



Measuring university size: A comparison of academic personnel versus scientific talent pool data

Benedetto Lepori¹ , Lutz Bornmann², and Félix de Moya Anegón³

¹Università della Svizzera italiana Facoltà di comunicazione cultura e società, Lugano, Switzerland

²Science Policy and Strategy Department, Administrative Headquarters, Max Planck Society for the Advancement of Science, Munchen, Germany

³University of Granada Faculty of Sciences, Granada, Spain

an open access  journal



Citation: Lepori, B., Bornmann, L., & de Moya Anegón, F. (2023). Measuring university size: A comparison of academic personnel versus scientific talent pool data. *Quantitative Science Studies*, 4(4), 800–819. https://doi.org/10.1162/qss_a_00273

DOI:
https://doi.org/10.1162/qss_a_00273

Peer Review:
https://www.webofscience.com/api/gateway/wos/peer-review/10.1162/qss_a_00273

Received: 9 March 2023
Accepted: 1 July 2023

Corresponding Author:
Benedetto Lepori
blepori@usi.ch

Handling Editor:
Vincent Larivière

Keywords: academic personnel, institutional size, normalized indicators, scientific talent pool

ABSTRACT

This paper compares two measures of the organizational size of higher education institutions (HEIs) widely used in the literature: the number of academic personnel (AP) measured according to definitions from international education statistics, and the scientific talent pool (STP) (i.e., the number of unique authors affiliated with the HEI as derived from the Scopus database). Based on their definitions and operationalizations, we derive expectations on the factors generating differences between these two measures, as related to the HEI's research orientation and subject mix, as well as to the presence of a university hospital. We test these expectations on a sample of more than 1,500 HEIs in Europe by combining data from the European Tertiary Education Register and from the SCImago Institutions Ranking. Our results provide support for the expected relationships and also highlight cases where the institutional perimeter of HEIs is systematically different between the two sources. We conclude that these two indicators provide complementary measures of institutional size, one more focused on the organizational perimeter as defined by employment relationships, the other on the persons who contribute to the HEI's scientific visibility. Comparing the two indicators is therefore likely to provide a more in-depth understanding of the HEI resources available.

1. INTRODUCTION

Most scholars in management and economics agree that organizational size is a key factor that affects the core characteristics of organizations (Kimberly, 1976), such as their internal governance (Cornforth & Simpson, 2002; Handy, 2007), responsiveness to external pressures (Baumann-Pauly, Wickert et al., 2013), innovation capacity (Damanpour, 1992), and efficiency (Daraio, Bonaccorsi, & Simar, 2015; Gralka, Wohlrabe, & Bornmann, 2019). For higher education institutions (HEIs), it is well known that their sizes are very skewed following the lognormal distribution, as foreseen by Gibrat's law (Vieira & Lepori, 2016).

Unfortunately, data on the size of HEIs, such as personnel or financial resources, have been problematic in terms of availability and comparability (Glänzel, Thijs, & Debackere, 2016; Lepori, Borden, & Coates, 2022). Traditionally, bibliometricians have attempted to circumvent this issue by using relative field-normalized indicators (van Raan, 2004). An example is the popular PP (top 10%) indicator; this is the institutional share of papers that belong to the

Copyright: © 2023 Benedetto Lepori, Lutz Bornmann, and Félix de Moya Anegón. Published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

10% most frequently cited papers in the corresponding disciplines and publication years (Waltman & van Eck, 2015). Yet, empirical evidence has been provided that the science system scales supralinearly with size at both the regional (Nomaler, Frenken, & Heimeriks, 2014) and organizational levels (van Raan, 2013). Accordingly, institutional positions in bibliometric-based rankings strongly correlate with the HEI's resources (Lepori, Geuna, & Mira, 2019).

A few years ago, Abramo and D'Angelo questioned the use of so-called scale-free indicators as indicators of performance (Abramo & D'Angelo, 2016). Despite basically agreeing with this critique, bibliometricians remarked that many issues remain with indicators measuring organizational size (e.g., Glänzel et al., 2016; Zitt, 2016); nevertheless, bibliometricians also argued that they "should improve their knowledge of the current availability of input data" and "explore possible ways of obtaining this data" (Waltman, van Eck et al., 2016, pp. 673–674).

We aim to contribute to this debate by building on recent methodological and data advances concerning the measurement of the number of personnel engaged in HEIs' activities. Compared with university budgets, personnel as a proxy of size has a number of advantages in terms of availability and comparability—although seemingly simple, university budgetary data are plagued with issues related to different accounting systems, the inclusion of ancillary services, and capital costs (Lepori et al., 2022).

We exploit in this respect advances in the availability of data from two sources. On the one hand, institutional data systems (Sivertsen, 2016) have started providing detailed data on academic personnel (AP) by adopting standardized definitions from international statistical organizations (UOE, 2013). Although the data are not free of comparability problems, they are now publicly available for the United States through the Integrated Postsecondary Data System (IPEDS; Jaquette & Parra, 2014) and for Europe through the European Tertiary Education Register (ETER; Lepori, Bonaccorsi et al., 2015). On the other hand, the introduction of unique author identifiers and extensive disambiguation of authors' names in bibliometric databases such as Scopus (Tekles & Bornmann, 2020) has made it feasible to count the number of authors in bibliometric databases affiliated to an HEI. This measure has been named the *scientific talent pool* (STP) and is in use in the SCImago Institutions Ranking (SIR: <https://www.scimagoir.com>; Bornmann, Gralka et al., 2020). For scholarly purposes, as a measure of organizational size, STP has attractive properties, including the ability to measure scientific talent (rather than simply personnel, possibly not engaged in research) and its global availability from a single source (such as Scopus), with potential advantages in terms of comparability.

In this paper, we systematically compare the AP and STP indicators for a large set of European HEIs. Although we expect that they are correlated, as both focus on institutional staff, the availability of several control variables from the ETER database allows us to test expectations concerning the factors that affect their relationship, such as the HEI's level of research intensity, the presence of a university hospital, and the HEI's subject composition. We empirically identify specific country differences as related to the structure of the national HEIs system, as well as outlying cases, which might reveal underlying comparability issues for either indicator.

Our aim is therefore to better understand what AP and STP measure—as related to different institutional characteristics—and to identify those methodological issues that might significantly affect institutional comparisons. We also advance some ideas on the kind of questions these indicators might be best suited to answering, depending on the purpose of the analysis, but also on the characteristics of the compared institutions.

The main contribution of this study is to allow for a sensible and context-aware usage of organizational size indicators to be used for different types of analyses, such as bibliometric

analyses, university rankings, and institutional efficiency analyses. We also provide an understanding of the indicators' respective robustness and possible limitations. Our main conclusion is that, Although no indicator is perfectly comparable, the currently available HEI size indicators can be used at least for the purpose of statistical analysis on a large scale, even if individual data points might still be biased.

2. COMPARING DEFINITIONS

Although counting personnel is seemingly simple, there are a number of issues that might potentially affect indicators computed at the HEI level. Some are general problems of institutional-level data (Lepori et al., 2022), and others are related to the data source and the methodology adopted for each indicator. Their impact might depend on country-specific characteristics and on the HEI's profile.

The main issues in this respect are

- How to identify an HEI unambiguously and to pick the right organizational level,
- How to define the HEI perimeter and decide which units to include or exclude,
- Which people to count, and
- How to count them.

In the following, we first introduce the methodology of AP and STP. We discuss how these issues might affect them and generate differences between the indicators.

2.1. Measuring Academic Personnel

The definition of AP and the approach to its measurement have been codified in the so-called UNESCO-OECD-EUROSTAT manual of educational statistics (UOE, 2013), which constitutes the methodological reference for the HEIs' data collection. AP broadly refers to the individuals within an HEI that are engaged in its core activities (i.e., teaching and research). More precisely, AP includes all individuals whose primary assignment is instruction and/or research, and/or individuals who hold an academic rank with titles such as professor, associate professor, assistant professor, instructor, lecturer, or researcher. AP excludes research and teaching assistants (RTAs), who support professors in their activities, as well as management staff, and administrative and operational personnel.

