



Multiplicative *versus* fractional counting methods for co-authored publications. The case of the 500 universities in the Leiden Ranking



Antonio Perianes-Rodríguez^a, Javier Ruiz-Castillo^{b,*}

^a Departamento de Biblioteconomía y Documentación, Universidad Carlos III, SClmago Research Group, Spain

^b Departamento de Economía, Universidad Carlos III, Spain

ARTICLE INFO

Article history:

Received 18 May 2015

Received in revised form 7 October 2015

Accepted 7 October 2015

Available online 19 November 2015

Keywords:

Co-authorship

Counting methods

Citation impact

ABSTRACT

This paper studies the assignment of responsibility to the participants in the case of co-authored scientific publications. In the conceptual part, we establish that one shortcoming of the full counting method is its incompatibility with the use of additively decomposable citation impact indicators. In the empirical part of the paper, we study the consequences of adopting the address-line fractional or multiplicative counting methods. For this purpose, we use a Web of Science dataset consisting of 3.6 million articles published in the 2005–2008 period, and classified into 5119 clusters. Our research units are the 500 universities in the 2013 edition of the CWTS Leiden Ranking. Citation impact is measured using the *Mean Normalized Citation Score*, and the *Top 10%* indicators. The main findings are the following. Firstly, although a change of counting methods alters co-authorship and citation impact patterns, cardinal differences between co-authorship rates and between citation impact values are generally small. Nevertheless, such small differences generate considerable re-rankings between universities. Secondly, the universities that are more favored by the adoption of a fractional rather than a multiplicative approach are those with a large co-authorship rate for the citation distribution as a whole, a small co-authorship rate in the upper tail of this distribution, a large citation impact performance, and a large number of solo publications.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

The assignment of responsibility to the participants in the case of co-authorship has been a vexing question since the beginning of Scientometrics (see Anderson et al., 1988, for an early discussion, as well as Albarrán, Crespo, Ortuño, & Ruiz-Castillo, 2010; Abramo, D'Angelo, & Rosati, 2013; Shen & Barabási, 2014; Waltman & Van Eck, 2015, and the references quoted therein). The continuous increase in co-authorship in all scientific disciplines exacerbates the problem with the passage of time.

In an important contribution, Waltman and Van Eck (2015) – hereafter WVE – focus on the comparison between the *fractional counting* and the *full counting* methods. The former assigns co-authored publications fractionally to each co-author, while the latter fully assigns co-authored publications to each co-author. WVE argue that there is a close connection between

* Corresponding author.

E-mail address: jrc@eco.uc3m.es (J. Ruiz-Castillo).

counting methods and field-normalization, that is, the correction for differences in citation practices between scientific fields. Based on an extensive theoretical and empirical analysis, they establish that properly field-normalized results cannot be obtained with full counting. In their own words, “Essentially, the problem of full counting is that co-authored publications are counted multiple times, once for each co-author, which creates an unfair advantage to fields with a lot of co-authorship and with a strong correlation between co-authorship and citations. For instance, the average full counting Mean Normalized Citation Score of all organizations or all countries active in these fields is significantly higher than one. On the other hand, fields in which co-authorship is less common or in which co-authorship does not correlate with citations are disadvantaged. Full counting yields results that are biased against organizations and countries whose activity is focused on these fields. Fractional counting does not suffer from this problem. In the case of fractional counting, each publication is counted only once, regardless of its number of co-authors, and this ensures that comparisons between fields can be made in an unbiased way” (p. 40). As for the practical implications of the choice of counting methods, WVE conclude “. . . this depends on the level of aggregation at which a bibliometric study is performed. In the case of a study at a high aggregation level, such as the level of countries or organizations (e.g., university rankings), we consider it absolutely essential to use fractional counting instead of full counting. At this level, there is a serious risk of misinterpretation of full counting results. Moreover, we believe that arguments in favor of full counting . . . are of limited relevance at a high aggregation level” (p. 40). Consequently, “We therefore recommend the use of fractional counting in bibliometric studies that require field normalization, especially in studies at the level of countries and research organizations.” (Abstract). Among fractional counting variants – all of which provide proper field-normalized results – WVE advocate the author-level or the address-line fractional counting.

However, WVE recall that in the *multiplicative counting* method co-authored publications are fully assigned to each co-author, like in full counting, but results are properly field-normalized, like in fractional counting (pp. 41–42). Both full and multiplicative counting extends as much as necessary the citation distributions of the units of analysis in question – authors, organizations, or countries. However, under full counting the overall citation distribution is maintained equal to the citation distribution of the original set of distinct articles, while in the multiplicative approach the overall citation distribution is made equal to the union of the units’ extended citation distributions.

This paper has two parts, one conceptual, and one empirical. In the conceptual part, we establish that, together with the arguments put forth by WVE, in our view a key problem with full counting is its incompatibility with the use of additively decomposable citation impact indicators. In the empirical part, following WVE’s recommendation (p. 42), we compare the fractional with the multiplicative approach. For this purpose, we use a Web of Science (WoS) dataset consisting of 3.6 million publications in the 2005–2008 period, the citations they receive over a 5-year citation window for each year in that period, and a classification system consisting of 5119 clusters (Ruiz-Castillo & Waltman, 2015). Our research units are the 500 universities in the 2013 edition of the CWTS Leiden Ranking (Waltman et al., 2012), referred to as the LR universities. There are 2.4 million distinct articles in which at least one author belongs to one of these universities. For reasons explained in Section 3, we assign these articles to the 500 LR universities following exclusively the address-line variant of the fractional and multiplicative approaches.

In the comparison between the two approaches, we investigate three issues.

- Firstly, assume that we order universities according to the percentage of co-authored publications with respect to the total, or the co-authorship rate, in the fractional case. Of course, a move from the fractional to the multiplicative approach will increase the co-authorship rate of all universities with some co-authored publications. The first question we investigate is whether this increase affects universities in a widely different manner. In other words, we investigate the importance of re-rankings when we order universities by the co-authorship rate in the multiplicative approach.
- Secondly, although changes in co-authorship patterns constitute a natural first step, we cannot stop here. We want to investigate whether the change in counting methods causes a great change in the ranking of universities by citation impact. For this purpose, we evaluate citation impact according to two commonly used indicators: the Mean Normalized Citation Score (MNCS hereafter) and the *Top 10%* indicator, defined as the percentage of an institution’s scientific output included in the set formed by the 10% of the most highly cited publications in the world.¹
- Thirdly, given the change in co-authorship and citation impact patterns, we investigate a new issue in this debate. We want to analyze which type of university is more likely to benefit from a move from the fractional to the multiplicative method (or *vice versa*). Naturally, there are several university characteristics worth investigating. For example, we can ask whether universities with a greater co-authorship rate, a greater citation impact, or a greater number of solo articles are the gainers or losers with the change from the fractional to the multiplicative approach. To study this issue involving several variables we use multiple regression techniques.

The rest of the paper is organized into three sections. Section 2 serves two purposes: it introduces the citation impact indicators and the counting methods studied in this paper, and it clarifies the nature of a new shortcoming precluding the use of full counting in practical applications. Section 3 presents the data, and the empirical results comparing the fractional

¹ The Top 10% indicator is used in the Leiden Ranking (www.leidenranking.com), and the SCImago Institutions Rankings (www.scimagoir.com).

Download English Version:

<https://daneshyari.com/en/article/523096>

Download Persian Version:

<https://daneshyari.com/article/523096>

[Daneshyari.com](https://daneshyari.com)