Cross-Cutting Analysis of Scientific Output versus other STI Indicators

David Campbell, Julie Caruso and Éric Archambault

david.campbell@science-metrix..com; julie.caruso@science-metrix.com; éric.archambault@science-metrix.com
Science-Metrix, 1335 Avenue Mont-Royal Est, Montréal, H2J 1Y6 (Canada)

Abstract

Investigations of existing relationships between R&D inputs and outputs from an econometric perspective have increased in past decades in response to the challenges faced by governments. As they are operating on increasingly tight budgets, governments are looking to maximise returns on investments; furthermore, accountability for public spending has become a primary issue for residents who expect to get the most value for their tax dollars. Most studies of economies and diseconomies of scale in scientific production have been performed with a view to providing evidence-based policy advice that will improve the allocation and management of resources in the research sector and, ultimately, enhance efficiency. This study adds to the growing knowledge base on the factors driving scientific productivity (i.e., the efficiency with which research inputs are converted into research outputs) at the national and regional levels by reporting on the results of an analysis performed using the most comprehensive dataset on STI indicators that is currently available for European Research Area (ERA) countries and NUTS2 regions. Diminishing returns were observed for R&D investment and expenditure indicators, whereas economies of scale were observed for human resource indicators. These results are discussed in light of their implications for research policy.

Introduction

The last two decades have seen a steady rise in the development of science, technology and innovation (STI) indicators. Their use is intended not only to better manage and govern the complex European system but to measure progress towards the achievement of an increasingly wide variety of social and economic objectives. A greater number and variety of actors are now involved in indicator development, contributing new guidelines, new data sources and new areas of inquiry. A major focus of these efforts has been to find appropriate, quantitative statistical tools that are comparable across systems (i.e., countries, regions, sectors, organisations and industries) and that can strike the best balance between internationally comparable and nationally relevant indicators (Edler & Flanagan, 2011). In order to create robust and meaningful measures, ‘positioning’ indicators must also account for the distinct contextual factors that underlie each system, which may be at least as important as formal inputs and outputs to their ultimate performance (Edler & Flanagan, 2011; Lepori, Barré & Filliatreau, 2008). In the face of often considerable underlying conceptual and methodological difficulties, newer indicators must both consider and attempt to confirm the specific drivers of research output and performance of countries and regions.

Unfortunately, an important aspect in the assessment of R&D performance that is often overlooked in bibliometric studies is the link between R&D inputs with outputs. Most of these studies focus on outputs, often as a result of a lack of comprehensive data on inputs, and thereby do not assess which entity is the most efficient at converting R&D inputs into outputs. Nevertheless, the cross-linking of R&D inputs with outputs from an econometric perspective has increased in the past decades, as governments operate on increasingly tight budgets and seek ways to maximise returns on investments, particularly as accountability for public spending has become a primary issue for residents who expect to get the most value for their tax dollars (OECD, 2008).

Most studies of economies and diseconomies of scale in scientific production have been performed with a view to providing evidence-based policy advice that will improve the allocation and management of resources in the research sector, with the ultimate goal of improving efficiency (i.e., productivity) (Bonaccorsi & Daraio, 2005). These studies have used various methods to measure research productivity in S&T systems, including regression analysis (e.g., the knowledge production function, as in Griliches, 1979) and the production frontier approach (i.e., Stochastic Frontier Analysis [SFA] and Data Envelopment Analysis [DEA]). These studies have also been performed at various levels of analysis, from the organisational level (Dundar & Lewis, 1995; Johnes, 1997; Pandit, Wasley & Zach, 2009; Xia & Buccola, 2005) to the country level (Meng, & *al.*, 2006; Rousseau & Rousseau, 1997; Sharma & Thomas, 2008; Wang & Huang, 2007).

This study adds to the growing knowledge base on factors driving the scientific productivity (i.e., the efficiency with which entities are converting research inputs into research outputs) at the national and regional levels by reporting on the results of an analysis performed using the most comprehensive dataset on STI indicators that is currently available for countries and regions (i.e., at the second level of the Nomenclature of Territorial Units for Statistics of the European Union [NUTS2]) of the European Research Area (ERA). Specifically, this study investigates the factors behind the publication output and productivity of countries/regions as revealed through a factorial and regression analysis. Please note that this study only considered one form of scientific output, namely peer-reviewed scientific publications in Scopus. Other forms of research outputs that were not considered include spin offs and scientific advice provided to governments and firms.

Methods & Results

Construction of Datasets

The R&D output indicator that was used to improve the understanding of differences between countries' and NUTS2 regions' scientific output and productivity is the total number of publications indexed in Scopus (the data covers the 2000-2009 period), defined as follows:

> Number of peer-reviewed scientific publications written by authors located in a given geographical entity (e.g., a country, a NUTS2 region). Fractional counting (FRAC) was used. The fractioning was done at the level of author addresses. Ideally, each author on a paper should be attributed a fraction of the paper that corresponds to his or her level of participation in the experiment compared to the other authors. Unfortunately, no reliable means exists for calculating the relative effort of authors on a paper, and thus each author is granted the same fraction of the paper. For example, if a paper is authored by two researchers from the Universidad Complutense de Madrid (UCM), one from the University College London (UCL) and one from the University of Liverpool, each author is given a quarter of the paper. It results that the UCM is attributed half of the papers (two authors), whereas UCL and the University of Liverpool are each attributed a quarter of the paper (each with one author). Similarly, half of the paper is attributed to Spain (two authors) and the other half to the UK (two authors in the UK). If an author has a double affiliation, half of its fraction is given to each of its affiliations. If there are two authors on a paper and one of them (author A) has a double affiliation, there are three addresses. Each of the addresses of author A is given a quarter of the paper, whereas the address of author B is given half of the paper.

