Evaluating the Effectiveness of Research Grants with Journal Bibliometrics

Marcin Oleksy\textsuperscript{1,*}, Przemysław Kazienko\textsuperscript{1} and Maciej Dzieżyc\textsuperscript{1}

\textsuperscript{1}Wroclaw University of Science and Technology, Department of Artificial Intelligence, Wyb. Wyspiańskiego 27, 50-370 Wroclaw, Poland

Abstract

We present and exploit WEIG – Wroclaw Effectiveness Indicator for Grants, a scientometric meta-measure that aggregates quality of scientific papers published within a given project normalized with its budget. Variations of WEIG based on general journal quality indicators: Impact Factor (IF), Article Influence Score (AIS), and Polish Ministerial Scoring of Journals (PMSJ) are applied to analyses of projects from two public agencies: European Research Council (ERC) and Polish National Science Centre (NSC). The studies on grants ended between 2014 and 2021 revealed that NSC projects are more effective than ERC ones. The efficiency of Science Branches and Scientific Panels over years are investigated. Some limitations of the proposed approach and observed phenomena are discussed as well.

Keywords

research efficiency, grant effectiveness, funding acknowledgement analysis, WEIG, IF, AIS, Polish Ministerial Scoring of Journals, European Research Council, National Science Centre

1. Introduction

Achievements of research activities or results like grants or scientific papers can be evaluated and compared by means of quantitative bibliometric measures. However, there is no commonly agreed measure for quality of scientific outcome or effectiveness of research grants. Anyway, we still expect a simple and possibly interpretable metric to compare research centres, journals, conferences, or funding schemas. \textit{Impact factor} (IF) appears to fulfill such a function for scientific journals, and \textit{h-index} for individual scientists. On the other hand, we should be aware that quantitative scientometric applied to evaluation of individual scientists, e.g. for the promotion purposes, still remains questionable. This issue was addressed in \textit{San Francisco Declaration on Research Assessment} \cite{1} and \textit{Leiden Manifesto} \cite{2}.

Supporting fundamental research by means of public funds is considered to have positive impact on the country’s economy \cite{3}. Therefore, public financing of basic studies is necessary in the recent world and “no nation can ‘free-ride’ on the world scientific system”. This paper concentrates on fundamental research, for which the primary expected outcome are scientific...
papers.

To support economical progress and competitiveness of their economy, most developed countries maintain their own dedicated public agencies to boost basic scientific studies. It includes NSF and NIH in USA, EPSRC in the UK, DFG in Germany, ANR in France, NSFC in China, NRF in Korea, JSPS and JST in Japan, as well as some international initiatives like ERC (European Research Centre), which is a part of the European Union’s flagship Research and Innovation programme (Horizon Europe), component of the EU-long-term Multiannual Financial Framework. These public bodies support hundreds and thousands of grants, so they want to monitor the project implementation and validate their results. Therefore, we present a measure that compare the quantity of scientific papers and effectiveness of the projects in a massive manner.

For that purpose Wroclaw Effectiveness Indicator for Grants (WEIG) – a meta-measure was initially proposed in [4] along with some preliminary analyses of Polish National Science Centre (NSC) grants. Next, it was further developed and applied to compare grants funded by European Research Centre (ERC) and NSC [5], which ended by 2020. Here, we present some other analyses related to effectiveness of research grants, i.e., we consider an additional journal metric – Polish Ministerial Scoring of Journals (PMSJ), year 2021 is appended, and publications without IF are investigated.

2. Related Work

2.1. Funding Acknowledgement Analysis

The Web of Science (recently by Clarivate Analytics) added two features to the indexed scientific papers in 2008: the funding agency and grant number. These fields were analyzed already in 2011, showing also an important role of non-govermental agencies in Germany and Japan [6]. Then, it was called "funding acknowledgement analysis".

John Rigby showed a weak correlation between the number of declared funding sources and the impact of the publications measured by the number of received citations [7]. This observation was based on 3,596 articles from the Journal of Biological Chemistry published in 2009. He also found out that Web of Science has a clear advantage over the MEDLINE (PubMed) database when it comes to funding sources. MEDLINE covers only a few sources of funding, mainly from the USA [8]. He also suggested that the funding text can be used for mapping research outputs and priorities of funding bodies.