An important characteristic of the AP definition is that it does not distinguish between personnel engaged in teaching and personnel engaged in research (HEIs personnel frequently engage in both activities). To measure personnel engaged in Research and Development (R&D), the Frascati manual (OECD, 2015) foresees a breakdown based on a time survey of AP. However, only a few countries have engaged in a regular survey, and there is extensive debate on the reliability of the measures obtained (Bentley & Kyvik, 2013; OECD, 2000).

First, the *HEI identification* is usually based on HEIs as legal units identified within national higher education systems. In most cases, this is quite straightforward, but there are also instances of HEIs with multiple campuses, particularly state and private HEIs in the United States. One can observe frequent differences in whether these campuses are considered as a single institution or are separated in databases. Such issues are less widespread in Europe than in the United States, with the exception of France, with its consortia of universities (Lepori et al., 2022). Multilevel structures need to be tracked carefully when combining different data sources in order to make consistent choices, as has been done, for example, in the European Register of Public-Sector Organizations (OrgReg; Lepori, 2020).

Second, the delimitation of the *HEI perimeter* is a more difficult issue, as HEIs frequently cooperate with other research institutions and might own subsidiaries engaged in technology transfer or, as in the United States, in sports and other facilities. Most relevant are linkages of HEIs with hospitals, where different constellations can be found in terms of hiring personnel, financial flows, and publications' affiliations (Calero-Medina, Noyons et al., 2020). In principle, AP measurement is related to the legal unit and to employment contracts. This implies the exclusion of all personnel affiliated with associated centers. Employees of hospitals are usually excluded from HEIs' AP, with the exception of institutional cases where the hospitals are owned by the university (such as in the Netherlands) and national cases where hospital personnel are included in higher education statistics, as in Germany. Most of the research in France is conducted in mixed research units between HEIs and Public Research Organizations (PROs). Although French research output is usually shared between PROs and HEIs, the universities' personnel only include people employed by the university, therefore generating potential mismatches between input and output.

Third, regarding the *people to be counted*, the number of personnel is based on employment contracts and, therefore, excludes categories of people potentially contributing to the HEI's output, such as emeritus professors or visiting researchers. A more important issue concerns the exclusion of research and teaching assistants from AP. A recent OECD survey has identified wide differences in how the AP definition is treated by national statistical authorities, specifically with respect to employed PhD candidates. Some authorities consider them as (partially independent) researchers and include them in AP data. Others exclude PhD candidates because they would purely support senior academics. Given the large number of employed PhD candidates in (research-intensive) universities, this issue might strongly affect comparability of data.

Fourth, with regard to *how to count personnel*, counting can be based on headcounts (HC, one person counts one) or on the average employment rate during the reference year (full-time equivalents). HC should be more comparable to author counts in bibliometric databases (STP), but depends on when people are counted (at the beginning or end of the reference year) and may be affected by large numbers of part-time teachers. Some national authorities exclude people below a certain employment threshold, but there is no common practice in that respect. Moreover, HC is frequently based on functions rather than people, which might generate double counting for people engaged in management roles. An empirical analysis of the ETER data showed that changes in how counting occurs are a frequent source of breaks in AP data series (Lepori et al., 2015). As FTE data are more stable and less impacted by such definitional issues, they might be considered a better measure of the resources available for research. However, they are not yet available for several European countries, including France and Italy.

2.2. Measuring the Scientific Talent Pool Indicator

On publications, authors are connected with certain institutional affiliations. In many cases, this connection can be interpreted as an employment relationship between an author and an institution. Therefore, counting the names of affiliated authors might provide a proxy of the number of employees of an institution.

Based on this reasoning, the SCImago group developed the STP indicator for their institutional ranking of universities and research-focused institutions (SIR, <https://www.scimagoir.com>). The STP indicator is defined as the total number of different authors affiliated with an institution in Scopus documents published during a given period. Indeed, based on data

for North America and Europe, it was shown that the correlation between STP and staff numbers was relatively high (Bornmann et al., 2020). Accordingly, it was proposed that the STP indicator can be used as an input indicator for measuring the efficiency of institutions.

For *HEI identification*, STP relies on name harmonization of research organizations in bibliometric databases. It is well known that HEI names occur in a large number of variants in publication affiliations. Large bibliometric databases invest considerable effort in standardizing organizations' names through features such as the Organization-Enhanced identifier (Web of Science) or Affiliation Identifier (Scopus; Purnell, 2022). Major providers of bibliometric services, such as the Centre for Science and Technology Studies (CWTS) and SCImago build extensive organization databases on top of these identifiers, also relying on manual checks and comparisons with registers such as ROR (see <https://ror.org>; Calero-Medina et al., 2020).

These efforts are relevant for the *definition of HEI perimeter*, as organizational identifiers in bibliometric databases, such as Scopus, also provide extensive information on the HEI's organizational hierarchy and on which subentities are included in the institutional perimeter when matching institutions with publications' affiliation data. This information is routinely updated based on fact-checking (for example from institutional websites) and feedback from the affiliated institutions (Calero-Medina et al., 2020). Although this process allows for detailed understanding of institutional perimeter, it might potentially lead to extended organizational perimeters because HEIs will strive to include as many entities as possible to enhance their visibility. Information on the organizational hierarchy could potentially be exploited to analyze the influence of the inclusion of some units, such as affiliated hospitals, in the organizational perimeter on the STP of the organization.

The STP definition implies that *individuals are counted* if they have published at least one paper in Scopus in the reference year. This poses two problems. On the one hand, guest scientists and other scientists who are only loosely connected to an institution are counted (although they are not in an employment relationship with the institution). This might increase the STP numbers of HEIs. On the other hand, only people publishing in Scopus are counted, implying that people not publishing (but supporting research) and people publishing in or producing outlets not covered in Scopus, such as national-language publications in the humanities (Jacso, 2005), will not be included. Varying coauthorship practices by field, which might affect STP, constitute another issue in this context. In domains with higher average numbers of authors the likelihood of individuals publishing in a certain year (and hence being counted in STP) is expected to be higher (Sivertsen, Rousseau & Zhang, 2019).

Finally, in terms of *how people are counted*, STP relies on full counting, independently from the extent of employment. Therefore, it is the analogue of HC for AP. As STP data are based on affiliation information, counting might be affected by name variants and critically depends on the quality of author disambiguation procedures (Tekles & Bornmann, 2020). Scopus has invested in automated and manual curation of authors' identifiers by connecting other information to sources such as ORCID (Baas, Schotten et al., 2020). Nevertheless, issues might still remain that affect author counts, such as homonyms. If different authors with the same name and affiliation are associated with the same institution, they cannot be differentiated and are counted as one author, thereby decreasing STP figures. As author name changes—name changes of one and the same author—are not considered in STP, they will increase the figures.

2.3. Scientific Talent Pool Versus Academic Personnel

Table 1 summarizes the main methodological issues that might affect AP and STP figures, and their respective relationships.

Table 1. Methodological issues for academic personnel (AP) and scientific talent pool (STP) compared

Issues	AP	STP	Implications
HEI identification	Mostly based on legal structures, problems for multi-campus HEIs (especially United States) and some consortia	Based on harmonization of HEI names in bibliometric databases	Careful comparison required to achieve consistency
HEI perimeter	Based on legal perimeters of HEIs, most affiliated units and hospitals are excluded.	Based on affiliations in papers, more inclusive of associated units and hospitals	Likely to affect comparisons particularly in certain national systems (e.g., France) and when there are associated hospitals
Personnel counted	Based on employment; might be lowered by employment thresholds and exclusion of PhD students	Based on papers' affiliations; excludes nonpublishing personnel in Scopus	Affects comparisons depending on research intensity (PhD students) and subject composition
How to count	Headcounts might be inflated by multiple functions	Counts might be inflated by name changes, but reduced by homonymy	

In general, the previous discussion suggests that AP and STP are correlated because both measure human resources involved in academic activities, as confirmed by previous analyses (Bornmann et al., 2020). We also expect that AP measures in headcounts are associated to a greater extent with STP than measures in FTEs, as, in principle, one author counts one in STP, independently of the employment percentage. However, we also identified several potential sources of differences, and possible explanatory factors for the differences as associated with HEIs' characteristics (Lepori, 2022).