To perform the cross-cutting analysis of R&D outputs versus R&D input indicators, the authors identified the type and quantity of data that could potentially be analysed, subject to data availability, in relation to the number of publications. These R&D input indicators fall under four broad categories: R&D investment and expenditure, human resources, innovation and research infrastructure (e.g., number of research infrastructures by domain).

In total, 16 R&D input indicators were considered, although some were not available for analysing NUTS2 regions. Data were downloaded in bulk from Eurostat[1] for 42 countries and 291 NUTS2 regions for which the bibliometric data were available. The downloaded data covered the years 2000 to 2009. The European Portal on Research Infrastructures' Services[2] was also used to download data for research infrastructure indicators. These data were then uploaded onto an SQL server and structured into a relational database that could be linked with the relational database of bibliometric indicators produced for the European Commission's Directorate-General for Research & Innovation (DG Research) (Campbell & *al.*, 2011). These indicators are as follows:[3]

R&D Investment and Expenditure

1. **GERD:** Gross Domestic Expenditure in R&D expressed in millions of PPS at 2000 prices (Source: Eurostat rd_e_gerdsc table [country level] and rd_e_gerdreg [NUTS2 level])
2. **HERD:** Higher Education Expenditure on R&D expressed in millions of PPS at 2000 prices (Source: Eurostat rd_e_gerdsc table [country level] and rd_e_gerdreg [NUTS2 level])
3. **GOVERD:** Government intramural Expenditure on R&D expressed in millions of PPS at 2000 prices (Source: Eurostat rd_e_gerdsc table [country level] and rd_e_gerdreg [NUTS2 level])
4. **BERD:** Business Expenditure in R&D expressed in millions of PPS at 2000 prices (Source: Eurostat rd_e_gerdsc table [country level] and rd_e_gerdreg [NUTS2 level])

Human Resources

5. **Researchers in the Higher Education Sector:** Number of researchers (both genders in all fields) in the higher education sector expressed in head count (Source: Eurostat rd_p_perssci table [country level] and rd_p_persreg [NUTS2 level])
6. **HRST with Tertiary Education:** Number of human resources (both genders in all fields; 15 to 74 years) in science and technology (HRST) with tertiary education (employed) expressed in thousands (Source: Eurostat hrst_st_nfiesex table [country level])
7. **PhD Students:** Number of PhD students (both genders in all fields) participating in tertiary education (ISCED 97: Level 6) expressed in thousands (Source: Eurostat hrst_fl_tepart table [country level])
8. **PhD Graduates:** Number of PhD graduates (both genders in all fields) from tertiary education (ISCED 97: Level 6) expressed in thousands (Source: Eurostat hrst_fl_tegrad table [country level])
9. **Foreign Students in Tertiary Education:** Number of foreign students (both genders in all fields) participating in tertiary education (ISCED 97: Levels 5 and 6) expressed in thousands (Source: Eurostat hrst_fl_tefor table [country level])
10. **Job-to-Job Mobility of HRST:** Job-to-Job Mobility of HRST (25-64 years; Employed) in all knowledge-intensive services expressed in thousands (Source: Eurostat hrst_fl_mobsect table [country level])

[1] http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/bulk_download, visited November 2011.
[2] http://www.riportal.eu/public/ index.cfm?fuseaction=ri.search, visited November 2011.
[3] For more details on these indicators, see Eurostat's metadata at: http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/metadata

Innovation (country level only)

11. **Employment in Technology and Knowledge-Intensive Sectors:** Employment in technology and knowledge-intensive sectors (all NACE activities; all occupations) expressed in thousands (Source: Eurostat htec_emp_nisco table)
12. **VCI (Expansion & Replacement):** Venture Capital Investments (VCI) for expansion & replacement expressed in millions of euro (Source: Eurostat htec_VCI_exre table)
13. **VCI (Buyout):** VCI for buyout expressed in millions of euro (Source: Eurostat htec_VCI_buyout table)
14. **VCI (Early Stage):** VCI for early stage expressed in millions of euro (Source: Eurostat htec_VCI_earl table)

Research Infrastructure (country level only)

15. **Research Infrastructures:** Number of new research infrastructures (Unit = All (no breakdown); Source: http://www.riportal.eu/public/index.cfm?fuseaction=ri.search)
16. **Average Lower Bound of RI investment:** Average lower bound of research infrastructure investment (i.e., for initial construction/setting up) expressed in millions of euro (Source: http://www.riportal.eu/public/index.cfm?fuseaction=ri.search)

Factor analysis for identifying the main dimensions (i.e., factors) among selected STI indicators and the publication output of countries