In 2017, Nicola Grassano et al. evaluated recall and precision for funding information provided by Web of Science and PubMed. Their study was based on manually annotated data about funding sources from 7,510 paper’s full-texts related to UK cancer research in 2011 [9]. They received a high recall value for WoS - 93% and 94% for precision. In contrast, recall for PubMed was 42% and precision - 96%. Also, based on the e-mail survey sent to the authors of the articles, they concluded that only 3% of papers did not reveal all funding sources.

Belén Álvarez-Bornstein et al. published a paper with the estimation of the completeness and the accuracy of Web of Science funding data [10]. In 87.8% of the articles, which included funding text data, both funding agency and grant number fields were correctly extracted. However, as they pointed out, there are problems with the precision of such data: (1) different funding
bodies can be recognised as one body, (2) there is a lack of the country for the funding body country, and (3) different grants from the same funding body can be recognised as one grant.

In 2020, Liu et al. reported improvement in funding information in Web of Science (WoS) [11]. Authors suggested that the funding acknowledgement data in WoS should be primarily used for the analysis of papers published in English. Also, a case study published in the same year revealed that funding information in WoS is more reliable than in Scopus [12].

In 2021, Álvarez-Bornstein and Montesi found in their literature review on analysed funding acknowledgements a lack of data normalisation, consistency, and completeness, especially for Pubmed and Scopus databases [13]. Although Web of Science is thought to possess the best funding acknowledgements data, it still has many problems related to papers in Social Sciences and Humanities.

2.2. Research Assessment

*The Declaration on Research Assessment (DORA)* was announced at the Annual Meeting of the American Society for Cell Biology in San Francisco in 2012 [1]. It contains the guidelines for funding agencies, institutions, publishers, organisations that supply metrics and researchers:

1. Be open and transparent by providing data and methods used to calculate all metrics.
2. Provide the data under a licence that allows unrestricted reuse, and provide computational access to data, where possible.
3. Be clear that inappropriate manipulation of metrics will not be tolerated; be explicit about what constitutes inappropriate manipulation and what measures will be taken to combat this.
4. Account for the variation in article types (e.g., reviews versus research articles), and in different subject areas when metrics are used, aggregated, or compared.

Polish National Science Centre (NSC) signed DORA in 2018 [14] and ERC in 2021 [15].

In 2015, *The Leiden Manifesto* for research metrics was announced in Nature [2].

It covers ten principles for research evaluation to avoid abuse of bibliometrics.

2.3. PBRF: Performance-based Research Funding

According to [16, 17], research funding can be treated as performance-based if: (1) is evaluated ex-post; (2) output or impact of scientific studies are estimated, and the funding or its part is based on such assessment; (3) funding is implemented either at national or regional level.

Implementation of the Performance-based research funding in Australia resulted in the greater number of scientific papers with simultaneous drop in quality of research [18]. Next, this finding was refuted in 2017 [19].

A positive effect of PBRF application on the research quantity was concluded from studies on the data from 31 countries and the 1996-2016 period [20]. The author relied primarily on bibliometric evaluation. Zacharewicz at al. discovered in [21] that 12 out of 28 EU countries exploited no PBRF, 3 had limited PBRF, 11 applied a quantitative, bibliometric assessment and 5 countries used peer review process.
3. Scientometric Measures for Journals and Grants

3.1. Journal Impact Factor, Derivatives, and Similar Measures

*Impact Factor (IF)* is a measure for scientific journal as the average number of citations normalized with the number of articles published in recent years. For a given year *y*, its *IF* sub *y* is the number of citations received in year *y* to papers published in two preceding years (*y* − 1 and *y* − 2), divided by the number of citable papers published in years *y* − 1 and *y* − 2:

\[
IF_y = \frac{\text{Citations}_{y-1} + \text{Citations}_{y-2}}{\text{CitableItems}_{y-1} + \text{CitableItems}_{y-2}}
\]  

(1)

Impact Factor is even criticized as a measure of journal quality. IF does not respect the quality of citations, includes unlimited self-citations, is incomparable between domains and concentrates on English journals. Moreover, a simple method of calculation incentivizes publishers to manipulate this factor by promoting self-citations, preferential for reviews and the decreasing number of Citable Items [22, 23]. It is also criticized for being unscientific [24].

The *Average Journal Impact Factor Percentile*, here referred as IF%, was introduced in 2016. It is a non-linear transformation of IF, which takes percentile positions of journal’s IF in each scientific category and averages it.

IF% was shown to have a smaller variation than IF and the distribution close to normal [25].

*Article Influence Score (AIS)* is defined as eigenfactor (the idea similar to Google Page Rank) divided by the number of citable articles. AIS is comparable to 5-year IF [26].