First, AP counts people involved in research and teaching, and STP counts only those publishing in academic journals (covered by Scopus). Therefore, STP might be a subset of AP.

Second, we expected that the relationship between AP and STP is impacted by the research orientation of the HEI considered in this study. Many HEIs have low or limited research activities and, accordingly, their STP is expected to be (much) lower than their AP. In contrast, in research-intensive universities, most AP will also actively publish, including PhD students. This will be the case especially in the sciences and health sciences, where early publishing with the PhD supervisor is the rule. For research-intensive universities, STP is expected to be near to AP or even higher if PhDs are not included in AP, as in many European countries.

Third, we expected differences related to HEI's subject-specific profiles (Thijs & Glänzel, 2008). More specifically, in social sciences and humanities (SSH), not all academics publish internationally (only nationally). This leads to a lower coverage of the HEI's publications in the literature database (Scopus), particularly of the publications in the humanities (Aksnes & Sivertsen, 2019). Therefore, HEIs more oriented towards SSH are expected to have lower STP values than HEIs oriented towards sciences with the same level of AP and research orientation.

Fourth, we assumed that the presence of associated university hospitals affects the relationship between STP and AP. In most cases, we expect that the presence increases the institutional STP to AP ratio. Hospitals associated with universities include many clinical workers who are usually not employed by the university but publish with the university's affiliation

(if they hold an academic title). The presence of hospitals therefore has a large impact on HEI's publications (Elizondo, Calero-Medina, & Visser, 2022).

Fifth, the presence of associated centers at HEIs is assumed to affect the ratio of STP and AP by generally increasing the ratio. In many cases, it is expected that the personnel of these centers are not employed by the HEI and, therefore, excluded from AP. However, researchers at these centers are counted as publishing authors for STP when they indicate HEI affiliations on publications.

3. DATA SOURCES AND VARIABLES

The main data source for this study is represented by the European Tertiary Education Register (ETER; Lepori et al., 2015): the reference database on European HEIs established by the European Commission. ETER provides a wealth of data at the institutional level for more than 3,000 HEIs in 41 European countries (EU-27, United Kingdom, European Economic Area/European Free Trade Agreement countries, as well as candidate and potential EU candidate countries) for the years 2011 to 2020. The data include descriptive information, location, numbers of students and graduates (with various breakdowns), numbers of PhD students and graduates, as well as data on personnel and finances (revenues and expenditures). ETER data are provided by national statistical authorities from the same data collection used for EUROSTAT educational statistics (UOE, 2013). Data are subject to extensive consistency and quality checks. The data set is extensively annotated for previously detected data problems, which could not be resolved with data providers. Coverage of tertiary education as compared with EUROSTAT ranges between 90% and 95% of the enrolled students in most countries. ETER's coverage extends well beyond universities, comprising also colleges, German Fachhochschulen, and other HEIs with limited or no research activity (Lepori, 2022)¹.

For the purposes of this study, the list of institutions from SIR has been matched with the register of public research organizations, OrgReg (Lepori, 2020). OrgReg is a facility that has been developed by the RISIS research infrastructure, which includes nearly 7,000 HEIs, public research organizations, and hospitals in Europe. Orgreg includes more institutions than ETER, such as HEIs nationally recognized, but for which there are no statistical data. It is therefore most suitable for matching. Matching was based on the English institutional name provided by SIR and located country. Proposed matches have been verified manually².

This process yielded 1,607 matched entities out of 1,648 entities included in the SIR for the countries covered in OrgReg. Twenty out of 41 unmatched cases have null STP, and the remaining ones are subsidiaries of U.S. universities, as well as some postgraduate schools and research institutes. The remaining cases account for less than 0.1% of STP. We can conclude that, at least for Europe, there is almost full correspondence between HEIs' delineation in Scopus and the HEIs' institutional databases.

Although only about half of the entities in ETER could be matched to SIR, these comprised 88% of the bachelors and masters students and 94% of the PhD students in 2019. Our sample is therefore largely representative of the full population of European HEIs.

¹ ETER is a public resource available at <https://www.eter-project.com>. Data can be searched and customized through the database interface and downloaded in different formats for further use.

² OrgReg is a public resource, which can be accessed upon registration with the RISIS Central Facility at <https://rcf.risis2.eu/>. Data can be searched by countries and organizational characteristics and downloaded in Excel format.

From ETER, we derived our main variable of interest (i.e., the number of AP in headcounts), but also a number of control variables. We expect these control variables to affect the relationship between AP and STP (see Table 1); the variables have been widely used in characterizing the diversity of HEIs profiles (Huisman, Lepori et al., 2015). The control variables are

- research intensity computed as the ratio between PhD students (level 8 of the International Standard Classification of International Degrees, ISCED) and undergraduate students (diploma, bachelors and masters, ISCED5-7);
- legal status defined in terms of institutional control and/or funding (in EUROSTAT);
- the right to award a PhD as an indicator of a research mandate; and
- the share of undergraduate students in natural sciences, informatics, and engineering (STEM orientation) as well as in health sciences based on the Fields of Education and Training Classification by EUROSTAT (FET).

We additionally included a binary variable for the presence of an affiliated or associated university hospital. The ETER data used in this study refer to the database version downloaded on February 21, 2023 from <http://www.eter-project.com>.

The STP data are derived from the SIR, which is produced by the SCImago Research Group and based on Scopus data.

For all variables, we used data from the year 2018 to maximize availability in ETER. AP data are available for 31 countries and 2,231 HEIs out of 2,851 in ETER. The largest missing countries are Denmark and Finland, for which only data in full-time equivalents are available. STP data are available for 1,380 HEIs in this sample.

3.1. Methods

To analyze the relationships between AP and STP in a first step, we employed descriptive statistics, including correlations and scatterplots. We also compared country totals for both variables to identify specific patterns that may hint at comparability differences related to country-specific factors. We finally performed a preliminary analysis of HEIs where STP is much larger than AP. This analysis might reveal underlying differences in the institutional perimeter between the two indicators. To this aim, we relied on descriptive information, on methodological remarks from ETER, and on a recent OECD analysis of AP data, as well as on detailed information on the HEI structure from the Scopus database.

As a second step, we performed multivariate analysis by regressing STP against AP and other variables that might affect the relationship of STP and AP. As our variables are skewed, we performed log-transformed variables for STP and AP and a square-root transformed variable for research intensity (to avoid dropping all zeros). Descriptive statistics showed that these transformations are effective in reducing nonnormality of the data (Table 2). To account for country-specific issues, we also ran a model with country fixed effects. The full model was specified as follows:

$$\ln(STP) = \alpha + \beta \ln(AP) + \gamma \sqrt{\text{research intensity}} + \delta \text{stemorientation} + \vartheta \text{healthorientation} + \varepsilon \text{ i.legalstatus} + \theta \text{ i.phdawarding} + \pi \text{ i.universityhospital} + \mu_j \text{ country}_j$$

We applied OLS regression with robust standard errors, clustered by country. To address potential collinearity issues, we have computed variance inflation factors (VIF); all VIF in

Table 2. Variables included in the statistical analyses

Variable	Definition	Source	Number of institutions
Academic personnel	Number of academic personnel in headcounts as defined by EUROSTAT	ETER	2,231
Scientific talent pool	Number of individual authors from HEI publications in Scopus	SIR	1,380
Research intensity	Number of ISCED8 students/number of ISCED5-7 students	ETER	2,401
Legal status	0 = under public control or mostly financed by the state 1 = under private control or mostly financed by private sources	ETER	2,799
PhD awarding	0 = no right to award a PhD 1 = right to award a PhD	ETER	2,654
STEM orientation	Share of ISCED5-7 students in natural sciences, informatics and engineering (FET fields 05, 06 and 07)	ETER	1,929
Health orientation	Share of ISCED5-7 students in health and welfare (FET field 08)	ETER	1,928
University hospitals	1 = affiliated hospital available 0 = no affiliated hospital available	ETER	2,836

our models are well below the rule-of-thumb threshold of five, suggesting that multicollinearity is not likely to bias results.