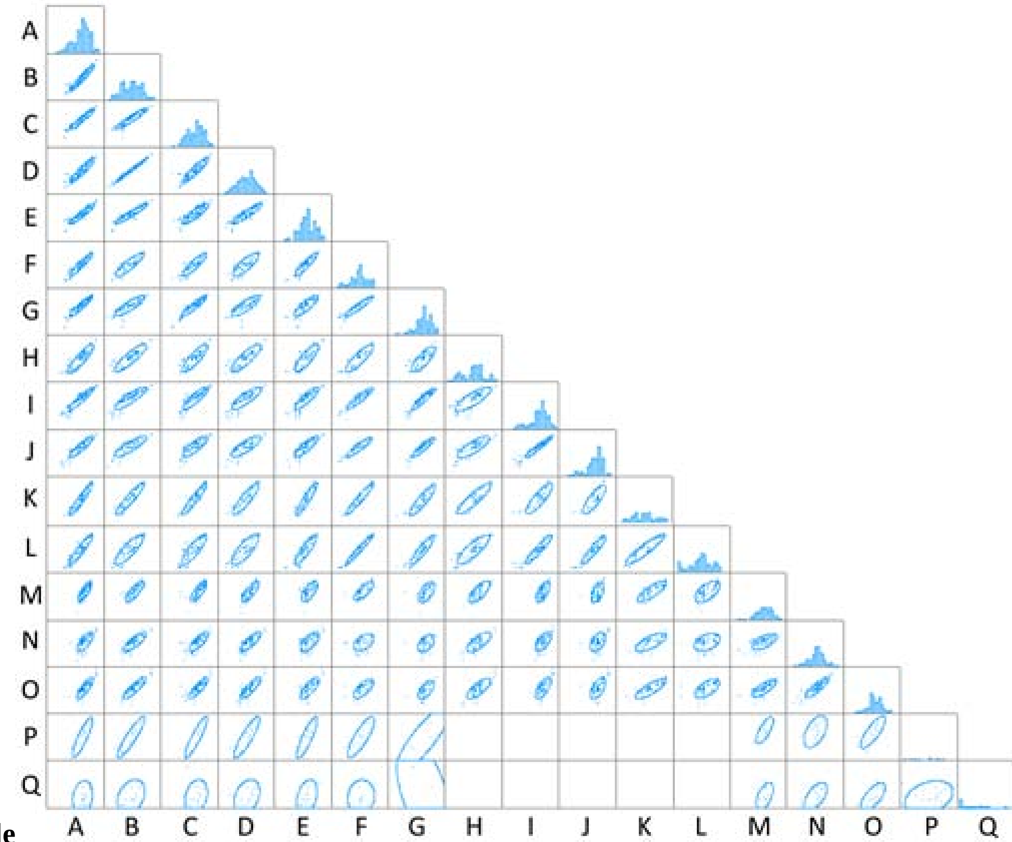
Exploratory Factor Analysis (EFA) was used for identifying the main dimensions (i.e., factors) explaining patterns of variation among selected STI indicators and the publication output (i.e., production) of countries. Prior to performing EFA, the frequency distributions of all indicators were examined to see whether or not the indicators needed to be transformed prior to undertaking the analyses. Indeed, the normality of individual variables is an assumption underlying many factoring methods used in performing EFA. As expected, all indicators had a high positive skew, with most countries having low scores and a few having high scores (e.g., similarly to heavy tails). As such, they were log transformed, and although the resulting distributions did not always satisfy the condition of normality (based on a Kolmogorov-Smirnov Test of normality, data not shown), neither did they appear to present a strong departure from normality (based on a visual inspection of the distributions, see Figure). Additionally, this transformation made (if it was not already the case) the relationships between all observed variables linear, which is another assumption often underlying EFA using different factoring methods (based on a visual inspection of the relationships, Figure 1).

The two indicators on research infrastructure (P and Q in Figure 1) were left aside, as the volume of data available for them was small (N = 31; many missing values). Additionally, GERD was also removed from the analysis because it is redundant with HERD, GOVERD and BERD. Therefore, the remaining indicators included 13 R&D input indicators to be cross-linked with scientific output (i.e., the number of publications). The sample size for each of the 14 variables submitted to the EFA is indicated in the note for Figure , and missing data were dealt with using pairwise deletion.

Because the variables did not fully satisfy the condition of normality, an attempt was first made to perform EFA using Iterated Principal Axis (IPA) factoring, which does not rely on any distributional assumption (Fabrigar & *al*., 1999). Unfortunately, the IPA procedure failed, as the correlation matrix consisted of a singular matrix that could not be inverted. This is due to the very small determinant (0) of the correlation matrix, which provided evidence of high multicollinearity in the dataset (Field, 2000).

Consequently, a second attempt was made using PCA factoring, which relies on the assumption of normality of the variables. Using the Kaiser criterion (i.e., eigenvalue > 1; a factor should be dropped when it carries less information than the average single input variable) for retaining the meaningful factors resulted, in this case, in an overextraction of factors, as is often the case with this approach (Costello & Osborne, 2005).

In identifying the meaningful factors, the scree plot approach was used in conjunction with the Kaiser criterion. Based on this approach, it was found that the 14 selected variables could be adequately summarised using a single factor. Indeed, although the two main factors have an associated eigenvalue greater than 1, there is a sharp break in the distribution of eigenvalues between these two factors, and the first factor alone explains 83% of the variance in the dataset (). In fact, all variables had at least 50% of their variance explained by the first factor, and the output variable (i.e., number of publications) was almost perfectly correlated (R^2 = 0.99) with the first factor (Table 1).



Table

Figure 1. Frequency distribution of selected STI indicators and matrix of the relationships between all pairs of indicators, 2000-2009.

