an open access journal

Citation: Abramo, G., Aksnes, D. W., & D'Angelo, C. A. (2022). Unveiling the distinctive traits of a nation's research performance: The case of Italy and Norway. *Quantitative Science Studies*, *3*(3), 732–754. https://doi.org/10.1162/qss_a_00198

DOI:
https://doi.org/10.1162/qss_a_00198

Peer Review:
https://publons.com/publon/10.1162/qss_a_00198

Received: 24 March 2022
Accepted: 9 May 2022

Corresponding Author:
Giovanni Abramo
giovanni.abramo@uniroma2.it

Handling Editor:
Ludo Waltman

The MIT Press

RESEARCH ARTICLE

Unveiling the distinctive traits of a nation's research performance: The case of Italy and Norway

Giovanni Abramo[1], Dag W. Aksnes[2], and Ciriaco Andrea D'Angelo[1,3]

[1]Laboratory for Studies in Research Evaluation, Institute for System Analysis and Computer Science (IASI-CNR). National Research Council, Rome, Italy
[2]Nordic Institute for Studies in Innovation, Research and Education, Oslo, Norway
[3]University of Rome "Tor Vergata," Dept of Engineering and Management, Rome, Italy

Keywords: bibliometrics, Fractional Scientific Strength, *FSS*, national R&D systems, professors, research evaluation, universities

ABSTRACT

In this study, we analyze the research performance of Italian and Norwegian professors using constituent components of the Fractional Scientific Strength (*FSS*) indicator. The main focus is on differences across fields in publication output and citation impact. The overall performance (*FSS*) of the two countries, which differ considerably in research size and profile, is remarkedly similar. However, an in-depth analysis shows that there are large underlying performance differences. An average Italian professor publishes more papers than a Norwegian, while the citation impact of the research output is higher for the Norwegians. In addition, at field level, the pattern varies along both dimensions, and we analyze in which fields each country has its relative strengths. Overall, this study contributes to further insights into how the research performance of different countries may be analyzed and compared to inform research policy.

1. INTRODUCTION

One of the principal objectives that governments and institutions pursue through research assessment is improving research effectiveness and efficiency. Effectiveness can be deployed along various dimensions, depending on the strategic goals of the country. These include among others, sustainability goals, research integrity, responsible research and innovation, open science, gender equality in research, geographical balance, international research collaboration, public–private collaboration, and rapid and widespread diffusion of new knowledge to potential users.

Efficiency essentially deals with the relation between research output and inputs to produce it. All inputs being equal, individual and organizational performance can be increased by producing more research products (alone or in collaboration), or of better quality (higher impact), or both. Providing the government, research institutions, and scientists with information along each single dimension of their overall research performance might serve better the aim of stimulating improvement. In this way, the single actors are informed about the dimensions with higher margins of efficiency gains.

Whether associated or not with financial rewards or other competitive mechanisms, the simple communication to individuals of their performance scores and ranks, if properly channeled, can be instrumental in continuous improvement. At the large-scale level of national scientific systems, governments and policymakers might also benefit from such information to complement that of their comparative world performance rank. It can inform decisions on the dimensions along which to focus and orient policy measures.

This study deals with research efficiency. The quintessential indicator of efficiency in any production systems is productivity (i.e., the rate of output per unit of input). In other words, it measures how efficiently production inputs are being used. Most bibliometricians define productivity as the number of publications in the period of observation. Because publications have different values (impact), and the resources employed for research are not homogeneous across individuals and organizations, we adopt the definition of productivity extracted from the economic theory of production: the value of output per euro spent in research (i.e., costs of labor and all other production factors, referred to as capital in the following).

Following on from our previous studies on the comparison of research productivity of Italian and Norwegian academics (Abramo, Aksnes, & D'Angelo, 2020, 2021), in this work, we present a methodology to measure the single components of research productivity, and apply it to further scan and contrast the two countries' academic systems.

In terms of their research production systems, Italy and Norway are rather different. First, the total scientific output of Italy measured by number of journal articles is approximately five times as large as Norway's. In 2018 (latest available data at time of first of our previous studies), Italy ranked as the eighth largest science producing country in the world and Norway was in 29th position (Norges forskningsråd, 2019). However, compared with the population size, per capita production is 150% higher for Norway than for Italy. Furthermore, the overall citation impact of research output per capita is also higher for Norway. Thus, Norway may be considered a more research-intensive country than Italy. The underlying reasons for these differences are interesting to analyze further, and we aim to contribute to additional knowledge on the issue.

Abramo et al. (2020) for the first time ever applied the Fractional Scientific Strength, or *FSS*, indicator of research performance of professors and universities to a country other than Italy (i.e., Norway). *FSS* (Abramo & D'Angelo, 2014) differs from the most popular indicators, such as the *h*-index (Hirsch, 2005), the MNCS (Waltman, van Eck et al., 2011), and their variants (Alonso, Cabrerizo et al., 2009; Waltman, 2016; Wildgaard, Schneider, & Larsen, 2014), essentially because it accounts for the resources used to produce research output. Furthermore, the *FSS* adopts the publications' fractional counting method, and values them through citation indicators.

We refer the reader to Abramo et al. (2020) for details about the academic systems in the two countries and the difficulties of achieving comparable measurements of performance, and the ways to overcome them. The reader will also find the procedure for operationalizing the measurements and all the limits and assumptions involved. Where appropriate, we extrapolate from that study (and report in this) the performance scores and ranks in the two countries, across disciplines.

In summary, the previous study showed the following:

- hardly any differences in the average research productivity of Italian and Norwegian professors;
- a higher concentration of Norwegian professors in the top and the bottom tails of the productivity distribution; and

- higher productivity of Norway in Mathematics and Earth and space sciences, and of Italy in Biomedical research and Engineering.

In this study, we complement the above results with more in-depth analysis of performance along the single dimensions of overall research productivity of Italian and Norwegian professors. In particular, we contrast the academics of the two countries in terms of yearly output, fractional output, average citations per paper, and average impact factor (IF) per paper.

The elaborations are aimed at answering the following research questions:

1. Do the academics of one country tend to publish more than those of the other?
2. In doing so, do they engage in more collaborative work?
3. Do they pursue higher quality research products in terms of citation impact?
4. Do they publish their results in more prestigious journals?

Such questions are important in the context of evaluative use of bibliometrics. For example, in an upcoming evaluation of Norwegian STEM research[1], the panels are asked to assess the production of scholarly publications, the research quality, and national and international cooperation—in addition to several other dimensions, such as research career and mobility, leadership, and research infrastructure. The Italian national research assessment exercise series, VQR[2], of which the results of the latest edition have just been released, emphasizes the importance of the quality of scientific production over volume, international research collaboration over national, etc. Thus, in-depth analyses of bibliometric patterns may be useful to assess the aspects related to research output, where a unit of analysis may perform differently across the dimensions. For example, high publication volume does not necessarily imply high citation impact and vice versa. Moreover, different indicators reflect different dimensions of performance, which are required to provide reliable assessments, a point which has long been emphasized by bibliometric professionals (van Raan, 1993). Therefore, we wish to underline that none of the indicators that we present and measure in this paper can be considered alone as performance indicators. Each one represents more or less a dimension of performance, but not the overall performance. They convey complementary information, useful to orient scientists in focusing their efforts for continuous improvement, and managers and for policymakers in formulating their interventions to the same aim.

Generally, a large number of factors may explain patterns identified through bibliometric analyses. Some can be investigated bibliometrically, but often other approaches are needed. For example, some might be related to different policy and behavioral mechanisms. Such explanatory attempts are not straightforward, as the systems are influenced by a large number of factors, and at best indirect evidence can be provided. For example, higher publication volume or quality may be due to more financial resources, more time for research, better management, changes in the staff composition, recruitment of talented individuals, responses to funding incentives, and so forth.

In this paper, we therefore mainly focus on the empirical aspects, and second-order interpretations are addressed to a limited extent. In presenting our study, we devote special attention to the definition of each single indicator and the description of the operationalization of

[1] See https://www.forskningsradet.no/en/statistics-evaluations/natural-sciences-2022-2023/ (accessed on May 9, 2022).

[2] https://www.anvur.it/wp-content/uploads/2020/09/Bando-VQR-2015-19_25-settembre_2020_signed.pdf (accessed May 9, 2022).

their measurement to make hopeful future replications of the exercise in other countries more straightforward.

The rest of the paper is structured as follows. In the next section we present the construction of the data set. In Section 3, we will recall the *FSS* definition and present those of the complementary indicators we are going to measure. In Section 4, we will present the results of the assessment. Section 5 will conclude the work with our considerations.

2. DATA

To follow on from the abovementioned previous studies of ours and their findings (Abramo et al. 2020, 2021), we observe the research activity of Italian and Norwegian professors in the same period, 2011–2015.