Based on an outlier analysis, we also performed regressions by excluding cases with large STP/AP ratios to avoid the influence of outliers on the results. We analyzed the regression results in terms of level of fit, coefficients, and relationships between STP observed and predicted values in the original scale. We also tested multilevel random intercept models, which provide very similar results to fixed effects models. Models using AP in FTEs were also tested: They provided similar results, but the models were difficult to compare as FTE data are not available for France and Italy. Following the suggestion that the relationships between STP and AP might be influenced by specific disciplinary profiles (Thijs & Glänzel, 2008), we also tested separate models for HEIs specializing in social sciences, humanities, natural sciences, and health sciences: The overall results and level of fit are similar to the general regression, suggesting that for the purposes of this study the effects of different subject compositions are adequately taken into account by the share of students in natural and technical sciences on the one hand, and in health sciences on the other hand.

4. RESULTS

4.1. Descriptive Statistics

Descriptive statistics show that, as expected, our main focal variables are highly skewed, because there are great differences between mean and median (p. 50, Table 3). We can also observe that the logarithmic transformation is effective in reducing skewness, following the well-known fact that organizational size tends to be lognormal (Gibrat's law; Vieira & Lepori, 2016). Research intensity is characterized by a large number of zeros (HEIs not awarding PhD titles) and by the presence of some outliers, notably a few graduate schools included in ETER. Accordingly, we dropped from the regression sample the five observations

Table 3. Descriptive statistics

Variable	N	Mean	SD	Min	p50	Max	Skewness	Kurtosis
Academic personnel HC	2,231	750.01	1,195.16	0.00	299.00	13,771.00	3.68	22.59
ln(academic personnel HC)	2,228	5.64	1.50	0.00	5.71	9.53	-0.08	2.31
academic personnel FTE	1,608	593.62	960.18	0.00	203.64	7,431.66	3.05	14.48
ln(academic personnel FTE)	1,606	5.32	1.56	-0.22	5.32	8.91	0.01	2.25
Scientific talent pool	1,380	877.13	1,449.04	0.00	303.00	12,987.00	3.29	17.65
ln(scientific talent pool)	1,313	5.80	1.53	1.61	5.81	9.47	-0.06	2.36
Research intensity	2,401	0.03	0.17	0.00	0.00	6.83	31.74	1,199.68
sqrt(research intensity)	2,401	0.10	0.14	0.00	0.00	2.61	4.36	57.09
STEM orientation (students)	1,929	0.19	0.24	0.00	0.10	1.00	1.52	4.75
Health orientation	1,928	0.12	0.23	0.00	0.00	1.00	2.49	8.91
Legal status	2,799	0.31	0.46	0.00	0.00	1.00	0.81	1.66
PhD awarding	2,654	0.50	0.50	0.00	1.00	1.00	-0.01	1.00
Scientific talent pool/Academic personnel	1,221	0.84	1.32	0.00	0.52	18.12	6.09	55.99

for which research intensity is greater than one, all of them being graduate schools and research institutes.

The correlation matrix in Table 4 confirms previous insights that STP and AP are correlated with about $r = 0.8$ (Bornmann et al., 2020). Notably, the correlation coefficient is larger with AP in FTEs. However, we have to take into account the large size variation in our sample. Therefore, a high correlation coefficient does not exclude the fact that the ratio of AP and STP varies widely by HEIs. The rather large correlation coefficient between research intensity

Table 4. Spearman correlation coefficients

Variable	Scientific talent pool	Academic personnel HC	Academic personnel FTE	Research intensity	STEM orientation (students)	Health orientation (students)
Scientific talent pool	1					
Academic personnel HC	0.79***	1				
Academic personnel FTE	0.84***	0.96***	1			
Research intensity	0.73***	0.52***	0.57***	1		
STEM orientation (students)	0.26***	0.32***	0.32***	0.06	1	
Health orientation (students)	0.13	0.15***	0.15***	0.01	-0.27	1

*** $p < 0.001$.

and STP suggests that research intensity might strongly affect STP (as expected). We notice that measures of AP in FTEs and HCs are highly correlated.

4.2. National Patterns and Outliers

For the HEIs for which we have AP and STP data, the sum of STP amounts to 76% of AP in headcounts. This can be expected as not all HEI personnel are actively publishing (with papers covered in international literature databases). Always for the same sample, the sum of FTEs is nearly equal to the sum of STP, as STP also counts as one unit publishing academics with a lower employment percentage.

However, a closer view reveals some distinct country patterns (Figure 1). On the one hand, several countries have much lower ratios. Most of them are Eastern European countries. The lower national ratios might be explained by the lower research intensity of the higher education systems compared to Western European countries, and by the tendency of researchers in these countries to publish in national language journals not covered by Scopus (Petr, Engels et al., 2021).

On the other hand, systematic differences are generated for some countries by methodological issues in the AP perimeter in ETER (and EUROSTAT). On the lower end, AP figures are inflated in Switzerland and Germany by the full inclusion of employed PhD candidates in AP—for Switzerland this will be corrected from 2020 onwards in ETER. Ongoing efforts to better harmonize the rules for inclusion/exclusion of employed PhD candidates are therefore highly relevant to improve AP comparability across countries. For Ireland, figures are affected both by lower coverage of AP in ETER (only core staff are included) and, more importantly, by the full inclusion of hospitals in Scopus, which strongly inflates STP figures for the largest HEIs in the country. Similarly, in Austria, higher STP/AP ratios are observed for medical universities but ratios for universities of applied sciences.

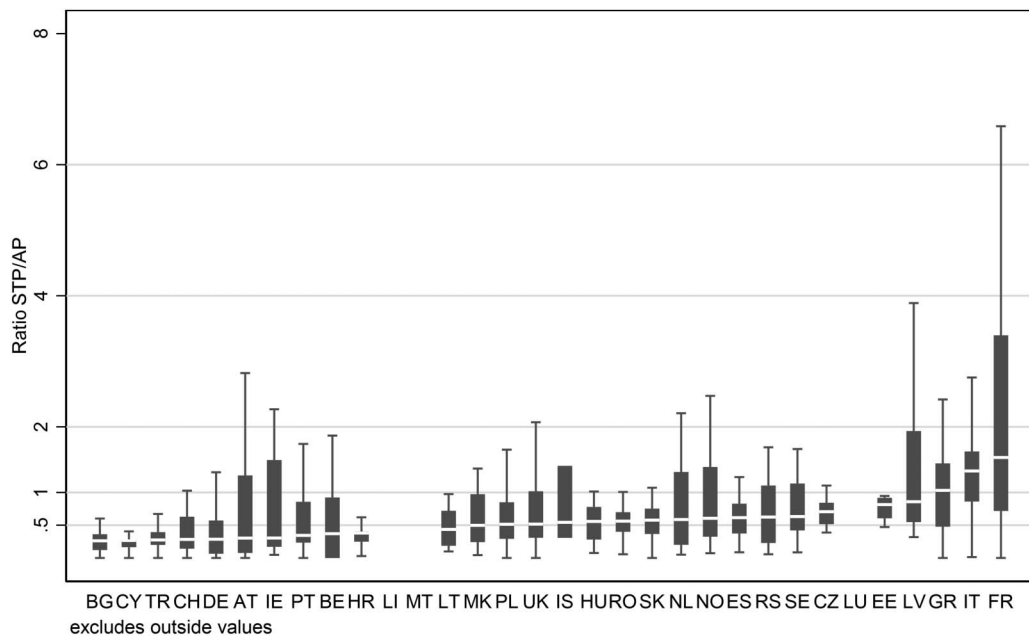


Figure 1. Boxplots of institutional ratios between scientific talent pool and academic personnel by country.