Note: A = Publications (N = 405), B = GERD (N = 339), C = HERD (N = 339), D = BERD (N = 336), E = GOVERD (N = 340), F = HRST with Tertiary Education (N = 218), G = Researchers in the Higher Education Sector (N = 263), H = Foreign Students in Tertiary Education (N = 237), I = PhD Graduates (N = 288), J = PhD Students (N = 262), K = Job-to-Job Mobility of HRST (N = 186), L = Employment in Technology and Knowledge-Intensive Sectors (N = 269), M = VCI (Buyout) (N = 183), N = VCI (Early Stage) (N = 201), O = VCI (Expansion & Replacement) (N = 210), P = Research Infrastructure (N = 31) and Q = Average Lower Bound of RI investment (N = 31)

To assess whether the departure from normality in the selected variables was sufficiently pronounced to create distortion in the results obtained using PCA factoring, an attempt was made to reduce the multicollinearity in the dataset to allow an EFA to be performed using IPA factoring, again using pairwise deletion. This was achieved by removing variables (i.e., job-to-job mobility of HRST, VCI Buyout, VCI Early Stage and VCI Expansion & Replacement) that appeared to cause problems. Among them, the job-to-job mobility of HRST was highly correlated with both the publication output and HERD of countries (data not shown) and was, with these two variables, almost perfectly correlated with the first factor based on PCA factoring. The IPA factoring confirmed the high multicollinearity of the dataset, and it provided results that were highly comparable to those obtained using PCA factoring (data not shown).

Table 1. Factor loadings of selected STI indicators on the first factor of the exploratory factor analysis using PCA factoring.

Indicator	**R**	**R^2**
Publications (FRAC)	1.00	0.99
Job-to-Job Mobility of HRST	0.97	0.94
HERD	0.97	0.95
HRST with Tertiary Education	0.95	0.89
BERD	0.95	0.91
GOVERD	0.94	0.89
PhD Graduates	0.94	0.88
Researchers in the Higher Education Sector	0.92	0.84
Employment in Technology and Knowledge-Intensive Sectors	0.92	0.85
Foreign Students in Tertiary Education	0.91	0.83
PhD Students	0.90	0.81
VCI (Expansion & Replacement)	0.84	0.71
VCI (Buyout)	0.78	0.61
VCI (Early Stage)	0.72	0.52
% of Total Variance Explained by the 1st Factor	**83%**	

All variables were again adequately summarised using a single dimension (only one meaningful factor with an eigenvalue above 1), and the output variable (i.e., the number of publications) was again almost perfectly correlated with the first factor. It should be noted that the dataset used to perform the EFA included repeated measures over time (from 2000 to 2009) for each country. Thus, the assumption of independence in the observations was violated. However, since the goal was to explore the data rather than to perform confirmatory factor analysis (CFA), this violation does not overly offset the main conclusion drawn from this analysis—that the selected indicators are highly collinear.

Nevertheless, given that there are differences in the way countries allocate R&D spending across sectors (e.g., higher education, government, private) and resources (e.g., human resources, infrastructure), it is of interest to investigate how the publication output of countries scale relative to individual R&D input indicators (12 indicators, including those on VCI but excluding the job-to-job mobility of HRST because of its very high correlation with HERD and the response variable). This was achieved through regression analysis.

Regression analysis for investigating the productivity of countries in terms of publication output per unit of the most relevant R&D input indicators

In investigating the impact of individual R&D input indicator on the production of countries (i.e., the output variable), the question arose as to whether the two variables were scaling linearly (i.e., no change in the ratio as one of the variables increases—an isometric pattern) or whether one variable was scaling exponentially relative to the other (i.e., a change in the ratio as one of the variables increases—an allometric pattern). In the latter case, the relationship between the two measured quantities is expressed as a power law:

$$y = kx^a$$

or, equivalently, as a linear relationship using the logarithm of the variables:

$$\log(y) = \mathrm{a} \log(x) + \log(\mathrm{k})$$

where y = output variable and x = explanatory variable.

When attempting to interpret the pattern of change in the ratio between two variables as one increases, estimating the regression coefficient using the logarithm of the two variables can yield the answer (Smith, 2009). In the present case, when the slope of the regression line using the logarithmic form equals 1, there is isometric scaling between the two variables, meaning that the relationship between the two variables is linear (e.g., if y is equal to twice the value of x, it will remain so for any value of x). When the slope of the regression line using the logarithmic form is smaller than 1, there is negative allometric scaling between the two variables, meaning that y increases less rapidly than x (e.g., the ratio of y to x decreases as x increases). In this study, negative allometric scaling will be referred to as "diminishing returns," whereby an increase in a factor of production (i.e., an R&D input indicator) while holding all others constant will yield lower per-unit returns (i.e., peer reviewed publications per unit of the R&D input indicator). Alternatively, when the slope of the regression line using the logarithmic form is greater than 1, there is positive allometric scaling between the two variables, meaning that y increases more rapidly than x (e.g., the ratio of y to x increases as x increases). In this study, positive allometric scaling will be referred to as "economies of scale," whereby an increase in a factor of production (i.e., an R&D input indicator) while holding all others constant will yield higher per-unit returns (i.e., peer-reviewed publications per unit of the R&D input indicator). When the 95% confidence interval of the slope does not overlap with the value of 1, which is indicative of an isometric scaling between the variables, it is concluded that there is a significant allometric scaling indicating either diminishing returns (i.e., a slope smaller than 1) or economies of scale (i.e., a slope greater than 1).

Multiple regression could not feasibly be used to investigate the productivity of countries in terms of outputs (i.e., publications) per unit of the most relevant R&D input indicators due to the high multicollinearity in the dataset. This could have resulted in spurious conclusions regarding the significance of the regression coefficients and led to coefficients of unexpected sign (Zar, 1999). Additionally, the dataset used in this report consisted of a panel data structure made up of cross-sections (i.e., countries) and time-series (i.e., years = 10), the latter being nested within the former. Thus, the input and output variables to be regressed have two dimensions. Each observation has a cross-sectional unit (i.e., country i) and a temporal reference (i.e., year t). The result is that the input and output variables, which consist of time-series, are likely to be autocorrelated as a result of the non-independence in the observations. For instance, the HERD of a country at time t+1 is likely dependent upon its HERD at time t such that it is correlated with itself; in other words, a lagged variable of HERD could likely be regressed with itself (i.e., autoregression). Due to autocorrelations in the data, estimating the regression coefficients on the pooled dataset (i.e., all countries and years) would likely compress the confidence intervals of the slopes, increasing the likelihood of falsely concluding that there are either diminishing returns or economies of scale.