Data on Italian professors have been retrieved from the database on university personnel, which is operated by the Italian Ministry of Universities and Research (MIUR). This database contains information on the name of each individual, their gender, affiliation, field classification, and academic rank at the end of each year[3].

Data on the Norwegian professors have been retrieved from a similar database, the Norwegian Research Personnel Register, operated by Nordic Institute for Studies in Innovation, Research and Education (NIFU) (the database underlying the official Norwegian R&D statistics).

For reasons of significance, the analysis is limited to those professors who held formal faculty positions for at least three years over the 2011–2015 period. Furthermore, the data set is limited to individuals with at least one publication during the time period (nonpublishing personnel are not registered in Norwegian databases).

The data on the publication output of the Italian professors are retrieved from the Italian Observatory of Public Research (ORP), a bibliometric database where data are available at the level of authors. The database is developed and maintained by Abramo and D'Angelo, and derived under license from the Clarivate Analytics Web of Science (WoS) Core Collection.

For the Norwegian professors, WoS publication data have been retrieved from a bibliographic database called *Cristin* (Current Research Information System in Norway). This is a national information infrastructure, providing data for all institutions in the higher education sector, research institutes and hospitals.

To perform a comparative assessment without distortions (see below), each individual is classified in one and only one WoS subject category (SC)[4].

The analysis covers professors in all fields, with the exception of arts and humanities and a few SCs in the social sciences, where WoS has largest limitations in terms of coverage (Aksnes & Sivertsen, 2019; Hicks, 1999; Larivière, Archambault et al., 2006). There are large variations in the size of the different SCs. In some cases there are few observations, which might cause random fluctuations in performance. Therefore, we have further removed the SCs that did not meet a minimum threshold of at least 10 professors in total, of both nationalities.

[3] https://cercauniversita.cineca.it/php5/docenti/cerca.php (accessed May 9, 2022).

[4] We assigned to each publication the SC or SCs of the hosting journal. We then classified each professor in the most recurrent SC in their publication portfolio. We refer the reader to Abramo et al. (2020) for more details of the procedure followed to classify professors.

Table 1. Data set of analysis

Discipline	No. of SCs	Italy Total professors	Italy Assistant (%)	Italy Associate (%)	Italy Full (%)	Norway Total professors	Norway Assistant (%)	Norway Associate (%)	Norway Full (%)
Mathematics	6	2,122	22.8	40.4	36.8	183	2.2	29.0	68.9
Physics	16	2,918	23.8	43.3	32.9	256	9.8	21.5	68.8
Chemistry	7	1,896	28.1	43.8	28.2	122	14.8	29.5	55.7
Earth and space sciences	11	1,873	29.6	41.6	28.8	413	14.8	29.1	56.2
Biology	28	5,635	34.9	37.6	27.5	736	20.7	28.0	51.4
Biomedical research	14	3,707	37.8	36.3	25.9	245	19.2	30.2	50.6
Clinical medicine	36	7,571	35.8	36.0	28.2	958	10.5	31.3	58.1
Psychology	6	475	31.2	39.8	29.1	148	2.7	42.6	54.7
Engineering	34	5,522	26.5	40.3	33.2	426	5.2	27.2	67.6
Political and social sciences	11	442	20.6	41.6	37.8	438	6.2	33.1	60.7
Economics	8	1,848	19.0	40.4	40.6	402	3.0	33.1	63.9
Total	*177*	*34,009*	*30.6*	*39.0*	*30.4*	*4327*	*10.9*	*30.1*	*59.0*

The final data set consists of 34,009 Italian and 4,327 Norwegian professors, falling in 177 SCs. Their distribution per academic rank and discipline[5] is shown in Table 1.

3. THE *FSS* INDICATOR AND ITS COMPONENTS

As already said, in the economic theory of production, productivity is defined as the rate of output per unit of input. Because publications (output) have different values (impact), and the resources employed for research are not homogeneous across individuals and organizations, in research systems, a more appropriate definition of productivity is the value of output per euro spent in research. The *FSS* indicator is a proxy measure of research productivity. A thorough description of the *FSS* indicator, and the theory underlying it, can be found in Abramo and D'Angelo (2014).

To measure the yearly average research productivity of Italian and Norwegian academics, Abramo et al. (2020) used the following formula[6]:

$$FSS = \frac{1}{\left(\frac{\mathrm{w}_r}{2}+k\right)} \cdot \frac{1}{t}\sum_{i=1}^{N}\frac{c_i}{\bar{c}} f_i \qquad (1)$$

[5] The SCs are classified and grouped into disciplines according to a system previously published on the webpage of the ISI Journal Citation Reports. This page is no longer available at the current Clarivate site. It should be noted that all SCs are assigned to only one discipline.

[6] The underlying assumption is that labor and capital contribute equally to production.

where

w_r = average yearly salary of the professor[7]
k = average yearly capital available for research to the professor[8]
t = number of years of work by the professor in period under observation
N = number of publications by the professor in period under observation
c_i = citations received by publication i, until 31 October 2018
$\bar{c}$ = average of distribution of citations received for all WoS-cited publications in same year and SC of publication i[9]
f_i = fractional contribution of professor to publication i.

In the fields where the common practice is to place authors in simple alphabetical order, the fractional contribution equals the inverse number of authors, while in other cases different weights are applied (see Waltman, 2012). More specifically, for Biology, Biomedical research, and Clinical medicine, the usual practice is to indicate the contributions of each author by their position in the bylines. Thus, for these disciplines, we have applied a system where different weights are credited to each coauthor depending on their position in the list of authors and the type of coauthorship (intramural or extramural)[10].

Looking at the formula, it can be seen that productivity is a function of output, or more precisely of individual contribution to output, its quality (impact), and the resources used for production.

In this work we want to assess how Italian professors compare to Norwegian professors along each dimension of productivity, namely, in terms of output, fractional output, average citations per paper, and, although not a dimension of productivity, average IF per paper. The average IF per paper is informative about the degree of prestige of the journals chosen by researchers to have their manuscripts published. The breakdown of the productivity indicator along the above dimensions allows us to unveil the possible different traits of professors in conducting research activities.

The relevant indicators are the following:

- Output (O), average yearly publications authored by the professor per euro spent in research:

$$O = \frac{1}{\left(\frac{w_r}{2} + k\right)} \frac{N}{t} \tag{2}$$

[7] We halved labor costs, because we assumed that 50% of professors' time is allocated to activities other than research. In Section 4.3, we will propose a specific analysis to test how sensitive the results are to such an assumption.
[8] Sources of input data and assumptions adopted in the measurement can be found in Abramo et al. (2020).
[9] Abramo, Cicero, and D'Angelo (2012) demonstrated that the average of the distribution of citations received for all cited publications of the same year and subject category is the most effective scaling factor.
[10] This means that if the publication is the result of a collaboration that is exclusively intramural (a single affiliation in the address list), 40% is credited to both the first and last authors, while the remaining 20% is shared among all other authors. One the other hand, if a publication involves extramural collaboration, 30% is credited to both the first and last authors; 15% to both the second and last but one authors; and the remaining 10% is divided among all other authors. This weighting system has been designed according to advice from senior Italian professors in the life sciences, and widely accepted by the customers of our research assessment services. The crediting values of the system could be changed according to different practices in other national contexts.

where

w = average yearly salary of the professor
k = average yearly capital available for research to the professor
N = number of publications by the professor in period under observation
t = number of years on staff of the professor during the period under observation

- Fractional Output (*FO*), average yearly total contribution to publications authored by the professor per euro spent in research:

$$FO = \frac{1}{\left(\frac{w_r}{2} + k\right)} \frac{1}{t} \sum_{i=1}^{N} f_i \tag{3}$$

where

f_i = fractional contribution of professor to publication i.

- Average Citation (*AC*), average standardized citations per publication:

$$AC = \frac{1}{N} \sum_{i=1}^{N} \frac{c_i}{\bar{c}} \tag{4}$$

where

c_i = citations received by publication i
$\bar{c}$ = average of distribution of citations received for all WoS-cited publications in same year and SC of publication i.

- Average IF (*AIF*), average standardized IF per publication:

$$AIF = \frac{1}{N} \sum_{i=1}^{N} \frac{IF_i}{\overline{IF}} \tag{5}$$

where

IF = IF of journal hosting publication i
$\overline{IF}$ = average of distribution of IFs of journals hosting all WoS-cited publications in same year and SC of publication i.

The performance scores of professors belonging to different SCs cannot be compared directly. In fact, scientists' intensity of publication remarkably varies across fields, in general (Lillquist & Green, 2010; Sandström & Sandström, 2009; Sorzano, Vargas et al., 2014), and in both countries in particular (D'Angelo & Abramo, 2015; Piro, Aksnes & Rørstad, 2013); citation behavior varies across fields (Stringer, Sales-Pardo, & Amaral, 2010; Vieira & Gomes, 2010); and the intensity of collaboration (i.e., the average number of coauthors per publication) also varies across fields (Abramo, D'Angelo, & Murgia, 2013; Glanzel & Schubert, 2004; Yoshikane & Kageura, 2004).