On the other hand, for some countries, STP totals are larger than AP totals. For France, this can be explained by the fact that the ETER perimeter includes only personnel employed by universities, whereas publication affiliations also include personnel working in joint research units with public research organizations, such as Centre National de la Recherche Scientifique (CNRS). As of Italy, ETER data include only permanent personnel, excluding most fixed-term personnel and, therefore, underestimating AP.

Our preliminary analyses suggest that country-specific variations in how AP is counted have an influence on the ratio between STP and AP and lead to systematic patterns. These patterns need to be controlled for in the statistical analysis. Moreover, our analyses provide some preliminary evidence of variation between individual HEIs as related to their research orientation and to the presence of a university hospital.

4.3. Institutional Patterns

In our data, we observe great variation in the ratio between STP and AP within our sample. For about 20% of the HEIs, STP is larger than AP; for nearly half of the HEIs considered, this ratio is below 0.5. There are 135 HEIs where STP is less than one-tenth of AP. Understanding the sources of this variation is the focus of the following empirical analysis.

Figure 2 hints at some systematic patterns. Nearly half of the HEIs have an STP/AP ratio below 0.5. These HEIs comprise one-third of all AP and more than half of the undergraduate students, but only 10% of STP and 20% of PhD students. At the other extreme, the nearly 300 HEIs for which STP is larger than AP comprise half of STP and about one-third of all PhD students in the sample, but only 20% of undergraduate students. These results confirm our expectation that differences in the relationships between STP and AP are closely associated with the research versus education orientation of HEIs. We also observe an association between the presence of university hospitals and high STP/AP ratios.

The group of 135 HEIs for which STP is less than one-tenth of AP is mostly comprised of universities of applied sciences in the Netherlands (Amsterdam UAS, HAN UAS) and Germany (Cologne, Munich, Darmstadt), online universities (Hellenic Open University, UNINETTUNO),

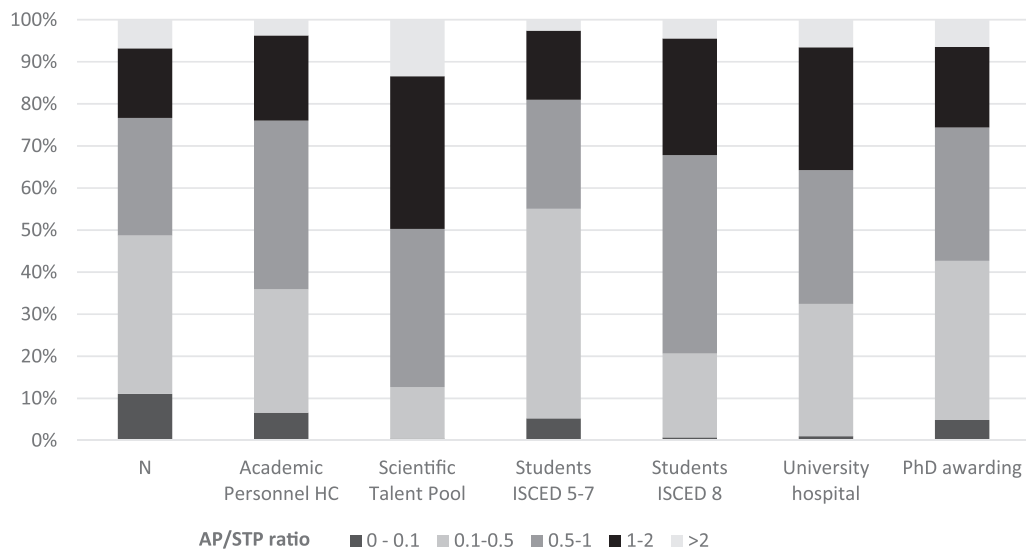


Figure 2. Higher education institutions by ratio between scientific talent pool and academic personnel.

and some HEIs specializing in humanities (University of the Arts, London), as well as universities in Eastern Europe and Turkey with low research activity (Beykent University, University of National and World Economy in Bulgaria). This group therefore reflects the presence in Europe of a sizeable number of HEIs with large educational activities (many HEIs in this group enrolled more than 10,000 students in 2018), but with limited research activity (Lepori, 2022).

The group of HEIs where STP is larger than AP is more heterogeneous; most cases can be related to issues of HEI perimeter discussed in Section 2.3. More precisely, we could identify 42 cases where the ratio between STP and AP is above 3, and 803 cases where the ratio is above 2. As there are 1,221 HEIs for which both indicator values are available, the number of outliers (high ratios) is small. The distribution of the outliers over countries shows distinctive country patterns: 39 cases are in France and 14 in Italy. In the remaining countries, there are only 1–2 cases, suggesting that these are individual issues.

A more precise analysis of these cases is informative to obtain some reasons why STP and AP are strongly deviating (Table 5). First, among the cases with the largest deviation, many can be explained by the existence of university hospitals with their large number of personnel but varying levels of integration in the university structures (Elizondo et al., 2022). A case-by-case analysis in Scopus showed that a number of university hospitals are included in the university affiliation hierarchy, because legally they are part of the respective university (such as the Dutch medical centers). In other cases (such as Swiss university hospitals, which are legally independent) only the medical faculty affiliations are included. Similarly, some specialized

Table 5. Selected outliers with high ratios of scientific talent pool and academic personnel

Name	Country	AP	STP	Ratio	Explanation
Sorbonne University	FR	2,784	12,987	4.66	The figure is inflated by hospitals and by UMR with CNRS
Erasmus University Rotterdam	NL	1,640	5,284	3.22	Erasmus medical center accounts for most of the authors
Trinity College Dublin	IE	743	3,129	4.21	Figure highly inflated by hospitals
Grenoble Institute of Technology (INP)	FR	393	2,555	6.50	STP seems to be inflated; fewer than 1,000 authors in Scopus
University of Liège	BE	697	2,357	3.38	ETER figures underestimated
National Polytechnic Institute of Toulouse	FR	310	1,863	6.01	Some large associated research institutes
West Pomeranian University of Technology, Szczecin	PL	41	743	18.12	Mistaken data in ETER
Campus Bio-Medico University	IT	199	662	3.33	Associated with a large hospital
École Nationale Supérieure de Chimie de Montpellier	FR	47	718	15.28	Some large associated research institutes
Gran Sasso Science Institute	IT	29	258	8.90	Research Infrastructure of the National Institute of Physics
Scuola Normale Superiore, PISA	IT	103	464	4.50	Graduate school
University Centre in Svalbard	NO	31	111	3.58	Arctic research base, mostly external authors

Downloaded from http://direct.mit.edu/qss/article-pdf/4/4/800/2339467/qss_a_00273.pdf by guest on 22 May 2024

universities or research centers associated with larger hospitals reveal a far larger number of authors than (university) AP, such as in the case of the Biomedical University in Rome.

Second, most French institutions are engineering schools that include research centers affiliated to either public research institutes such as the CNRS or to universities in the same region, such as the Polytechnic Institute in Toulouse and the High National School of Chemistry in Montpellier. This phenomenon is slightly less visible for universities, given their large size, with the exception of the Sorbonne University, which includes both affiliated research centers and many hospitals in the Paris region.

Third, a few institutions are entities with specific tasks and structures with a small core of employed personnel but many external affiliations. These include graduate schools such as the Normale Supérieure in Pisa and the Scuola Internazionale Superiore di Studi Avanzati (SISSA) in Trieste, and research infrastructures such as the Gran Sasso Physics Observatory and the Arctic Research Base on the Svalbard Islands.

Fourth, a few institutional cases point to database problems on either source. The STP data for the Grenoble Institute of Technology (Grenoble INP) seem to be inflated, as the HEI has fewer than 1,000 affiliated authors in Scopus. In some cases, it is known that ETER data are misspecified, such as in the case of the University of Liège (only professors are included in AP) and the Technical University in Szczecin (2018 data are misspecified). The fact that we found only a few such cases suggests that the technical quality of both data sets is good and that most of the deviations are related to deeper institutional issues.