Several regression models have been developed to deal with the peculiarities of panel datasets, in particular with the autocorrelation that often occurs in time-series as well as with unbalanced panels as in this study (i.e., data were not always available for all countries and years among the set of retained variables). Common models in panel data analysis include the fixed-effects model, the between-effects model, the random-effects model and the dynamic panel data model. When the analysis investigates the population response means, as in this study, the need to account for the within cross-section (i.e., country) variation and autocorrelation does not matter as much as when the analysis aims to investigate subject-specific (i.e., country-specific) effects of explanatory variables. In such cases, one can go for robust inference (Gardiner & *al*., 2009) using, for example, a between-effects model, which fits a group-mean regression.

In this study, group-mean regressions were fitted by means of S-estimators (robust regression) (Rousseeuw & Yohai, 1984). This method is adequate for fitting a regression line when outliers might be present in both the response and explanatory variables, which is highly likely with the data used in this report. Furthermore, this regression technique is also robust to violations of the assumptions of normality and homoscedasticity of the residuals, which was the case in several of the fitted regressions. The regressions were performed using a c-value of 2.937, which provided a good compromise between the breakdown point (i.e., the percentage of outliers above which the estimator is likely to be biased; 25%) and efficiency (i.e., 75%) of the estimator (Rousseeuw & Yohai, 1984).

Figure 2 shows the regression coefficients obtained by fitting a group-mean regression between the scientific production (i.e., number of publications [FRAC]) of countries and each of the R&D input indicators retained in the factor analysis in decreasing order of the strength of the correlation between the scientific output of countries and each of these variables (correlation coefficients based on group means ranged from 0.72 to 0.96). In nearly half of the regressions, it was concluded that there was a significant allometric scaling indicative of either diminishing returns or economies of scale.

Significant diminishing returns in terms of publication output are observed for five out of six R&D input indicators related to expenditures (i.e., GOVERD, BERD and all three VCI indicators). Diminishing returns appear to be stronger with the VCI indicators, followed by BERD and GOVERD, whereas there appears to be isometric scaling with HERD. However, there might also be slight diminishing returns with respect to HERD. Indeed, the slope for HERD at the country level is below 1 (i.e., 0.93), and the 95% confidence intervals only slightly overlap with the value of 1 indicative of isometric scaling. This overlap may have been inexistent with a larger system size. As will be shown later with NUTS2 regions, moderate decreasing returns with respect to HERD are confirmed.

Moderate diminishing returns in terms of publication output with respect to the number of students participating in a doctoral program are also very likely. Indeed, the regression coefficient is equal to 0.87, and although the 95% confidence interval for the slope of the regression includes the value of 1, which is indicative of isometric scaling, it does not overlap with it (i.e., it is its upper boundary).
In contrast, significant economies of scale in terms of publication output are observed with employment in technology and knowledge-intensive services, which includes the education sector and all occupations (i.e., professionals, technicians and other occupations). The scaling coefficient is high (1.20), and its confidence interval does not overlap the value of 1. Regarding the number of researchers in the higher education sector, the results at the country level suggest isometric scaling. However, as will be seen later with NUTS2 regions, there appear to be moderate “economies of scale” in terms of publication output as the number of researchers increases. Since this result is based on a much larger system (i.e., there are more NUTS2 regions than countries within the ERA), it is considered more reliable.