To avoid distortions, then, the performance rankings of professors are constructed at the SC level.

For comparisons at higher levels of aggregation (i.e., discipline and overall), we normalize performance scores to the average score of all professors of the same SC. To exemplify, an *FSS* score of 1.10 means that the professor's performance is 10% above average, in his or her own SC[11].

[11] Note that the scaling is not referred to world distributions. As we are comparing Italy vs. Norway, the "average" used to rescale original distributions is calculated by collapsing Italian and Norwegian performance distributions only.

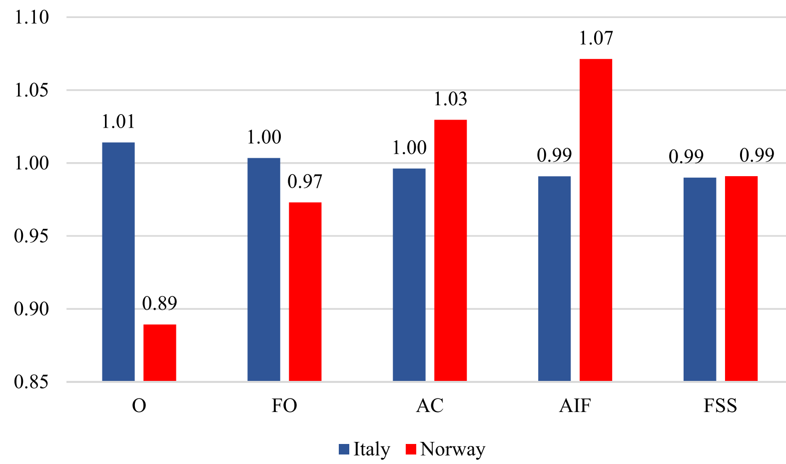

Figure 1. 2011–2015 average normalized research output and fractional output per euro spent (*O*, *FO*), impact (*AC*, *AIF*), and productivity (*FSS*), at overall country level.

In the following tables, figures and text, all performance scores are normalized, but we keep the same denomination for the relevant indicators.

4. RESULTS

In the following, we present the comparison of Italian and Norwegian professors along each of the above indicators, per discipline and overall.

Figure 1 shows the average normalized scores of each indicator, for the overall 34,009 Italian and 4,327 Norwegian professors. While the two populations show practically the same productivity, *FSS*[12], noticeable differences occur for the other indicators. On average, Italian professors publish more (1.4% above average) than Norwegian professors (11.1% below average). Accounting for the real contribution to each publication, the gap decreases, but it is still there: Italy's *FO* equals 1.00, while Norway's is 0.97.

Because the *FSS* is the same, it follows that Norway's *AC* needs to be higher (1.03) than Italy's (1.00). Furthermore, Norwegian academics on average publish in more prestigious journals (*AIF* = 1.07) than Italians (*AIF* = 0.99).

In short, while the two countries present the same research productivity, Italians publish more, with larger research teams, while Norwegians publish higher quality products in more prestigious journals.

We wondered whether the same traits could be found among top-performing scientists as well. Limiting the analysis to top 10% professors by *FSS* at SC level, we observe similar differences for all considered indicators (Figure 2). Norwegian top professors (*FSS* = 4.15) are slightly more productive than Italians (*FSS* = 4.10). The first build up their supremacy by producing higher impact publications. In fact, the average *O* by Italian top professors is 6.8% higher than Norwegian (2.57 vs. 2.39), and although it diminishes by *FO* to 5.2% (2.77 vs.

[12] Note that for *FSS*, both countries show average figures below one, because in rescaling the *FSS* of individuals, we exclude nil values. This implies that the average of the overall distribution is below one, in general and for both countries. As for impact indicators (*AC*, *AIF*), the effect of removing nil values does not produce the same "visual" effect, while for output indicators (*O*, *FO*), we consider only professors with at least one publication authored in the period under observation; consequently there are no nil values.

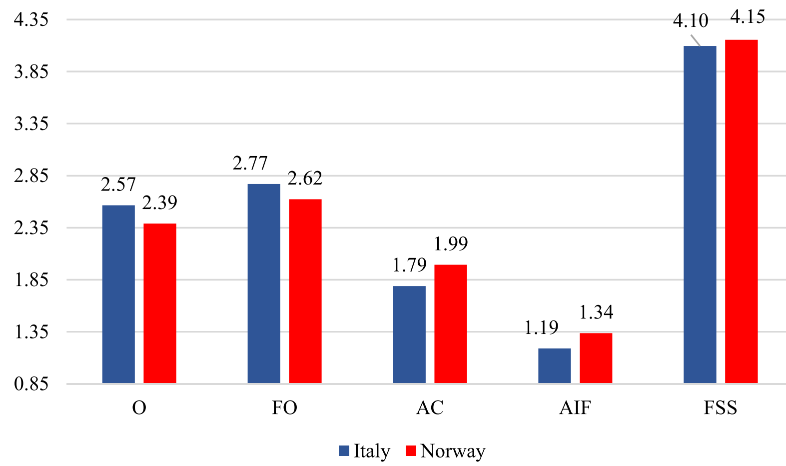

Figure 2. 2011–2015 average normalized research output and fractional output per euro spent (*O*, *FO*), impact (*AC*, *AIF*), and productivity (*FSS*), at overall country level for top 10% productive professors.

2.62), the roles swap in terms of impact, whereby Norwegian top professors register an *AC* (*AJI*) 11.4% (12.4%) higher than Italians.

In the following, we present the results at field level (discipline and SC), to assess possible differences across fields between the two populations. We will start with output indicators; impact will follow.

4.1. Output and Fractional Output per Euro Spent, at Field Level

Italian professors' *O* is higher than that of Norwegians in all 11 disciplines except Mathematics, as shown in Table 2. The biggest gap occurs in Biomedical research, where the 3,707 Italian professors' *O* is 1.02, while the 245 Norwegian professors' is 0.77. Noticeable differences occur also in Psychology (1.05 vs. 0.84), Biology (1.02 vs. 0.82), Engineering (1.01 vs. 0.82), and Clinical medicine (1.02 vs. 0.85).

When accounting for the real contribution to each paper, Italian professors' contribution (*FO*) is greater in only five disciplines, especially in Biomedical research (1.01 vs. 0.88) and in Biology (1.01 vs. 0.90). Norwegians noticeably prevail in Mathematics (1.12 vs. 0.99), Chemistry (1.13 vs. 0.99), and Earth and space sciences (1.08 vs. 0.98).

Next, we analyze the research output at the level of SCs. As noted in Section 2, SCs differ considerably in size, and some of them include a rather small number of individuals, particularly for Norway. Nevertheless, an analysis at this level may reveal interesting differences across the nations.

Table 3 shows how the differences by *O* and *FO* observed at discipline level vary across the SCs of each single discipline. The *FSS* scores are reported as a reference. We note that Norwegian professors outperform Italians' *FSS* in 40% of SCs, in 24% by *O*, and in 38% by *FO*. In particular, by *O* Norwegians never outperform Italians in Psychology, and they do so in only two of the 28 SCs of Biology and in two of the 14 SCs of Biomedical research. On the other hand, in Mathematics, Italians outperform Norwegians in two out of 6 SCs. In terms of *FO*, Norway recovers and, differently from other disciplines, in Chemistry it outperforms Italy in five out of seven SCs.

Table 2. 2011–2015 normalized average research output and fractional output per euro spent (*O*, *FO*), and productivity (*FSS*) at discipline level.

	Italy				Norway			
Discipline	No. of professors	*O*	*FO*	*FSS*	No. of professors	*O*	*FO*	*FSS*
Mathematics	2,122	0.99	0.99	0.94	183	1.10	1.12	1.25
Physics	2,918	1.01	1.00	1.00	256	0.94	1.06	0.97
Chemistry	1,896	1.01	0.99	1.00	122	0.91	1.13	0.98
Earth and space sciences	1,873	1.00	0.98	0.94	413	0.98	1.08	1.27
Biology	5,635	1.02	1.01	1.01	736	0.82	0.90	0.90
Biomedical research	3,707	1.02	1.01	1.01	245	0.77	0.88	0.89
Clinical medicine	7,571	1.02	1.01	0.99	958	0.85	0.93	1.03
Psychology	475	1.05	1.02	0.98	148	0.84	0.94	1.03
Engineering	5,522	1.01	1.01	1.00	426	0.82	0.94	0.88
Political and social sciences	442	1.03	0.99	0.91	438	0.97	1.01	0.89
Economics	1,848	1.01	1.00	0.98	402	0.95	1.00	0.96
Overall	34,009	1.01	1.00	0.99	4327	0.89	0.97	0.99

Table 3. Number and proportion of fields (SCs) per discipline where 2011–2015 normalized average research output and fractional output per euro spent (*O*, *FO*) and productivity (*FSS*) of Norway is higher than Italy.