3.2. Polish Ministerial Scoring of Journals (PMSJ)

Universities and public scientific bodies in Poland are evaluated by the Ministry of Education and Science separately in 47 scientific disciplines that are grouped into branches of science: humanities (7 disciplines), engineering and technology (9), medical and health sciences (4), agricultural sciences (5), social sciences (11), natural sciences (7), theology (1), the arts (3) [27]. For the evaluation purpose, scientific entities report among others the papers published by the author affiliated at their organisation. Polish Ministerial Scoring of Journals (PMSJ) is assigned to each paper according to the list announced by the Minister of Education and Science [28]. Each journal is assigned one of the score: 200 (best), 140, 100, 70, 40, 20 (worst) as well as the list of disciplines a given journal is relevant to. In principle, the scoring is based on IF of the journal, however, it is not the absolute rule. Since PMSJ is not a linear transformation of IF, we cannot claim that sum of IF of two journals with PMSJ=100 is close to IF of one journal with PMSJ=200. To some extent, PMSJ may be considered similar to IF% but with less granularity and some manual corrections.

Results of the evaluation for years 2017-2021 are published in [29], however, the entire process has not been finished yet due to some appeals against this decision.

3.3. Wroclaw Effectiveness Indicator for Grants

*Wroclaw Effectiveness Indicator for Grants – W E I G(p)* for project *p* is a meta-measure defined as a calculable outcome *O(p)* of *p* divided by *p*’s funding *F(p)*, here expressed in millions of euros:
Table 1
Example of two projects selected from panel PE6 within Physical Sciences and Engineering, which terminated in 2017, together with their respective metrics

<table>
<thead>
<tr>
<th>Agency</th>
<th>NSC</th>
<th>ERC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grant id</td>
<td>2013/09/B/ST6/02317</td>
<td>278212</td>
</tr>
<tr>
<td>Budget [€]</td>
<td>154579</td>
<td>1638175</td>
</tr>
<tr>
<td>#papers</td>
<td>25</td>
<td>33</td>
</tr>
<tr>
<td>$\sum$ IF</td>
<td>46,266</td>
<td>184,708</td>
</tr>
<tr>
<td>WEIG-IF</td>
<td>327.841</td>
<td>112,752</td>
</tr>
<tr>
<td>$\sum$ AIS</td>
<td>13,931</td>
<td>95,960</td>
</tr>
<tr>
<td>WEIG-AIS</td>
<td>98,72</td>
<td>58,577</td>
</tr>
<tr>
<td>$\sum$ PMSJ</td>
<td>2150</td>
<td>4870</td>
</tr>
<tr>
<td>WEIG-PMSJ</td>
<td>15234.922</td>
<td>2972.820</td>
</tr>
</tbody>
</table>

$1\text{ EUR} = 4.7\text{ PLN}$

\[
WEIG(p) = \frac{O(p)}{F(p)} \tag{2}
\]

Assuming the project $p$ has a set $A_p$ of related scientific articles $a$, which are the effect of a given project $p$, and which acknowledge $p$. We can use a bibliometric quality measure $M(a)$ for each such paper $a$. By aggregating $M(a)$ for all papers resulting from project $p$, we obtain the general output quality of $p$: $O(p) = \sum_{a \in A_p} M(a)$. Then, $WEIG(p)$ is:

\[
WEIG(p) = \frac{\sum_{a \in A_p} M(a)}{F(p)} \tag{3}
\]

In this study, as $M(a)$ we analyzed three measures: IF, AIS (Sec. 3.1), and PMSJ (Sec. 3.2), however, also any other measure can be considered like IF% or the number of $a$’a citations. As a result, we tested WEIG-IF, WEIG-AIS, and WEIG-PMSJ, respectively.

WEIG can also be generalised for the set of projects $P$, i.e., for all grants from a given scientific discipline (panel) or year:

\[
WEIG(P) = \frac{O(P)}{F(P)} = \frac{\sum_{p \in P} O(p)}{\sum_{p \in P} F(p)} = \frac{\sum_{p \in P} \sum_{a \in A_p} M(a)}{\sum_{p \in P} F(p)} \tag{4}
\]

where $O(P)$ are aggregated outcomes (sum) of all projects $p \in P$, and $F(P)$ - their total funding, e.g. in millions of euros.