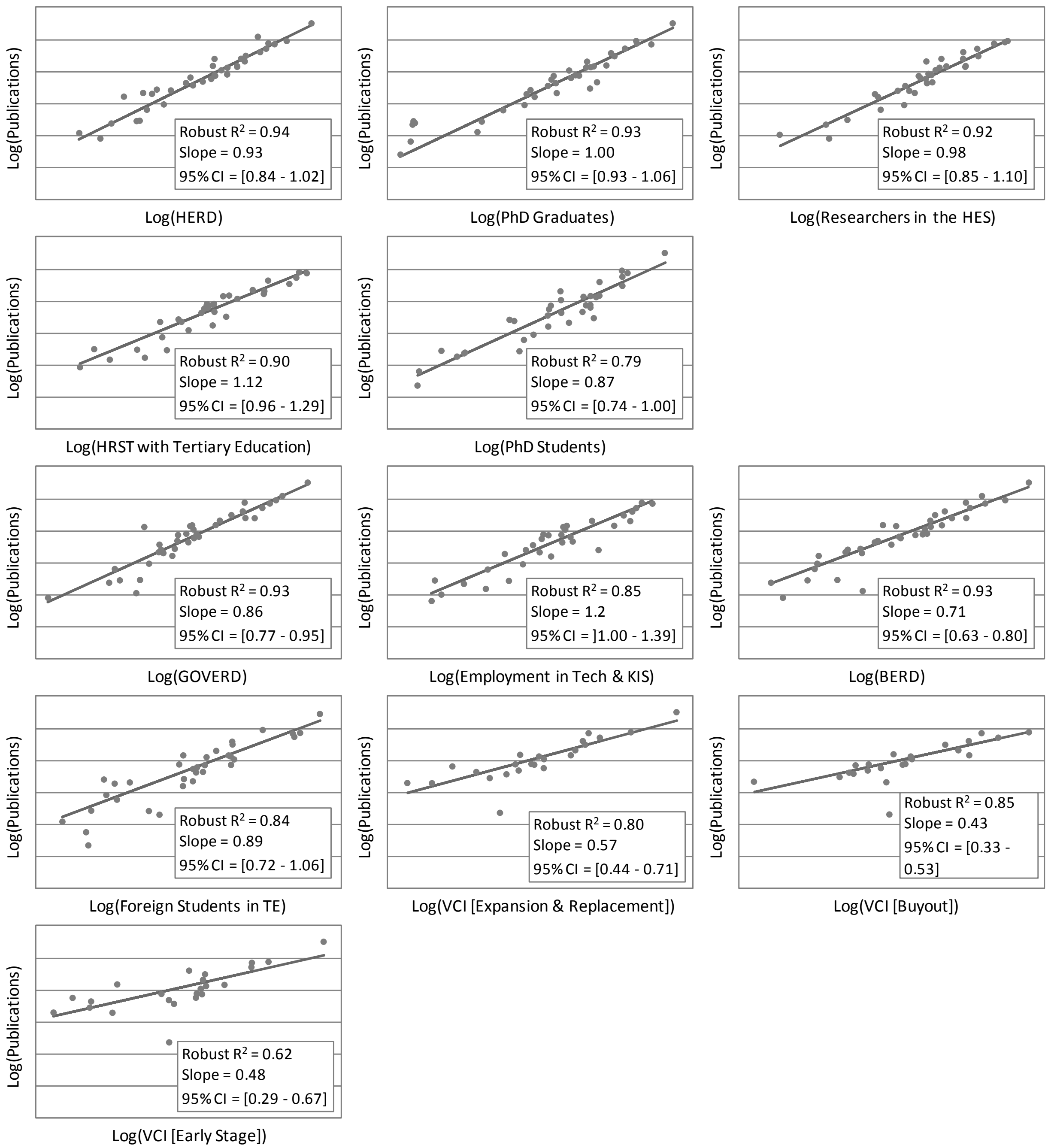


Figure 2. Robust group mean regressions between the scientific output of countries and selected R&D input indicators, 2000-2009.

Regression analysis for investigating the productivity of NUTS2 regions in terms of publication output per unit of the most relevant R&D input indicators

As data were only available for four R&D input indicators at the NUTS2 level (i.e., HERD, BERD, GOVERD and the number of researchers in the higher education sector), the exploratory factor analysis was not performed again, and it was assumed that the four indicators were again highly collinear. Thus, to investigate the productivity of NUTS2 regions in terms of outputs (i.e., publications) per unit of these four R&D input indicators, the same approach was applied to the regions as that applied at the country level.

Again, there appear to be significant diminishing returns in terms of publication output with increases in BERD and GOVERD. Diminishing returns are also significant, although more moderate, in HERD at the NUTS2 level. Similarly to observations at the country level, diminishing returns are strongest with respect to BERD, followed by GOVERD and HERD (Figure 3). Finally, there are significant but moderate economies of scale in terms of publication output with respect to the number of researchers in the higher education sector (Figure 3). It should be noted that the R&D input indicators that have the strongest allometric relationship with the number of peer-reviewed publications are the same indicators that explain the least variation in the scientific production of NUTS2 regions, namely BERD (robust R^2 = 0.54) and GOVERD (robust R^2 = 0.67). These variables are therefore less relevant to the analysis of the scientific output of NUTS2 regions; this is expected, given that they are less tightly linked, conceptually, with this type of output than HERD and the population of researchers in the higher education sector.

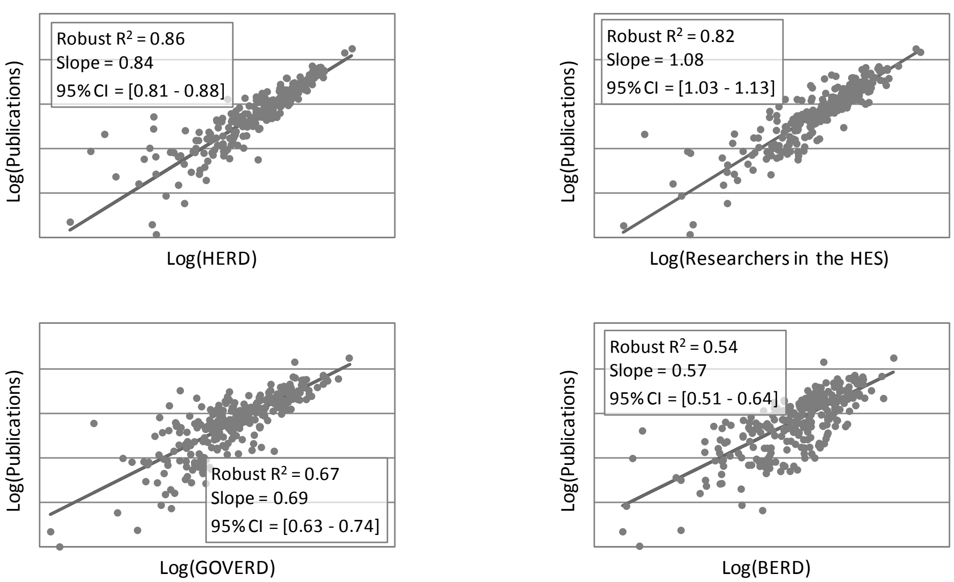


Figure 3. Robust group mean regressions between the scientific output of NUTS2 regions and selected R&D input indicators, 2000-2009.