		No. of SCs where Norway outperforms Italy		
Discipline	No. of SCs	*O*	*FO*	*FSS*
Mathematics	6	4 (67%)	4 (67%)	3 (50%)
Physics	16	7 (44%)	10 (63%)	8 (50%)
Chemistry	7	2 (29%)	5 (71%)	3 (43%)
Earth and space sciences	11	4 (36%)	4 (36%)	8 (73%)
Biology	28	2 (7%)	7 (25%)	7 (25%)
Biomedical research	14	2 (14%)	4 (29%)	5 (36%)
Clinical medicine	36	7 (19%)	11 (31%)	17 (47%)
Psychology	6	0 (0%)	1 (17%)	2 (33%)
Engineering	34	9 (26%)	13 (38%)	11 (32%)
Political and social sciences	11	4 (36%)	4 (36%)	5 (45%)
Economics	8	2 (25%)	4 (50%)	2 (25%)
Overall	177	43 (24%)	67 (38%)	71 (40%)

Table 4 shows the 10 SCs where the gap by *O* between Norway and Italy is highest, and vice versa. At the top of the list, in line with the above findings, we find an SC of Mathematics (Mathematics, interdisciplinary applications), where the six Norwegian professors show an *O* (3.77) about six times as high as that of their 46 Italian counterparts (0.64). The other top SCs fall in four disciplines, namely, Engineering (Remote sensing; Construction & building technology), Clinical medicine (Anesthesiology; Substance abuse), Physics (Acoustics; Thermodynamics; Physics, particles & fields), and Political and social sciences (Area studies; Political science). On the other side, the 10 SCs with the highest gap in favor of Italy fall in Engineering (seven), Clinical medicine (two), and Biomedical research (one).

Table 5 presents the same analysis by *FO*. The results are similar to those by *O*. Six of the top 10 SCs where Norway outperforms Italy remain the same, while only one changes among those where Italy outperforms Norway.

Table 4. Top 10 fields (SCs) by gap of output (*O*) in favor of Norway, and the same for Italy.

	Italy		Norway		
SC	No. of professors	*O*	No. of professors	*O*	Δ
Mathematics, interdisciplinary applications	46	0.64	6	3.77	−3.13
Remote sensing	102	0.92	8	2.04	−1.12
Acoustics	25	0.78	9	1.61	−0.83
Anesthesiology	57	0.96	3	1.79	−0.83
Construction & building technology	67	0.91	11	1.52	−0.61
Substance abuse	9	0.69	10	1.28	−0.59
Thermodynamics	97	0.98	4	1.52	−0.54
Physics, particles & fields	556	0.97	30	1.50	−0.52
Area studies	4	0.51	73	1.03	−0.52
Political science	90	0.85	43	1.32	−0.48
…					
Engineering, multidisciplinary	11	1.13	2	0.27	0.87
Materials science, composites	59	1.03	2	0.16	0.87
Computer science, hardware & architecture	10	1.08	1	0.19	0.89
Computer science, interdisciplinary applications	44	1.17	10	0.27	0.89
Medicine, research & experimental	66	1.03	2	0.12	0.91
Computer science, cybernetics	6	1.39	4	0.42	0.97
Andrology	9	1.10	1	0.13	0.97
Rehabilitation	47	1.47	38	0.42	1.05
Medical informatics	6	1.72	6	0.28	1.43
Information science & library science	13	2.04	24	0.44	1.61

Table 5. Top 10 fields (SCs) by gap of fractional output (*FO*) in favor of Norway, and the same for Italy.

	Italy		Norway		
SC	No. of professors	*FO*	No. of professors	*FO*	Δ
Mathematics, interdisciplinary applications	46	0.69	6	3.41	−2.72
Hematology	340	0.99	4	2.19	−1.21
Remote sensing	102	0.92	8	2.02	−1.09
Chemistry, inorganic & nuclear	242	0.98	4	1.92	−0.93
Thermodynamics	97	0.96	4	1.86	−0.89
Anesthesiology	57	0.96	3	1.79	−0.83
Substance abuse	9	0.60	10	1.36	−0.75
Instruments & instrumentation	184	0.97	7	1.71	−0.74
Reproductive biology	66	0.98	2	1.58	−0.60
Acoustics	25	0.85	9	1.42	−0.58
…					
Medical informatics	6	1.38	6	0.62	0.76
Computer science, interdisciplinary applications	44	1.14	10	0.37	0.77
Computer science, hardware & architecture	10	1.07	1	0.27	0.81
Medicine, research & experimental	66	1.02	2	0.20	0.82
Engineering, aerospace	75	1.01	1	0.17	0.84
Materials science, composites	59	1.03	2	0.15	0.88
Computer science, cybernetics	6	1.37	4	0.45	0.92
Engineering, multidisciplinary	11	1.15	2	0.16	0.99
Andrology	9	1.11	1	0.05	1.06
Information science & library science	13	1.87	24	0.53	1.35

4.2. Impact Analysis at Field Level

We now turn to the average impact analysis per discipline. Figure 3 presents the normalized scores of *AC*. Norwegian academics outperform Italians in six disciplines, most noticeably in Earth and space sciences (0.96 for Italy vs. 1.16 for Norway), Mathematics (0.98 vs. 1.15), and Clinical medicine (0.99 vs. 1.09). Italians outperform Norwegians in five, most noticeably in Engineering (1.01 vs. 0.90), Psychology (1.01 vs. 0.95), and Physics (1.00 vs. 0.95).

Norwegian professors publish on average in more prestigious journals than Italians in eight disciplines, the only exceptions being Physics, Chemistry, and Engineering (Figure 4).

As shown in Figure 1, Norway produces higher impact publications on average, but from Table 6 we see that this occurs in fewer than 50% of the SCs under observations (86 out of 177): in particular, in 10 of the 11 SCs of Earth and space sciences, in four of the six SCs of Mathematics, and in seven of the 11 SCs of Political and social sciences. Conversely, Italians

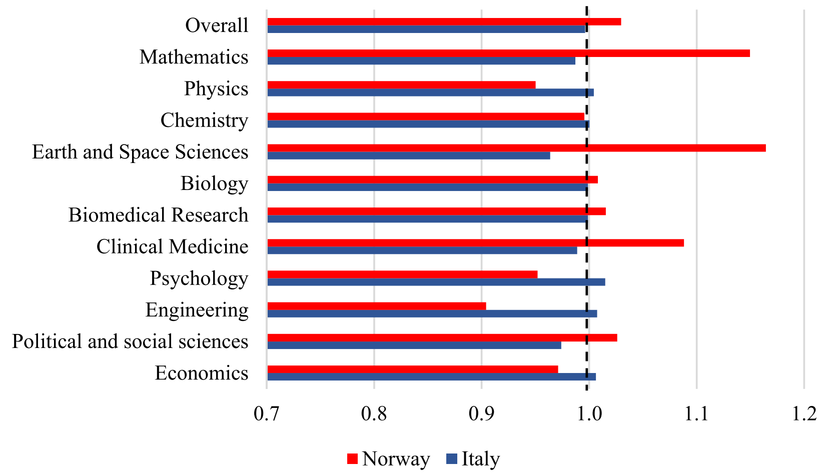

Figure 3. 2011–2015 normalized average impact of publications (*AC*) per discipline. The dashed line represents the expected value.

outperform Norwegians in six of the eight SCs of Economics, five of the seven SCs of Chemistry, and 10 of the 14 SCs of Biomedical research.

Norwegian research products are published in higher *IF* journals in 112 SCs (63%).

Table 7 reports the 10 SCs with the highest difference of *AC* for each country. In five SCs the average *AC* of Norwegian professors is above 2, while for Italians it is below 1. The highest gap is registered in History & philosophy of science, with an average *AC* for 62 Italian professors of 0.74, against 2.59 for 10 Norwegians. Of the top 10 SCs where Italy outperforms Norway, five SCs fall in Engineering (Robotics; Engineering, aerospace; Transportation science & technology; Materials science, composites; Computer science, cybernetics), and three in Economics/Political and social sciences (Public administration; Urban studies; Area studies).