4.4. Regression Results

Table 6 presents the results of the regression analyses for three models: (1) an OLS regression for $\ln(\text{STP})$ as a function of $\ln(\text{AP})$, (2) a model including the covariates on the level of HEIs suggested by our previous discussion, and (3) a model including country fixed effects.

Table 6. Results of three regression models with the logarithmized scientific talent pool as dependent variable. Cases with $\text{STP}/\text{AP} > 3$ and cases where $\text{sqrt}(\text{research_intensity}) > 1$ have been excluded

	Model 1			Model 2			Model 3		
	C	SE	Sig.	C	SE	Sig.	C	SE	Sig.
$\ln(\text{academic personnel HC})$	1.091	0.037	0.000	0.736	0.048	0.000	0.796	0.050	0.000
$\text{sqrt}(\text{research intensity})$				2.696	0.996	0.012	3.813	0.630	0.000
Legal status				-0.188	0.176	0.296	-0.150	0.185	0.425
STEM orientation (students)				0.693	0.214	0.003	0.827	0.164	0.000
Health orientation (students)				0.316	0.241	0.202	0.326	0.199	0.113
PhD awarding				0.605	0.300	0.054	0.290	0.216	0.191
University Hospital				0.645	0.167	0.001	0.519	0.118	0.000
_cons	-1.374	0.294	0.000	-0.292	0.349	0.411	-0.638	0.329	0.063
Country-level fixed effects		No			No			Yes	
Rsquare		0.61			0.74			0.79	
N		1,129			764			764	

The model fit increased significantly by adding HEIs' characteristics and—to a lesser extent—by including country dummy variables. We notice that in Model 1, AP is essentially proportional to STP, but in Models 2 and 3 the coefficient is below 1. However, this is accounted for by the fact that research intensity is correlated with HEI size as well.

As expected, with the same number of AP, STP is larger with increasing research intensity and for HEIs that are oriented towards natural and technical sciences. The orientation towards health sciences is not statistically significant, but the university hospital dummy is. This result shows that the impact of medicine on AP is mostly due to the presence of hospitals with their personnel not included in the HEI AP perimeter.

The variable for legal status is not statistically significant, but one must consider that the sample is very unbalanced towards public HEIs. The PhD awarding variable is also not statistically significant, suggesting that research intensity is more relevant than a (legal) research mandate to account for differences in the STP/AP ratio.

Therefore, the results by large confirm our hypotheses.

Because we used a log-log regression, coefficients of variables can be interpreted as elasticities of STP over AP. In such a perspective, a research intensity of 0.25 (i.e., enrolling one PhD student per four undergraduate students) would increase STP by nearly four times as compared with a null research intensity. A research intensity of 0.25 is the level of top-ranked European universities such as Oxford, Cambridge, and ETH Zurich. As an example, the largest non-PhD awarding HEI in our sample (the UAS of Western Switzerland) has 3,800 AP and only 411 STP, but ETH Zurich has 9,000 AP and 7,848 STP. Accordingly, STP really measures the scientific potential of HEIs rather than their aggregate size (as associated also with education and, perhaps, third mission).

Conversely, an HEI enrolling all students in natural and technical sciences would have an increase of STP of 71% compared with an HEI with 19% of students in that area (our sample

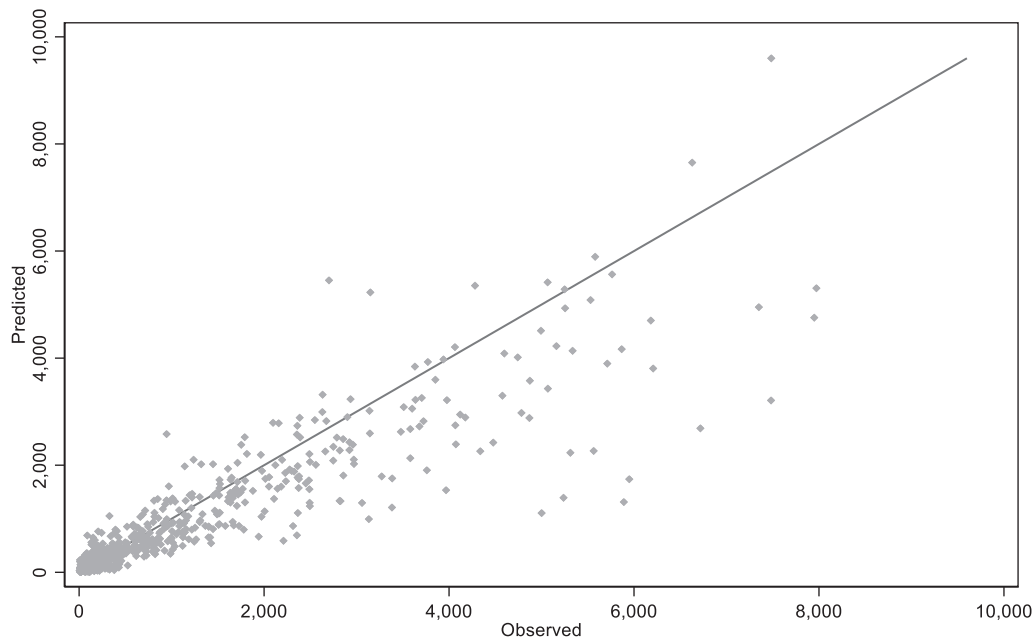


Figure 3. Predicted versus observed STP values (estimations from the full regression model including country dummy variables).

average). The presence of a university hospital entails an increase in STP by 68% with the same number of AP. These institutional characteristics therefore account for large differences in the STP/AP ratios.

Following the regression analyses, we examined the predictive ability of the models in terms of the original variable STP. In other words, we investigated the extent to which the regression models are able to predict correctly the (untransformed) values of STP. The correlation between predicted and observed STP values is 0.78 for Model 1, 0.80 for Model 2, and 0.90 for Model 3. As shown by Figure 3, the fit is very good along the whole range of institutional size. Thus, once we take into account country specificities and influencing factors, such as the presence of a hospital, the STP value of an HEI can be predicted from AP with high precision. The other way around, it would be possible to compute from STP statistically good proxies for AP when further institutional covariates (such as research intensity) are available.

5. DISCUSSION

In the introduction to this paper, we suggested that there are good reasons to complement output indicators, such as those derived from bibliometric databases, with input indicators measuring institutional size, such as budgets or numbers of academic personnel (Moed & Halevi, 2015). From the perspective of institutional leaders and managers, it is relevant to benchmark their HEI with institutional peers of roughly similar size. This would provide a sensible analysis of whether their institution is performing better or worse than HEIs having similar resources (Meek & van der Lee, 2005). From the perspective of economists, institutional efficiency can only be analyzed by linking input (resources) and output (e.g., publication counts; Abramo & D'Angelo, 2016). The resulting high efficiency values mean that institutions completely utilize their resources and produce maximal outputs (Agasisti and Gralka (2019); see also Rhaïem (2017) and Ghimire, Amin, and Wardley (2021)).

We suggest that measuring university size is relevant for institutional bibliometric analyses and for the practice of international rankings as well (Gadd, Holmes, & Shearer, 2021). In the last two decades, rankings have made universities and research-focused institutions measurable and comparable with respect to global competitions for reputation and students (Hammarfelt, De Rijcke, & Wouters, 2017). The issue with the institutional standardization practiced in rankings is the heterogeneity of universities: Universities have different sizes, missions, and institutional contexts (Waltman, Wouters, & van Eck, 2017). This is an attempt by size-independent performance analyses to control for institutional heterogeneity with respect to size (Waltman et al., 2017). Yet, empirical analyses reveal that such indicators are still correlated with size. It seems that rankings largely reflect the wealth of universities rather than their performance (Lepori et al., 2019). To focus on performance (and not wealth), comparable information on resources is required to interpret their results.