WEIG was initially proposed in [4] along with some preliminary analyses of NSC grants and then in-depth investigated for both considered agencies in [5]. Here, we extend it with additional analyses, a new year – 2021 and a new measure PMSJ.
4. Effectiveness of Research Grants Funded by ERC and NSC

4.1. Research Setup

A detailed description of all processing steps can be found in [5]. Overall, the process consists of: (1) collecting the data about scientific grants funded by ERC and NSC from their online databases, (2) using grant ids, the outcome papers are identified from Web of Science web service, (3) for each such paper published in the journal, a bibliometric measure is gathered from Journal Citation Report (IF, AIS) or from the online list published by Polish Ministry of Education [28]. Having bibliometric measures (IF, AIS, PMSJ) for each paper assigned to each grant, appropriate WEIG values are computed, including relative ones, Sec. 3.3. Aggregated WEIG values are calculated for scientific panel, years, and funding schemas.

The project year was defined as the year in which a given grant had ended. We considered all project’s identified publications registered by the data collection date (April 2022) while calculating the WEIG values. We believe that a reliable comparison between agencies or panels should be carried out only between grants completed in the same year.

Year 2014 was treated as a reliable starting point for our studies, since there were relatively few grants ended before this year for both agencies. Their funding schemas (many finished projects) yielded first mature results as late as 2014.

Figure 1: General effectiveness for funding agencies over time: WEIG-IF (a), WEIG-AIS (b), and WEIG-PMSJ (c), for branches of science: Physical Sciences and Engineering - PE (green), Life Sciences - LS (red), and Social Sciences and Humanities - SH (blue), separately for NSC projects (dotted lines) and ERC projects (solid lines). Mind that X denotes the last year of the grants.
4.2. Effectiveness of Science Branches

Both financing agencies distinguish three separate science branches (see [5]), which correlate:

- Social Sciences and Humanities (\(SH_{ERC}=HS_{NSC}\)),
- Physical Sciences and Engineering (\(PE_{ERC}=ST_{NSC}\)),
- Life Sciences (\(LS_{ERC}=NZ_{NSC}\)).

The WEIG values for science branches are relatively stable regardless of the measure taken (see Fig. 1). However, we can observe a general small drop over time for Life Sciences (LS) both for NSC and ERC projects. A sudden drop is characteristic of almost all branches and it is related to the COVID-19 pandemic – loss of effectiveness, which started in 2020, can be explained by the requested projects extensions. A slight increase for Physical Sciences and Engineering projects funded by NSC may be related to growing efficiency of Production and Processes Engineering and Astronomy and Space Science grants. Interestingly, there is one significant difference between WEIG-PMSJ and other metrics: Social Sciences and Humanities (SS&H) effectiveness does not differ from other branches as significantly as in the case of WEIG-IF or WEIG-AIS. On the other hand, a very large part of SS&H scientific papers is published in periodics with no IF (see Fig. 3). Overall IF may not be an adequately good metric to quantify output in this science branch. WEIG-PMSJ may be complementary to other presented metrics.

4.3. Effectiveness of Scientific Panels and Funding Schemes

The values of WEIG-IF, WEIG-AIS, WEIG-PMSJ, and funding for the projects, which ended between 2014 and 2021, were presented for Physical Sciences and Engineering (PE) as well as Life Sciences (LS) in division into individual panels, Fig. 2. Social Sciences and Humanities (SH) were disregarded as there is no correspondence between most of NSC and ERC panels for HS (see [5]). There is a similarity in the effectiveness of the respective NSC and ERC panels. NSC grants are overall more effective than ERC ones. They are also characterized by significantly smaller funding. The Universe Sciences (PE9) both for NSC and ERC provide the highest overall effectiveness. Simultaneously, Mathematics (PE1) reveals a big difference between WEIG-IF and WEIG calculations based on different measures. Computer Science and Informatics and Systems (CS&IS) and Communication Engineering (CE) grants both for NSC and ERC appear to be less effective in comparison to other Physical Sciences and Engineering panels. It may be connected with the specificity of scientific activity. A significant part of the articles for CS&IS and CE have been published in the journals with no IF (see Fig. 3). In fact, they are very often conference proceedings.