Table 8 proposes the same view by the average *IF* of hosting journals per paper. Along this indicator, five of the top 10 SCs where Norway outperforms Italy fall in Clinical medicine (Andrology; Urology & nephrology; Medicine, general & internal; Anesthesiology; Medicine, legal). Similarly, five of the top 10 SCs for Italy fall in Engineering (Robotics; Medical

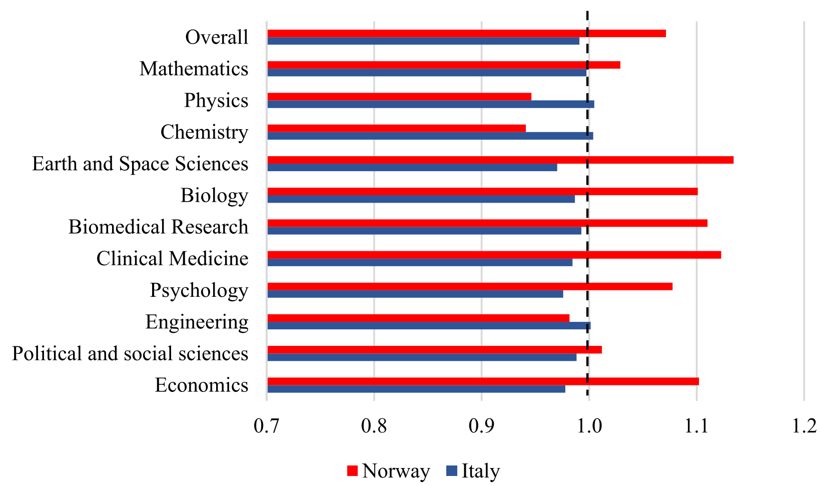

Figure 4. 2011–2015 normalized average impact factor of journals (*AIF*) per discipline. The dashed line represents the expected value.

Table 6. Number of fields (SCs) and proportion per discipline where 2011–2015 average citations per publication (*AC*), average *IF* per publication (*AIF*), and productivity (*FSS*) of Norway is higher than Italy.

		No. of SCs where Norway outperforms Italy		
Discipline	No. of SCs	*AC*	*AIF*	*FSS*
Mathematics	6	4 (67%)	5 (83%)	3 (50%)
Physics	16	6 (38%)	6 (38%)	8 (50%)
Chemistry	7	2 (28%)	3 (43%)	3 (43%)
Earth and space sciences	11	10 (91%)	10 (91%)	8 (73%)
Biology	28	16 (57%)	21 (75%)	7 (25%)
Biomedical research	14	4 (29%)	11 (79%)	5 (36%)
Clinical medicine	36	20 (56%)	28 (78%)	17 (47%)
Psychology	6	3 (50%)	4 (67%)	2 (33%)
Engineering	34	12 (35%)	15 (44%)	11 (32%)
Political and social sciences	11	7 (64%)	5 (45%)	5 (45%)
Economics	8	2 (25%)	4 (50%)	2 (25%)
Overall	177	86 (49%)	112 (63%)	71 (40%)

informatics; Computer science, interdisciplinary applications; Nanoscience & nanotechnology; Transportation science & technology).

4.3. Sensitivity Analysis

The results of any evaluation exercise depend on the adopted methodology and, in particular, on the assumptions underlying the measurement indicators used. As reported in Section 3, in this study we used both output to input indicators (*O*, *FO*, and *FSS*) and average impact indicators (*AC* and *AIF*) that are independent of input.

The typical limits embedded in bibliometric analyses apply to our study too. In particular, the new knowledge produced is not only that encoded in publications. Moreover, bibliographic repertories (such as WoS, used here) do not register all publications. The measurement of the impact of publications, before the end of their life cycle, using citation-based indicators is a prediction of their overall impact, and therefore not definitive. Citations can also be negative or inappropriate, and in any case they certify only scholarly impact, forgoing other types of impact. Finally, the results could be sensitive to the field classification schemes used for publications and, especially in the proposed work, for professors.

To these limits, for *O*, *FO*, and *FSS* we must add those associated to the assumptions on input data. When assessing the performance of individual scientists, if there are differences in the production factors available to each, then one should normalize for them. Unfortunately, relevant data are not available at the individual level, either in Italy or in Norway. So, we are forced to assume that the same resources are available to all professors within the same field. The second assumption is that the hours devoted to research are the same for all professors.

Table 7. Top 10 fields (SCs) by gap of average citation per publication (*AC*) in favor of Norway, and the same for Italy.

	Italy		Norway		
SC	No. of professors	*AC*	No. of professors	*AC*	Δ
History & philosophy of science	62	0.74	10	2.59	−1.84
Engineering, multidisciplinary	11	0.79	2	2.13	−1.33
Urology & nephrology	234	0.98	3	2.21	−1.23
Medicine, legal	92	0.94	5	2.10	−1.16
Anesthesiology	57	0.94	3	2.06	−1.12
Statistics & probability	390	0.94	29	1.83	−0.89
Remote sensing	102	0.95	8	1.66	−0.71
Mathematics, interdisciplinary applications	46	0.92	6	1.62	−0.70
Medicine, general & internal	18	0.97	1	1.62	−0.65
Physics, atomic, molecular & chemical	98	0.94	12	1.52	−0.58
…					
Robotics	56	1.01	1	0.43	0.58
Engineering, aerospace	75	1.01	1	0.39	0.61
Public administration	21	1.26	12	0.55	0.71
Urban studies	26	1.14	6	0.37	0.77
Transportation science & technology	41	1.06	3	0.24	0.81
Physics, multidisciplinary	157	1.05	11	0.23	0.82
Materials science, composites	59	1.03	2	0.18	0.85
Computer science, cybernetics	6	1.35	4	0.48	0.86
Developmental biology	22	1.04	1	0.13	0.91
Area studies	4	1.98	73	0.95	1.04

Given the characteristics of the Italian and Norwegian academic systems, this assumption appears to be acceptable. In fact, in keeping with the Humboldtian model, there are no "teaching-only" universities included in the analysis, as all professors are required to carry out both research and teaching, with a breakdown of the overall workload that, on average, seems to justify the assumption that time available for research is 50% of the total work time. However, lacking a normative reference or a qualified on-field survey attesting the validity of this assumption, we here propose a specific analysis to test how sensitive evaluation outcomes are to such an assumption. In particular, for *O*, *FO* and *FSS* we repeat the analysis at an overall level in four distinct scenarios:

- Scenario 0: the one based on the assumption adopted in this study, and whose results are shown in Figure 1;
- Scenario 1: where the time available for research is set at 40% of the total work hours for all professors in the data set;

Table 8. Top 10 fields (SCs) by gap of average journal impact (*AIF*) in favor of Norway, and the same for Italy.

	Italy		Norway		
SC	No. of professors	*AIF*	No. of professors	*AIF*	Δ
History & philosophy of science	62	0.82	10	2.14	−1.32
Andrology	9	0.92	1	1.73	−0.82
Urology & nephrology	234	0.99	3	1.72	−0.73
Agriculture, multidisciplinary	67	0.96	4	1.61	−0.65
Business, finance	101	0.90	20	1.53	−0.63
Medicine, general & internal	18	0.97	1	1.54	−0.58
Anesthesiology	57	0.97	3	1.54	−0.57
Metallurgy & metallurgical engineering	47	0.92	9	1.44	−0.52
Medicine, legal	92	0.97	5	1.49	−0.52
Horticulture	162	0.99	2	1.50	−0.50
…					
Robotics	56	1.01	1	0.72	0.29
Medical informatics	6	1.15	6	0.85	0.29
Communication	25	1.18	34	0.87	0.32
Biophysics	17	1.02	1	0.69	0.33
Computer science, interdisciplinary applications	44	1.06	10	0.72	0.34
Nanoscience & nanotechnology	32	1.01	1	0.65	0.36
Medicine, research & experimental	66	1.01	2	0.62	0.40
Transportation science & technology	41	1.03	3	0.60	0.43
Physics, multidisciplinary	157	1.03	11	0.53	0.50
Developmental biology	22	1.02	1	0.46	0.56

- Scenario 2: where the time available for research is set to 60% of the total work hours for all professors in the data set;
- Scenario 3: where the time available for research is set to 40% of the total work hours for full professors, 50% for associate and 60% for assistant (assuming that the responsibilities for management, coordination and institutional representation activities (and therefore the time dedicated to them) grows with academic rank.

It should be noted that differentiation by country would have the sole effect of increasing the productivity of professors in one country with respect to the other.

In Figure 5 we show the results of the sensitivity analysis. The variability of the average value of the indicators among the four scenarios is very limited for Italy, while it is more evident for Norway. This result certainly depends on the different size of the two countries: Norwegian professors represent 11% of the total data set, Italians 89%. In particular, for Norway

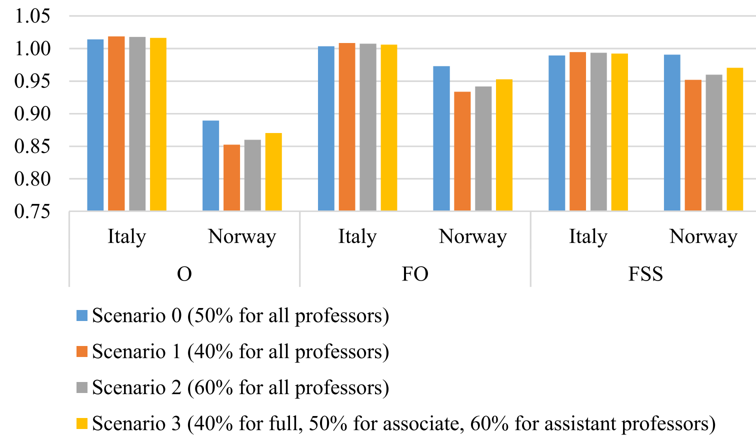

Figure 5. 2011–2015 average normalized research output, fractional output, and productivity per euro spent (*O*, *FO*, *FSS*) at overall country level, varying the share of time devoted to research.