Whereas output indicators such as institutional publication counts are available in well-known literature databases (e.g., Web of Science), input indicators (reflecting different institutional sizes) are difficult to obtain (Waltman et al., 2017). Comparable data on AP were usually available for only a few countries, but the situation has improved significantly in recent decades regarding the United States and Europe (Lepori et al., 2022).

In this study, we deal with a new input indicator, the STP indicator from SIR, which delivers worldwide comparable data reflecting personnel resources. We empirically compared the indicator with AP data derived from institutional statistics for a large sample of HEIs in different European countries. This comparison is meant to provide better

understanding of the robustness and methodological limitations of both indicators; the factors affecting their relationships; and how they could be applied in size-independent institutional analyses, such as bibliometric performance analyses, university rankings, and institutional efficiency analyses.

Our results show that STP and AP are correlated with a large effect size. As we also observed an extreme variation of institutional size, the overall high correlation still implies large differences at the HEI level. We were therefore also interested in factors possibly affecting the relationship of STP and AP. In the corresponding regression analyses, we included several factors that we expected to be relevant for the relationship: research intensity, legal status, STEM and health sciences orientation, and university hospitals. Our expectations regarding these factors have been largely confirmed. We showed that they account for a large share of the differences between HEIs in the STP/AP ratio.

Furthermore, we have provided an in-depth analysis of deviant cases at the institutional and country level. We showed that large differences between the two measures at the individual level are related to very specific institutional structures, such as HEIs with large numbers of visiting researchers, graduate schools, or large associated research centers. These cases are few; they can be identified and excluded from the analysis. Country differences are related to institutional structures (such as the pervasive presence of joint units in France), but also to (still remaining) differences in the AP definition adopted by countries.

Our results, first, suggest that measures of institutional size are not only available, but usable for statistical analyses such as the study of efficiency. Methodological issues can be identified and controlled for, for example, by introducing country dummies in statistical models and by excluding outliers. The conclusions are more nuanced for institutional comparisons, as we observed cases where the number of STP and/or AP was significantly affected by specific local conditions. We therefore suggest avoiding the computation of size-normalized indicators, as they might hide sources of bias. Size measures should be provided alongside (size-normalized) output measures, as is already done by some providers of university rankings. Creating size classes to allow users to compare HEIs of similar size might be another option, as already done by the U-Multirank project.

Second, our results demonstrate that STP has some advantages in measuring institutional personnel resources, particularly in nested settings such as the presence of joint units (as, for example, in France). The STP indicator, however, becomes problematic when research units do not contribute to education to the same extent as other units. This problem is especially relevant in efficiency analyses considering both research and education. Another issue in the empirical use of the STP indicator (e.g., for efficiency analyses) concerns university hospitals. Hospitals are responsible for large differences between AP and STP. Based on our empirical results, we conclude therefore that controlling for university hospitals is key in institutional performance analyses (producing results that should be size-independent).

Third, our analysis provides guidance for methodological improvements of both indicators. As for STP, the analysis suggests advancing the identification of authors affiliated to an institution: Specifically, to deal with multiple institutional affiliations and visiting scholars, authors with a low share (or number) of publications affiliated to an institution could be excluded from STP.

As for AP, the analysis confirms known issues in counting academic personnel and, particularly, the need for a more comparable treatment of employed PhD candidates. These issues are currently under debate at OECD and EUROSTAT in view of a revision of the

international methodological guidelines. For both measures, AP and STP, our results emphasize the need for a more systematic delineation of associated units and hospitals, as already done in Europe from the public-sector register OrgReg (<https://www.risis2.eu/orgreg-data>) (Lepori, 2020).

Fourth, we suggest that the availability of both STP and AP variables is the optimal situation for future institutional performance analyses. The comparison of the empirical results from both variables might be highly informative. The comparison will not only contribute to an understanding of what is measured exactly (with respect to personnel resources) but also point to problematic institutional cases. These cases can be detected and possibly excluded from the empirical analysis. Moreover, our results suggest exploring the construction of composite size indicators combining AP and STP, as this may limit the impact of institutional characteristics and allow developing more comparable size measures.

The results of our study, therefore, emphasize not only the relationship but also the complementarity of the AP and STP.

AUTHOR CONTRIBUTIONS

Benedetto Lepori: Conceptualization, Data curation, Methodology, Writing—original draft. Lutz Bornmann: Conceptualization, Writing—review & editing. Félix de Moya Anegón: Data curation.

COMPETING INTERESTS

The authors have no competing interests.

FUNDING INFORMATION

The authors acknowledge support from the European Commission Horizon2020 RISIS2 project (Grant agreement ID: 82409).

DATA AVAILABILITY

European Tertiary Education Register data are available from <https://www.eter-project.com>. Scimago Institutions Ranking data are available from the authors.

REFERENCES

- Abramo, G., & D'Angelo, C. A. (2016). A farewell to the MNCS and like size-independent indicators. *Journal of Informetrics*, *10*(2), 646–651. <https://doi.org/10.1016/j.joi.2016.04.006>
- Agasisti, T., & Gralka, S. (2019). The transient and persistent efficiency of Italian and German universities: A stochastic frontier analysis. *Applied Economics*, *51*(46), 5012–5030. <https://doi.org/10.1080/00036846.2019.1606409>
- Aksnes, D. W., & Sivertsen, G. (2019). A criteria-based assessment of the coverage of Scopus and Web of Science. *Journal of Data and Information Science*, *4*(1), 1–21. <https://doi.org/10.2478/jdis-2019-0001>
- Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*, *1*(1), 377–386. https://doi.org/10.1162/qss_a_00019
- Baumann-Pauly, D., Wickert, C., Spence, L. J., & Scherer, A. G. (2013). Organizing corporate social responsibility in small and large firms: Size matters. *Journal of Business Ethics*, *115*, 693–705. <https://doi.org/10.1007/s10551-013-1827-7>
- Bentley, P. J., & Kyvik, S. (2013). Individual differences in faculty research time allocations across 13 countries. *Research in Higher Education*, *54*(3), 329–348. <https://doi.org/10.1007/s11162-012-9273-4>
- Bornmann, L., Gralka, S., de Moya Anegón, F., & Wohlrabe, K. (2020). Efficiency of universities and research-focused institutions worldwide: An empirical DEA investigation based on institutional publication numbers and estimated academic staff numbers. *CESifo Working Paper*, 8157. <https://doi.org/10.2139/ssrn.3555584>
- Calero-Medina, C., Noyons, E., Visser, M., & De Bruin, R. (2020). Delineating organizations at CWTS—A story of many pathways. In C. Daraio & W. Glänzel (Eds.), *Evaluative informetrics: The art of metrics-based research assessment* (pp. 163–177). Cham: Springer. https://doi.org/10.1007/978-3-030-47665-6_7