Discussion

After extensive analyses using EFA, it was found that 15 of the selected indicators could be adequately summarised using a single factor; a single variable (the primary factor) explained 83% of the variance in the dataset. The high level of multicollinearity observed in the dataset does not come as a surprise, as all of the indicators considered are, to varying extent, intrinsically linked with the total R&D expenditures of countries (i.e., GERD). Indeed, as a country increases its investment in R&D, it is likely to gain more resources (e.g., human resources, infrastructure) in S&T, such that other STI indicators are expected to correlate positively with the GERD.

In investigating the impact of individual R&D input indicators on the productivity of countries and NUTS2 regions (i.e., the output variable), economies of scale in terms of publication output at the country level were observed with employment in technology and knowledge-intensive services, which includes the education sector and all occupations (i.e., professionals, technicians and other occupations). Regarding the number of researchers in the higher education sector, the result at the country level suggested isometric scaling. However, there appears to be moderate economies of scale in terms of publication output as the number of researchers in the higher education sector increases at the NUTS2 regional level. As the latter result is based on a much larger system (i.e., there are more NUTS2 regions than countries within the ERA), it is considered more reliable (no data were available for employment in technology and knowledge-intensive services at the NUTS2 level).

Potential mechanisms for explaining the increased productivity of human capital (i.e., employment in technology and knowledge-intensive services and researchers in the higher education sector) as a country's or NUTS2 region's pool of human resources increases include, for example, the diversification and sharing of complementary expertise and competencies, as well as an increase in the specialisation and division of labour. Bonaccorsi and Daraio (2005) also found preliminary evidence of economies of scale as the size of teams in laboratories increases. In terms of policy implications, the authors asserted that specific policies regarding the growth of laboratories within institutions would be required if economies of scale were to be realised through the creation of mega-organisations. This is because the relevant unit through which the mechanisms behind economies of scale would operate, with respect to human resources, is the laboratory.

Significant diminishing returns in terms of publication output were observed for five out of six R&D input indicators related to expenditures (i.e., GOVERD, BERD and all three VCI indicators) at the country level. Diminishing returns were stronger with the VCI indicators, followed by BERD and GOVERD. Similar findings were obtained at the NUTS2 level for BERD and GOVERD (no data were available for VCI indicators at this aggregation level). Diminishing returns are also likely with respect to HERD at the country level and were confirmed at the level of NUTS2 regions.

The observed order in the intensity of diminishing returns in terms of publication output for BERD, GOVERD and HERD does not come as a surprise, as the tradition to publish scientific results in peer-reviewed journals is strongest in the academic sector, followed by the government and private sectors. In fact, the private sector is mostly oriented towards development rather than research, and there are stronger incentives to keep results secret. The fact that the regression coefficients for the VCI indicators are closer to that of BERD than to those of GOVERD and HERD is not unexpected, as both early- and expansion-stage venture capital are captured in BERD, making them somewhat redundant with this indicator.

A potential mechanism for explaining the observed reduction in the productivity of countries and NUTS2 regions in terms of publications produced per euro investment in R&D would be that the number of researchers of a given entity (i.e., its units of production) does not increase as rapidly as its financial resources; as financial resources increase, the maximum production capacity of researchers would be reached. Interestingly, the population of researchers in the higher education sector was shown to scale less rapidly than GERD and HERD (data not shown) at the NUTS2 level. A rationale for awarding smaller grants to a larger population of researchers logically follows from this finding in order to increase the productivity of a given entity as the size of its financial resources increases. However, research teams operating on larger budgets are more likely to carry out projects that could not be conducted with fewer resources (e.g., the Human Genome Project [HGP]). Although the cost of publications produced from such projects likely exceeds that of publications produced by less expensive projects, they impact a much larger community. In turn, these publications are likely to have a higher scientific impact (as measured by citations), such that entities with larger R&D expenditures might generally have a higher scientific impact per euro investment in R&D. In fact, Hung, Lee and Tsai (2009) found that human capital carries more weight in terms of the quantity of academic research, whereas capital accumulation plays a more important role in the citation impact of academic research. Future research efforts will look at how citations scale relative to HERD at the country and NUTS2 levels to test the above hypothesis.

The publication output of countries might also show very slight diminishing returns with respect to one of the selected R&D input indicators in the human resources category, namely the number of students participating in a doctoral program. A potential hypothesis that could explain diminishing returns as the size of a population of PhD students increases would be a concomitant decrease in the amount of researchers per student if the population of PhD students scale more rapidly than the population of researchers in the higher education sector. Indeed, if students receive, on average, less supervision from their thesis director, it seems likely that fewer students would succeed to publish the results of their research. However, based on the data analysed in this study, the population of PhD students appears to scale at about the same rate as the population of researchers in the higher education sector (isometric scaling; data not shown).

The publication output of countries may also show very slight diminishing returns with respect to one of the selected R&D input indicators in the human resources category, namely the number of students participating in a doctoral program. A hypothesis that could potentially explain diminishing returns as the size of a population of PhD students increases would be a concomitant decrease in the amount of researchers per student if the population of PhD students scales more rapidly than the population of researchers in the higher education sector. Indeed, if students receive, on average, less supervision from their thesis director, it seems likely that fewer students would successfully publish the results of their research. However, based on the data analysed in this study, the population of PhD students appears to scale at about the same rate as the population of researchers in the higher education sector (i.e., isometric scaling; data not shown).
Within the research policy context, any attempt at increasing the productivity of a country or region should take account of the complex interplay between the many factors that contribute to their efficiency, such as the country's or region's characteristics (e.g., funding schemes and disciplinary portfolios) and development stages (Leydesdorff & Wagner, 2009). Intervening variables not considered in this study may also be at play, such as languages spoken (English is widely used only in certain European Union countries, and publishing in English may be a particularly relatively new development in former East Bloc countries). The analysis also assumes that all ERA countries have the same scientific/scholarly mix of disciplines. For instance, Archambault and Larivière (2010) showed that the average cost per publication was higher in the basic medical sciences compared to the humanities. Thus, if a larger share of its

R&D budget is allocated to the humanities, a country might exhibit stronger productivity in terms of publications per dollar investment in R&D but lesser productivity in terms of received citations per dollar investment in R&D than another country.