Scenario 0 is certainly the most favorable. With respect to the single indicators, for *O* and *FO* the differences between the two countries are always quite noticeable. In contrast, for *FSS*, Scenario 0 shows practically no difference in average performance between the two countries, while in the other three scenarios the difference is more noticeable.

We repeat the same analysis at the discipline level. For each country and discipline, we provide the average value and error bars of normalized research output (Figure 6), fractional output (Figure 7), and productivity (Figure 8) for the four scenarios. The graphs confirm that the variability among the four scenarios is practically undetectable for Italy, while more evident for Norway. However, in no case does a ranking inversion between the two countries emerge. We

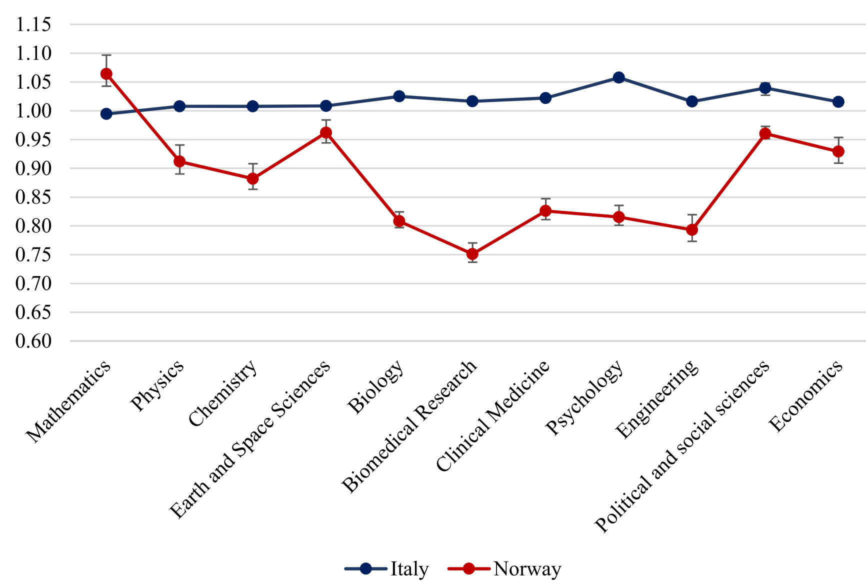

Figure 6. 2011–2015 average normalized research output per euro spent (*O*) by UDA, with error bars obtained by varying the share of time devoted to research*. *Scenario 0 (50% for all professors); Scenario 1 (40% for all professors); Scenario 2 (60% for all professors); Scenario 3 (40% for full, 50% for associate, and 60% for assistant professors).

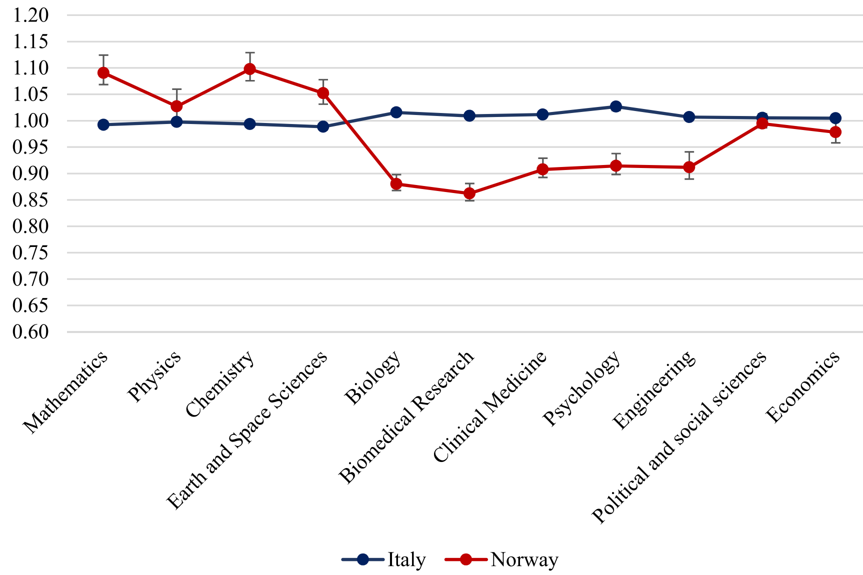

Figure 7. 2011–2015 average normalized research fractional output per euro spent (*FO*) by UDA, with error bars obtained by varying the share of time devoted to research*. *Scenario 0 (50% for all professors); Scenario 1 (40% for all professors); Scenario 2 (60% for all professors); Scenario 3 (40% for full, 50% for associate, and 60% for assistant professors).

can conclude that at the discipline level, the sensitivity of results of an evaluation exercise to assumptions about the time available for research is very limited.

Instead, at the overall level, the analysis reveals a higher sensitivity to input data and suggests attention when setting assumptions and caution in the interpretation of results.

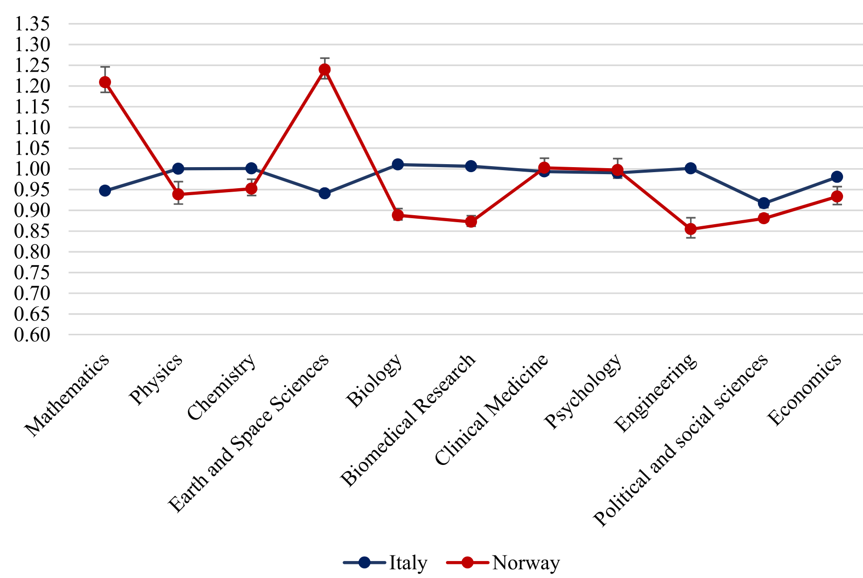

Figure 8. 2011–2015 average normalized research productivity (*FSS*), by UDA, with error bars obtained varying the share of time devoted to research*. *Scenario 0 (50% for all professors); Scenario 1 (40% for all professors); Scenario 2 (60% for all professors); Scenario 3 (40% for full, 50% for associate, and 60% for assistant professors).

5. DISCUSSION AND CONCLUSIONS

Cross-national comparison of research performance is one of the main areas of application of bibliometrics. For example, such analyses have a central position in science and technology indicator reports for monitoring scientific development and performance (see, for example, National Science Board, 2020; Norges forskningsråd, 2019; OECD, 2019). Through the great attention that these reports receive, the bibliometric indicators play an important role in the public perception of the scientific performance of countries. Our study provides new insight into more advanced methods for comparing national research performance than is provided through such reports.

Despite Italy and Norway differing considerably along such dimensions as research size and profile, and the organization of their research systems, there are hardly any differences at all in the average research productivity (*FSS*) of Italian and Norwegian professors (Abramo et al., 2020).

In this work, we delve into the various dimensions of research performance. Overall, the study shows that an average Italian professor publishes more papers than a Norwegian. This difference is quite large when measuring publications on a whole count basis, but is smaller using fractionalized measures (in which credit for a publication is allocated according to the contribution of the participating coauthors). This means that Italian publications must have more coauthors, which suggests that Italians tend to work in larger teams. This might be related to the fact that the Italian research system is much larger than Norway's. Possibly, Italian researchers may benefit from larger universities, which can favor larger institutional collaborations. The different collaboration behavior of Italian and Norwegian professors might reveal an interesting topic for future investigation. Further future research may repeat this study, restricting the observation to top performers only, to assess whether their distinctive traits along each productivity component differ from the average population.