WEIG measure can indicate some important differences between various types of projects. Undoubtedly, larger projects allocate relatively more resources to infrastructure development etc. in comparison with smaller projects. This translates into effectiveness of research grants measured with journal bibliometrics. The results of the comparison of the effectiveness of the small projects for young scientists and larger projects for advanced researchers are significant in this matter (see Fig. 4). Both PRELUDIUM and Starting Grant (AdG) are more effective than bigger projects (OPUS and Advanced Grant (AdG) respectively) in terms of WEIG measure. However, smaller projects have much greater restrictions on financing the purchase or manufacture of scientific and research equipment.
Figure 2: (a-b) WEIG-IF, (c-d) WEIG-AIS, (e-f) WEIG-PMSJ, and (g-h) funding in millions of euros (1 EUR = 4.3 PLN), for all projects ended in 2014-21. Only panels from Physical Sciences and Engineering (PE) and Life Sciences (LS) are shown, separately for NSC (a, c, e) and ERC projects (b, d, f). The data refers to grants from all funding schemes and years. Panels with green bars are discussed in Sec. 4.3.
5. Discussion: Other WEIG Variations

We are aware of the limitations and problems associated with the use of journal metrics, e.g. the IF measure [30]. First of all, journal impact factor could not be simply translated into individual achievements, e.g. actual citations of individual articles. Second, IF is affected by various factors such as research field (e.g. favouring the research fields with literature that rapidly becomes obsolete) or journal and article type (e.g. review articles are heavily cited). Some limitations are related to the database, e.g. its coverage. It is necessary therefore to take into consideration the wider context while analysing and interpreting the results. It is difficult or even impossible to include all of the factors in one measure (e.g. different grant results, the characteristics of the
discipline etc.). However, some of them have such potential (e.g. exchange rate fluctuations, costs of living etc.). Simple experiments like adding cost of living index (CLI) to the formula may be the way to introduce the relevant determinants into the WEIG measure:

$$WEIG(p) = \frac{O(p)}{F(p) \cdot COI}$$ (5)

In this case, one would have to consider e.g. that the CLI varies depending on the place of living and the time of the estimation. Referring to the example in the Table 1, after assuming adequate CLI for UK and Poland, the gap between the presented projects would decrease: WEIG-IF (CLI) for grant 2013/09/B/ST6/02317 would be 12490.768 instead of 327.841, WEIG-IF (CLI) for grant 278212 would be 7836.285 instead of 112.752. Thus, the NSC-funded grant would have a score 159.40% better than the ERC grant, instead of 290.76%. Such experiments could nevertheless be a step toward improving the WEIG measure.

6. Conclusions and Future Work

There are three main findings derived from our analyses of the WEIG values for grants funded by two public agencies European Research Council (ERC) and Polish National Science Centre (NSC):

1. Polish NSC grants are commonly more effective than those supported by European Union via ERC (two times for physical and engineering as well as humanities, ca. 50% for life sciences). The difference is particularly noticeable when considering WEIG-PMSJ: three times greater;
2. WEIG calculations using different measures provide similar results, i.e., they reveal the same phenomena. However, they are complementary in some aspects;
3. There is a similarity in the effectiveness of the respective NSC and ERC panels for Physical Sciences and Engineering (PE) and Life Sciences (LS). It is connected with the specificity of scientific activity including publishing in conference proceedings.

In many cases, it is not possible to analyze all the data in detail. WEIG, in combination with other project information (such as scientific discipline, subject matter, implementing entity, project size), can be used to filter information on ongoing grants. This can apply to the searching for scientists to collaborate with, identification of promising research topics, or policy-making in the field of scientific research and technological development. Thus, it could be useful in tackling information overload in various contexts.

Future work will focus on further development and analysis of derived WEIG measures, e.g. starting from the one suggested in Sec. 5. Along with the WEIG-based measures already tested in this paper, they would provide a more comprehensive insights into research projects and can be used to minimize the impact of some selected factors like local salary level.

Acknowledgments

The authors would like to thank Web of Science Group, Clarivate Analytics for giving consent for usage and aggregation of data acquired from their databases.
This work was financed by (1) the National Science Centre, Poland, project no. 2021/41/B/ST6/04471; (2) the Polish Ministry of Education and Science, CLARIN-PL; (3) the European Regional Development Fund as a part of the 2014-2020 Smart Growth Operational Programme, CLARIN – Common Language Resources and Technology Infrastructure, project no. POIR.04.02.00-00C002/19; (4) the statutory funds of the Department of Artificial Intelligence, Wroclaw University of Science and Technology; (5) the Polish Ministry of Education and Science within the programme "International Projects Co-Funded"; (6) the European Union under the Horizon Europe, grant no. 101086321 (OMINO). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor European Research Executive Agency can be held responsible for them.

References