Acknowledgments

The authors are grateful to Carmen Marcus, Grégoire Côté, Guillaume Roberge and Matthieu Delescluse for their thoughtful comments and advice as well as to three anonymous reviewers for their constructive comments on the paper. This study was performed for the European Commission, DG Research and Innovation, Unit C6–Economic analysis and Indicators and funded by FP7. The "Analysis and Regular Update of Bibliometric Indicators" is a study designed, commissioned and owned by the European Commission (Contract No. RTD/DirC/2009/S 158-229751). Any use, right of dissemination or copyright belongs to the European Commission. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Commission.

References

Archambault, É. & Larivière, V. (2010). Individual Researchers' Research Productivity: A Comparative Analysis of Counting Methods. *Book of Abstracts of the 11th International Conference on Science and Technology Indicators* (pp. 22-24). Leiden: Universiteit Leiden.

Bonaccorsi, A. & Daraio, C. (2005). Exploring Size and Agglomeration Effects on Public Research Productivity. *Scientometrics*, 63 (1), 87-120.

Campbell, D., Picard-Aitken, M., Côté, G., Trépanier, M., Ventimiglia, A. & Archambault, E. (2011). *Analysis and Regular Update of Bibliometric Indicators: Country and Regional Scientific Production Profiles*. Report prepared for European Commission, Directorate-General for Research.

Costello, A.B. & Osborne, J.W. (2005). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment, Research & Evaluation*, 10 (7). Available online: http://pareonline.net/getvn.asp?v=10&n=7.

Dundar, H. & Lewis, D.R. (1995). Departmental Productivity in American Universities: Eco-Nomies of Scale and Scope. *Economics of Education Review*, 14 (2), 119-144.

Edler, J. & Flanagan, K. (2011). Indicator Needs for the Internationalisation of Science Policies. *Research Evaluation*, 20 (1), 7-17.

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C. & Strahan, E.J. (1999). Evaluating the Use of Exploratory Factor Analysis in Psychological Research. *Psychological Methods*, 4 (3), 272-299.

Field, A.P. (2000). *Discovering Statistics using SPSS for Windows*. London/Thousand Oaks/ New Delhi: Sage publications.

Gardiner, J.C., Luo, Z. & Roman, L.A. (2009). Fixed Effects, Random Effects and GEE: What Are the Differences? *Statistics in Medicine*, 28 (2), 221-239.

Griliches, Z. (1979). Issues in Assessing the Contribution of Research and Development to Productivity Growth. *Bell Journal of Economics*, 10 (1), 92-116.

Hung, W.C., Lee, L.C. & Tsai, M.H. (2009). An International Comparison of Relative Contributions to Academic Productivity. *Scientometrics*, 81 (3), 703-718.

Lepori, B., Barré, R. & Filliatreau, G. (2008). New Perspectives and Challenges for the Design and Production of S&T Indicators. *Research Evaluation*, 17 (1), 2-3.

Leydesdorff, L. & Wagner, C. (2009). Macro-Level Indicators of the Relations between Research Funding and Research Output. *Journal of Informetrics*, 3 (4), 353-362.

Johnes, G. (1997). Costs and Industrial Structure in Contemporary British Higher Education. *The Economic Journal*, 107 (442), 727-737.

Meng, W., Hu, Z. & Liu, W. (2006). Efficiency Evaluation of Basic Research in China. *Scientometrics*, 69 (1), 85-101.

Organisation for Economic Co-operation and Development (OECD). (2008). *Performance Budgeting: A Users' Guide.* Policy Brief Retrieved from www.oecd.org/dataoecd/53/27/37704120.pdf

Pandit, S., Wasley, C. E. & Zach, T. (2009). The Effect of Research and Development (R&D) Inputs and Outputs on the Relation between the Uncertainty of Future Operating Performance and R&D Expenditures.

Rousseuw, P. & Yohai, V. (1984). Robust Regression by Means of S-Estimators. In . Franke, W. Härdle, and D. Martin (Eds.), *Robust and Nonlinear Time Series Analysis*, Lecture Notes in Statistics No. 26 (pp. 256-272*)*. New-York: Springer-Verlag.

Rousseau S. & Rousseau, R. (1997). Data Envelopment Analysis as a Tool for Constructing Scientometrics Indicators. *Scientometrics*, 40 (1), 45-56.

Sharma, S. & Thomas, V.J. (2008). Inter-Country R&D Efficiency Analysis: An Application of Data Envelopment Analysis. *Scientometrics*, 76 (3), 483-501.

Smith R.J. (2009). Use and Misuse of the Reduced Major Axis for Line-Fitting. *American Journal of Physical Anthropolgy*, 140 (3), 476-486.

Wang, E.C. & Huang, W. (2007). Relative Efficiency of R&D Activities: A Cross-Country Study Accounting for Environmental Factors in the DEA Approach. *Research Policy*, 36 (2), 260-273.

Xia, Y. & Buccola, S.T. (2005). University Life Science Programs and Agricultural Biotechnology. *American Journal of Agricultural Economics*, 87 (1), 229-243.

Zar, J.H. (1999). *Biostatistical Analysis*, Fourth Edition. New Jersey: Prentice Hall.