The discussion of the appropriateness of whole versus fractionalized measures of publication output has a long history in bibliometrics, dating back to the dispute on the decline of British science during the 1980s (Braun, Glänzel, & Schubert, 1989; Irvine, Martin et al., 1985). Moed (2005) argues that the two measurement principles should be seen as complementary, where whole counts measure participation and fractionalized counts measure the number of creditable papers. Our analysis shows that the different measurement principles yield quite different results, which has also been documented in previous studies (Aksnes, Sivertsen et al., 2016; Gauffriau, Olesen Larsen et al., 2008). Currently, fractionalized measurement seems to have gained increased popularity (Waltman & van Eck, 2015), and this is also the principle underlying the *FSS* indicator.

While the number of publications produced per individual is higher on average for Italian than for Norwegian professors, the pattern is reversed when looking at the citation impact and journal profile. We observe national patterns where Italian academics produce more papers but with lower impact and in less prestigious journals. Thus, there is a quantity versus quality tension in the national publication patterns, where the first dimension apparently is more emphasized by Italians and the latter by Norwegians. However in the combined *FSS* indicators these differences tend to level out.

We are not able to provide any final answer on the possible reasons underlying these differences. The quality of the research is one obvious dimension, but there are also mechanisms in the research systems that might favor certain publication behaviors more than others. We will discuss this further below.

In Norway, there is a performance-based funding system where a bibliometric indicator is applied for the allocation of funding across institutions (Sivertsen, 2018). Although this model is designed to work at an overall national level, previous research has shown that it is sometimes applied at lower levels and may have an incentive effect at the level of individual researchers (Aagaard, 2015). Two important elements of the model should be emphasized. First, extra points are given for publications in the most prestigious publication channels; and second, all journal and book publications in accredited publication channels are included, not only publications indexed in WoS. This means that there is an incentive to publish in prestigious publication outlets and no incentive to publish in WoS journals specifically. Moreover, citations are not included as indicators in the Norwegian model.

In Italy, a proportion of public funds started being allocated to universities on the basis of the outcome of the second national research assessment exercise, VQR 2006–2010[13]. The evaluation was based on a restricted number of research products submitted by universities, and the bibliometric indicator applied to assess their relative quality was a combination of citations and IF. The underlying incentive seems to have professors focus on the production of a few very high-quality outputs, to be published in prestigious journals. In the same period, the "national scientific habilitation" (ASN) for university appointments was introduced in Italy[14]. The ASN required passing the threshold values in all or part of (depending on the field) three bibliometric indicators, namely number of articles, number of citations, and contemporary *h*-index[15].

Based on these national differences, where citations are used as a component indicator in Italy only, one might have expected Italian professors to be more highly cited. However, our overall results show the contrary. The empirical paradox may suggest that the incentive effect at the level of individual researchers may be limited, or that it takes a longer time for incentive systems to show effects. Still, it should be emphasized that the Norwegian division of publication channels by quality levels means that the Norwegian model is not a pure productivity measure. Moreover, the funding models are only a part of the picture, and citations indicators are sometimes used in other contexts as well, such as in evaluations of research and in the assessment of research proposals and candidates (Wilsdon, Allen et al., 2015). The national differences in the number of publications are also relevant to discuss in the light of the two funding models. While the Italian model is limited to WoS-indexed publications, the Norwegian system covers all publications in channels accredited as scientific or scholarly. A previous analysis showed that the WoS Core Collection covered 69% of the total Norwegian publication output, with large variations across fields (Aksnes & Sivertsen, 2019). This means that a relatively large number of the publications of the professors analyzed would not be included in the present analysis. This presumably holds for both Italy and Norway. However, as there is no specific incentive to publish in WoS journals in Norway, this issue might affect Norway more than Italy.

Furthermore, previous research has shown that there is a language bias in the coverage of WoS, which affects various countries differently (van Leeuwen, Moed et al., 2001). Although journals in national languages as a general rule are not covered by WoS, at least not in the core collection, there are many exceptions. While coverage of national journals will increase the publication numbers of a country, it will have an opposite effect on the citation impact, as

[13] https://www.anvur.it/en/homepage/ (accessed May 9, 2022).

[14] https://abilitazione.miur.it/public/index.php?lang=eng (accessed May 9, 2022).

[15] The contemporary index was introduced by Sidiropoulos, Katsaros, and Manolopoulos (2007).

articles in these journals generally tend to be little cited. There are hardly any national Norwegian journals indexed in the core collection of WoS, while for Italy there are several, and becoming more and more numerous, especially after the introduction of the VQR. Possibly, this factor too might explain some of the observed country differences across the indicators analyzed.

In interpreting the results of the performance analysis, all the usual limits, caveats, assumptions, and qualifications of evaluative scientometrics apply. In addition to the bibliographic data source and relative country coverage, these include general bibliometric issues, such as that publications are not representative of all knowledge produced, and that citations have limitations as performance measures. In addition, the results are sensitive to the classification schemes applied for the publication output and professors. Finally, due to limitations in the availability of comparable input data, some adaptations have been introduced. The sensitivity of results to varying input data suggests care when setting assumptions and caution in interpreting findings. We have shown in fact that varying the assumed 50% of professors' time devoted to research to 40% or 60% would tilt research productivity in the two countries from being equal to favoring Italy. This is particularly true at overall level, while at discipline level the sensitivity of results is very limited and the relative performance of the two countries seems very robust.

From this, there arises a call for those governments and research institutions that intend to benefit from ever more precise and reliable performance evaluations, useful for their decisions and policy-making: They should be prepared to give scientometricians the underlying input data necessary for the job. Subject to the availability of input data, the present study can be replicated to include other countries.

Studies of the output dimension in other countries would be of particular interest. For a long time, various normalization procedures have been applied for calculating citation indicators (Moed, 2005). However, there are no cross-national reference standards for comparing publication output, although previous studies have shown the field dimension to be of particular importance. For example, a study of publication patterns in social sciences and humanities in eight European countries showed large differences across fields but also across nations (Kulczycki, Engels et al., 2018). In addition, publication output has been shown to be influenced by individual variables such as gender (Abramo et al., 2021; Sugimoto, Larivière et al., 2013; Elsevier, 2020), age (Abramo, D'Angelo, & Murgia, 2016; Gingras, Larivière et al., 2008; Kyvik, 1990; Levin & Stephan, 1989), and academic rank (Abramo, D'Angelo, & Di Costa, 2011; Blackburn, Behymer, & Hall, 1978; Ventura & Mombrù, 2006). Thus, further research would be required to provide better fundamentals for cross-national analyses along this dimension of performance.

Analyses based on comparable research assessment data may give stakeholders, such as governments, policymakers, universities, business enterprises, and prospective students, valuable information on how various higher education institutions perform. In particular, in-depth investigations along single dimensions of research performance can reveal those with higher potential efficiency gains, and therefore orient the incentive systems for continuous improvement.

AUTHOR CONTRIBUTIONS

Giovanni Abramo: Conceptualization, Investigation, Methodology, Supervision, Validation, Writing—original draft, Writing—review & editing. Dag W. Aksnes: Data curation, Funding acquisition, Investigation, Methodology, Supervision, Validation, Visualization, Writing—review & editing. Ciriaco Andrea D'Angelo: Data curation, Formal analysis, Investigation, Methodology, Validation, Writing—original draft.

COMPETING INTERESTS

The authors have no competing interests.

FUNDING INFORMATION

The research project received funding by the Nordic Institute for Studies Innovation, Research and Education (NIFU), in Oslo, Norway for open access publishing.

DATA AVAILABILITY

Being subject to Clarivate-WoS license restrictions, the raw data cannot be made publicly available, but are available from the authors upon request by the reader for personal interest only.

REFERENCES

Aagaard, K. (2015). How incentives trickle down: Local use of a national bibliometric indicator system. *Science and Public Policy*, *42*(5), 725–737. https://doi.org/10.1093/scipol/scu087

Abramo, G., Aksnes, D. W., & D'Angelo, C. A. (2020). Comparison of research productivity of Italian and Norwegian professors and universities. *Journal of Informetrics*, *14*(2), 101023. https://doi.org/10.1016/j.joi.2020.101023

Abramo, G., Aksnes, D. W., & D'Angelo, C. A. (2021). Gender differences in research performance within and between countries: Italy vs Norway. *Journal of Informetrics*, *15*(2), 101144. https://doi.org/10.1016/j.joi.2021.101144

Abramo, G., Cicero, T., & D'Angelo, C. A. (2012). Revisiting the scaling of citations for research assessment. *Journal of Informetrics*, *6*(4), 470–479. https://doi.org/10.1016/j.joi.2012.03.005

Abramo, G., & D'Angelo, C. A. (2014). How do you define and measure research productivity? *Scientometrics*, *101*(2), 1129–1144. https://doi.org/10.1007/s11192-014-1269-8

Abramo, G., D'Angelo, C. A., & Di Costa, F. (2011). Research productivity: Are higher academic ranks more productive than lower ones? *Scientometrics*, *88*(3), 915–928. https://doi.org/10.1007/s11192-011-0426-6