- Cornforth, C., & Simpson, C. (2002). Change and continuity in the governance of nonprofit organizations in the United Kingdom: The impact of organizational size. *Nonprofit Management and Leadership*, 12(4), 451–470. <https://doi.org/10.1002/nml.12408>
- Damanpour, F. (1992). Organizational size and innovation. *Organization Studies*, 13(3), 375–402. <https://doi.org/10.1177/017084069201300304>
- Daraio, C., Bonaccorsi, A., & Simar, L. (2015). Efficiency and economies of scale and specialization in European universities: A directional distance approach. *Journal of Informetrics*, 9(3), 430–448. <https://doi.org/10.1016/j.joi.2015.03.002>
- Elizondo, A. R., Calero-Medina, C., & Visser, M. S. (2022). The three-step workflow: A pragmatic approach to allocating academic hospitals' affiliations for bibliometric purposes. *Journal of Data and Information Science*, 7(1), 20–36. <https://doi.org/10.2478/jdis-2022-0006>
- Gadd, E., Holmes, R., & Shearer, J. (2021). Developing a method for evaluating global university rankings. *Scholarly Assessment Reports*, 3(1), 2. <https://doi.org/10.29024/sar.31>
- Ghimire, S., Amin, S. H., & Wardley, L. J. (2021). Developing new data envelopment analysis models to evaluate the efficiency in Ontario Universities. *Journal of Informetrics*, 15(3), 101172. <https://doi.org/10.1016/j.joi.2021.101172>
- Glänzel, W., Thijs, B., & Debackere, K. (2016). Productivity, performance, efficiency, impact—What do we measure anyway? Some comments on the paper “A farewell to the MNCS and like size-independent indicators” by Abramo and D’Angelo. *Journal of Informetrics*, 10(2), 658–660. <https://doi.org/10.1016/j.joi.2016.04.008>
- Gralka, S., Wohlrabe, K., & Bornmann, L. (2019). How to measure research efficiency in higher education? Research grants vs. publication output. *Journal of Higher Education Policy and Management*, 41(3), 322–341. <https://doi.org/10.1080/1360080X.2019.1588492>
- Hammarfelt, B., De Rijcke, S., & Wouters, P. (2017). From eminent men to excellent universities: University rankings as calculative devices. *Minerva*, 55(4), 391–411. <https://doi.org/10.1007/s11024-017-9329-x>, PubMed: 29200503
- Handy, C. (2007). *Understanding organizations*. London: Penguin.
- Huisman, J., Lepori, B., Seeber, M., Frölich, N., & Scordato, L. (2015). Measuring institutional diversity across higher education systems. *Research Evaluation*, 24(4), 369–379. <https://doi.org/10.1093/reseval/rvv021>
- Jacso, P. (2005). As we may search—Comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced databases. *Current Science*, 89(9), 1537–1547. <https://www.jstor.org/stable/24110924>
- Jaquette, O., & Parra, E. E. (2014). Using IPEDS for panel analyses: Core concepts, data challenges, and empirical applications. In M. B. Paulsen (Ed.), *Higher education: Handbook of theory and research* (pp. 467–533). Dordrecht: Springer. https://doi.org/10.1007/978-94-017-8005-6_11
- Kimberly, J. R. (1976). Organizational size and the structuralist perspective: A review, critique, and proposal. *Administrative Science Quarterly*, 21(4), 571–597. <https://doi.org/10.2307/2391717>
- Lepori, B. (2020). A register of public-sector research organizations as a tool for research policy studies and evaluation. *Research Evaluation*, 29(4), 355–365. <https://doi.org/10.1093/reseval/rvaa012>
- Lepori, B. (2022). The heterogeneity of European Higher Education Institutions: A configurational approach. *Studies in Higher Education*, 47(9), 1827–1843. <https://doi.org/10.1080/03075079.2021.1968368>
- Lepori, B., Bonaccorsi, A., Daraio, A., Daraio, C., Gunnes, H., ... Wagner-Schuster, D. (2015). *Establishing a European tertiary education register*. Final Report. Brussels: European Commission. <https://doi.org/10.2766/755061>
- Lepori, B., Borden, V. M., & Coates, H. (2022). Opportunities and challenges for international institutional data comparisons. *European Journal of Higher Education*, 12, 373–390. <https://doi.org/10.1080/21568235.2022.2094817>
- Lepori, B., Geuna, A., & Mira, A. (2019). Scientific output scales with resources. A comparison of US and European universities. *PLOS ONE*, 14(10), e0223415. <https://doi.org/10.1371/journal.pone.0223415>, PubMed: 31613903
- Meek, V. L., & van der Lee, J. J. (2005). Performance indicators for assessing and benchmarking research capacities in universities. In *APEID, UNESCO Bangkok Occasional Paper Series Paper no. 2: United Nations Educational, Scientific and Cultural Organisation (UNESCO)*.
- Moed, H. F., & Halevi, G. (2015). Multidimensional assessment of scholarly research impact. *Journal of the Association for Information Science and Technology*, 66(10), 1988–2002. <https://doi.org/10.1002/asi.23314>
- Nomaler, Ö., Frenken, K., & Heimeriks, G. (2014). On scaling of scientific knowledge production in U.S. metropolitan areas. *PLOS ONE*, 9(10), e110805. <https://doi.org/10.1371/journal.pone.0110805>, PubMed: 25353686
- OECD. (2000). *Measuring R&D in the higher education sector. Methods in use in the OECD/EU member countries*. Paris: OECD.
- OECD. (2015). *Frascati manual 2015. Guidelines for collecting and reporting data on research and experimental development*. Paris: OECD. <https://doi.org/10.1787/9789264239012-en>
- Petr, M., Engels, T. C., Kulczycki, E., Dušková, M., Guns, R., ... Sivertsen, G. (2021). Journal article publishing in the social sciences and humanities: A comparison of Web of Science coverage for five European countries. *PLOS ONE*, 16(4), e0249879. <https://doi.org/10.1371/journal.pone.0249879>, PubMed: 33831115
- Purnell, P. J. (2022). The prevalence and impact of university affiliation discrepancies between four bibliographic databases—Scopus, Web of Science, Dimensions, and Microsoft Academic. *Quantitative Science Studies*, 3(1), 99–121. https://doi.org/10.1162/qss_a_00175
- Rhaiem, M. (2017). Measurement and determinants of academic research efficiency: A systematic review of the evidence. *Scientometrics*, 110(2), 581–615. <https://doi.org/10.1007/s11192-016-2173-1>
- Sivertsen, G. (2016). Data integration in Scandinavia. *Scientometrics*, 106, 849–855. <https://doi.org/10.1007/s11192-015-1817-x>
- Sivertsen, G., Rousseau, R., & Zhang, L. (2019). Measuring scientific contributions with modified fractional counting. *Journal of Informetrics*, 13(2), 679–694. <https://doi.org/10.1016/j.joi.2019.03.010>
- Tekles, A., & Bornmann, L. (2020). Author name disambiguation of bibliometric data: A comparison of several unsupervised approaches. *Quantitative Science Studies*, 1(4), 1510–1528. https://doi.org/10.1162/qss_a_00081
- Thijs, B., & Glänzel, W. (2008). A structural analysis of publication profiles for the classification of European research institutes. *Scientometrics*, 74(2), 223–236. <https://doi.org/10.1007/s11192-008-0214-0>
- UOE. (2013). *UOE data collection on education systems. Volume 1. Manual. Concepts, definitions, classifications*. Montreal, Paris, Luxembourg: UNESCO, OECD, Eurostat.

- van Raan, A. F. J. (2004). Measuring science. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of quantitative science and technology research* (pp. 19–50). Dordrecht: Springer. https://doi.org/10.1007/1-4020-2755-9_2
- van Raan, A. F. J. (2013). Universities scale like cities. *PLOS ONE*, 8(3), e59384. <https://doi.org/10.1371/journal.pone.0059384>, PubMed: 23544062
- Vieira, E. S., & Lepori, B. (2016). The growth process of higher education institutions and public policies. *Journal of Informetrics*, 10(1), 286–298. <https://doi.org/10.1016/j.joi.2016.01.001>
- Waltman, L., & van Eck, N. J. (2015). Field-normalized citation impact indicators and the choice of an appropriate counting method. *Journal of Informetrics*, 9(4), 872–894. <https://doi.org/10.1016/j.joi.2015.08.001>
- Waltman, L., van Eck, N. J., Visser, M., & Wouters, P. (2016). The elephant in the room: The problem of quantifying productivity in evaluative scientometrics. *Journal of Informetrics*, 10(2), 671–674. <https://doi.org/10.1016/j.joi.2015.12.008>
- Waltman, L., Wouters, P., & van Eck, N. J. (2017). Ten principles for the responsible use of university rankings. <https://www.universiteitleiden.nl/binaries/content/assets/algemeen/onderzoek/responsible-use-of-university-rankings.pdf>
- Zitt, M. (2016). Paving the way or pushing at open doors? A comment on Abramo and D’Angelo “Farewell to size-independent indicators.” *Journal of Informetrics*, 10(2), 675–678. <https://doi.org/10.1016/j.joi.2016.04.003>