann, Efficiency measures and productivity indexes in the context of university library benchmarking , *Applied Economics,* 2010, 42,311-323. **(*)**

Salleh, M. I., An empirical analysis of efficiency and productivity changes in Malaysian public higher education institutions, Doctor of Philosophy thesis, School of economics, University of Wollongong, 2012, http://ro.uow.edu.au/theses/3708

Sav, Thomas G., Data Envelopment Analysis of Productivity Changes in Higher Education For-profit Enterprises Compared to Non-profits, *International Business Research*, 2012a, 5.

Sav, Thomas G., Productivity Growth and Efficiency Changes in Publicly Managed U.S. Comprehensive Universities: Data Envelopment Analysis and Malmquist Decompositions, *Journal of Business Management and Applied Economics*, 2012b, 4.

Sav, Thomas G., Productivity, Efficiency, and Managerial Performance Regress and Gains in United States Universities: A Data Envelopment Analysis, *Advances in Management & Applied Economics,* 2012c, 2, 13-32.

Thanassoulis, E., M. Kortelainen, G. Johnes and J. Johnes, Costs and efficiency of higher education instistutions in England: A DEA analysis, *Journal of the Operational Research Society*, 2010, 61, 1-16.

Worthington, A. C., and B. L. Lee. Efficiency, technology and productivity change in Australian universities, 1998-2003. *Economics of Education Review*, 2008, 27, 285-298.

Yi-Chung Hsu and Jen-Chin Chan, The operating efficiency of universities and vocational colleges in Central Taiwan: An application of DEA and the Malmquist productivity index, (master thesis at the Chaoyang university of technology, department of accounting).

Ying Chu NG and Sung-ko LI, Efficiency and productivity growth in Chinese universities during the post-reform period, *China Economic Review*, 2009, 20, 183-192.

Yongmei, H. and L. Wenyan, Malmquist Index analysis of the dynamic changes in scientific research productivity of some Chinese universities before and after merger, *Frontiers of Education in China* 2008, 3, 429–447. **(*)**