Abramo, G., D'Angelo, C. A., & Murgia, G. (2013). The collaboration behaviors of scientists in Italy: A field level analysis. *Journal of Informetrics*, *7*(2), 442–454. https://doi.org/10.1016/j.joi.2013.01.009

Abramo, G., D'Angelo, C. A., & Murgia, G. (2016). The combined effects of age and seniority on research performance of full professors. *Science and Public Policy*, *43*(3), 301–319. https://doi.org/10.1093/scipol/scv037

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

Aksnes, D. W., Sivertsen, G., van Leeuwen, T. N. & Wendt, K. K. (2016). Measuring the productivity of national R&D systems: Challenges in cross-national comparisons of R&D input and publication output indicators. *Science and Public Policy*, *44*(2), 246–258. https://doi.org/10.1093/scipol/scw058

Alonso, S., Cabrerizo, F. J., Herrera-Viedma, E., & Herrera, F. (2009). *h*-Index: A review focused in its variants, computation and standardization for different scientific fields. *Journal of Informetrics*, *3*(4), 273–289. https://doi.org/10.1016/j.joi.2009.04.001

Blackburn, R. T., Behymer, C. E., & Hall, D. E. (1978). Research notes: Correlates of faculty publication. *Sociology of Education*, *51*(2), 132–141. https://doi.org/10.2307/2112245

Braun, T., Glänzel, W., & Schubert, A. (1989). Assessing assessments of British science—Some facts and figures to accept or decline. *Scientometrics*, *15*(3–4), 165–170. https://doi.org/10.1007/BF02017195

D'Angelo, C. A., & Abramo, G. (2015). Publication rates in 192 research fields. In A. Salah, Y. Tonta, A. A. A. Salah, & C. Sugimoto (Eds.), *Proceedings of the 15th International Society of Scientometrics and Informetrics Conference* (pp. 909–919). Istanbul: Bogazici University Printhouse.

Elsevier. (2020). *Report: The researcher journey through a gender lens*. https://www.elsevier.com/research-intelligence/resource-library/gender-report-2020?dgcid=_EC_Connect (accessed May 9, 2022).

Glanzel, W., Schubert, A. (2004). Analyzing scientific networks through co-authorship. In H. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of bibliometric indicators for science and technology*. Dordrecht: Kluwer Academic Publishers.

Gauffriau, M., Olesen Larsen, P., Maye, I., Roulin-Perriard, A., & von Ins, M. (2008) Comparisons of results of publication counting using different methods. *Scientometrics*, *77*(1), 147–176. https://doi.org/10.1007/s11192-007-1934-2

Gingras, Y., Larivière, V., Macaluso, B., & Robitaille, J. P. (2008). The effects of aging on researchers' publication and citation patterns. *PLOS ONE*, *3*(12), 1–8. https://doi.org/10.1371/journal.pone.0004048, PubMed: 19112502

Hicks, D. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics*, *44*(2), 193–215. https://doi.org/10.1007/BF02457380

Hirsch, J. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, *102*(46), 16569–16572. https://doi.org/10.1073/pnas.0507655102, PubMed: 16275915

Irvine, J., Martin, B., Peacock, T., & Turner, R. (1985). Charting the decline in British science. *Nature*, *316*(6029), 587–590. https://doi.org/10.1038/316587a0

Kyvik, S. (1990). Age and scientific productivity. Differences between fields of learning. *Higher Education*, *19*(1), 37–55. https://doi.org/10.1007/BF00142022

Kulczycki, E., Engels, T. C. E., Polonen, J., Bruun, K., Duskova, M., ... Zuccala, A. (2018) Publication patterns in the social sciences and humanities: Evidence from eight European countries. *Scientometrics*, *116*(1), 463–486. https://doi.org/10.1007/s11192-018-2711-0

Larivière, V., Archambault, É., Gingras, Y., & Vignola-Gagné, É. (2006). The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities. *Journal of the American Society for Information Science and Technology*, *57*(8), 997–1004. https://doi.org/10.1002/asi.20349

Levin, S. G., & Stephan, P. E. (1989). Age and research productivity of academic scientists. *Research in Higher Education*, *30*(5), 531–549. https://doi.org/10.1007/BF00992202

Lillquist, E., & Green, S. (2010). The discipline dependence of citation statistics. *Scientometrics*, *84*(3), 749–762. https://doi.org/10.1007/s11192-010-0162-3

Moed, H. F. (2005). *Citation analysis in research evaluation*. Dordrecht: Springer.

National Science Board. (2020). *Science and Engineering Indicators 2020. NSB-2020-1*. Alexandria, VA: National Science Foundation. Available at https://ncses.nsf.gov/pubs/nsb20201 (accessed May 9, 2022).

Norges forskningsråd. (2019). *Det norske forsknings- og innovasjonssystemet – statistikk og indikatorer 2019. [Science & Technology Indicators for Norway 2019]*. Oslo. Available at https://www.forskningsradet.no/indikatorrapporten/ (accessed May 9, 2022).

OECD. (2019). *Main science and technology indicators*. Available at https://www.oecd-ilibrary.org/science-and-technology/main-science-and-technology-indicators_2304277x (accessed May 9, 2022).

Piro, F. N., Aksnes, D. W., & Rørstad, K. (2013). A macro analysis of productivity differences across fields: Challenges in the measurement of scientific publishing. *Journal of the American Society for Information Science and Technology*, *64*(2), 307–320. https://doi.org/10.1002/asi.22746

Sandström, U., & Sandström, E. (2009). Meeting the micro-level challenges: Bibliometrics at the individual level. *12th International Conference on Scientometrics and Informetrics*, Rio de Janeiro, Brazil.

Sidiropoulos, A., Katsaros, D., Manolopoulos, Y. (2007). Generalized Hirsch *h*-index for disclosing latent facts in citation networks. *Scientometrics*, *72*(2), 253–280. https://doi.org/10.1007/s11192-007-1722-z

Sivertsen, G. (2018). The Norwegian model in Norway. *Journal of Data and Information Science*, *3*(4), 3–19. https://doi.org/10.2478/jdis-2018-0017

Sorzano, C. O. S., Vargas, J., Caffarena-Fernández, G., & Iriarte, A. (2014). Comparing scientific performance among equals. *Scientometrics*, *101*(3), 1731–1745. https://doi.org/10.1007/s11192-014-1368-6

Stringer, M. J., Sales-Pardo, M., & Amaral, L. A. N. (2010). Statistical validation of a global model for the distribution of the ultimate number of citations accrued by papers published in a scientific journal. *Journal of the American Society for Information Science*, *61*(7), 1377–1385. https://doi.org/10.1002/asi.21335, PubMed: 21858251

Sugimoto, C. R., Larivière, V., Ni, C., Gingras, Y., & Cronin, B. (2013). Global gender disparities in science. *Nature*, *504*(7479), 211–213. https://doi.org/10.1038/504211a, PubMed: 24350369

van Raan, A. F. J. (1993). Advanced bibliometric methods to assess research performance and scientific development: Basic principles and recent practical applications. *Research Evaluation*, *3*(3), 151–166. https://doi.org/10.1093/rev/3.3.151

van Leeuwen, T. N., Moed, H. F., Tijssen, R. J., Visser, M. S., & Raan, A. F. (2001) Language biases in the coverage of the science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics*, *51*(1), 335–346. https://doi.org/10.1023/A:1010549719484

Ventura, O. N., & Mombrù, A. W. (2006). Use of bibliometric information to assist research policy making. A comparison of publication and citation profiles of full and associate professors at a school of chemistry in Uruguay. *Scientometrics*, *69*(2), 287–313. https://doi.org/10.1007/s11192-006-0154-5

Vieira, E. S., & Gomes, J. A. N. F. (2010). Citations to scientific articles: Its distribution and dependence on the article features. *Journal of Informetrics*, *4*(1), 1–13. https://doi.org/10.1016/j.joi.2009.06.002

Waltman, L. (2012). An empirical analysis of the use of alphabetical authorship in scientific publishing. *Journal of Informetrics*, *6*(4), 700–711. https://doi.org/10.1016/j.joi.2012.07.008

Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, *10*(2), 365–391. https://doi.org/10.1016/j.joi.2016.02.007

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., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. J. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, *5*(1), 37–47. https://doi.org/10.1016/j.joi.2010.08.001

Wildgaard, L., Schneider, J. W., & Larsen, B. (2014). A review of the characteristics of 108 author-level bibliometric indicators. *Scientometrics*, *101*(1), 125–158. https://doi.org/10.1007/s11192-014-1423-3

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., ... Johnson, B. (2015). *The metric tide: Report of the independent review of the role of metrics in research assessment and management*. HEFCE. https://doi.org/10.4135/9781473978782

Yoshikane, F., & Kageura, K. (2004). Comparative analysis of coauthorship networks of different domains: The growth and change of networks. *Scientometrics*, *60*(3), 433–444. https://doi.org/10.1023/B:SCIE.0000034385.05897.46