Journal of Informetrics 9 (2015) 974–989

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

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

Antonio Perianes-Rodriguez[a], Javier Ruiz-Castillo[b,*]

[a] *Departamento de Biblioteconomía y Documentación, Universidad Carlos III, SCImago 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.

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).

http://dx.doi.org/10.1016/j.joi.2015.10.002

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).

and the multiplicative approaches, while Section 4 discusses the results and offers some conclusions. To save space, some statistical and graphical material is available in Supplementary Material Section, SMS hereafter.

2. Citation impact indicators and counting methods

2.1. Notation and citation impact indicators

Suppose we have an initial citation distribution $\mathbf{C} = \{c_l\}$ consisting of N distinct publications, indexed by $l = 1, \ldots, N$, where c_l is the number of citations received by publication l. A *classification system* is an assignment of publications in $\mathbf{C}$ to J fields, indexed by $j = 1, \ldots, J$. Let I be the number of research units, indexed by $i = 1, \ldots, I$. In this section, the assignment of publications in $\mathbf{C}$ to the I research units is taken as given. Let c_{ijk} be the number of citations received by the kth article of unit i in field j. Then $\mathbf{c_{ij}} = \{c_{ijk}\}$ denotes the *citation distribution of unit i in field j*, while $\mathbf{c_j}$ denotes the *citation distribution of field j*, that is, the union of all research units' citation distributions in that field: $\mathbf{c_j} = \cup_i\{\mathbf{c_{ij}}\}$. We assume that the assignment of publications in $\mathbf{C}$ to the I research units is such that the set of distributions $\mathbf{c_{ij}}$, $i = 1, \ldots, I$, form a partition of $\mathbf{c_j}$. Of course, $\mathbf{C} = \cup_i\cup_j\{\mathbf{c_{ij}}\} = \cup_j\{\mathbf{c_j}\}$, and the total number of articles in the overall citation distribution is $N = \sum_i\sum_j N_{ij} = \sum_j N_j$, where N_{ij} is the number of articles in distribution $\mathbf{c_{ij}}$, and $N_j = \sum_i N_{ij}$ is the total number of articles in field j.

In our context, where in every field j we have $\mathbf{c_j} = \cup_i\{\mathbf{c_{ij}}\}$, the evaluation of any citation distribution is done by taking into account a key characteristic of distribution $\mathbf{c_j}$, say θ_j. Thus, a *citation impact indicator* is a function F defined in the product space of all citation distributions and the characteristic space, so that – given the characteristic $\theta_{\mathbf{j}}$ – the expression $F_{ij} = F(\mathbf{c_{ij}}; \theta_{\mathbf{j}})$ denotes the citation impact of unit i in field j, while $F_j = F(\mathbf{c_j}; \theta_{\mathbf{j}})$ denotes the citation impact of field j as a whole. To clarify this notion, consider the following two indicators that will be used in this paper. In order to be able to compare citation distributions of different size of research units working in the same field, as well as the citation impact achieved by research units publishing in different fields, both indicators are size- and scale-independent.

1. Let μ_{ij} and μ_j be the mean citation of distributions $\mathbf{c_{ij}}$ and $\mathbf{c_j}$, respectively. The *Relative Citation Rate, RCR*, is defined as

$$RCR_{ij} = RCR(c_{ij}; \mu_j) = \frac{\mu_{ij}}{\mu_j}. \tag{1}$$

   In this case, $\theta_{\mathbf{j}} = \mu_j$. For field j as a whole, $RCR_j = RCR(\mathbf{c_j}; \mu_{\mathbf{j}}) = \mu_j/\mu_j = 1$.
2. Let $X_{\mathbf{j}}$ be the set of the 10% most cited articles in citation distribution $\mathbf{c_j}$, and let x_{ij} be the sub-set of articles in $X_{\mathbf{j}}$ corresponding to unit i, so that $X_{\mathbf{j}} = \cup_i\{x_{ij}\}$ with x_{ij} non-empty for some i. If n_{ij} is the number of articles in $X_{\mathbf{ij}}$, then the *Top 10% indicator, T*, is defined as

$$T_{ij} = T(c_{ij}; X_j) = \frac{n_{ij}}{N_{ij}}. \tag{2}$$

   In this case, $\theta_{\mathbf{j}} = X_j$. If $n_j = \sum_i n_{ij}$ is the number of articles in $X_{\mathbf{j}}$, then for field j as a whole, $T_j = T(\mathbf{c_j}; X_{\mathbf{j}}) = n_j/N_j = 0.10$.

2.2. The additive decomposability property

The following property, introduced by Foster, Greeer, and Thorbecke (1984) in the context of economic poverty, plays a key role in this paper. Given θ, an indicator F is said to be *additively decomposable* if for any partition of a citation distribution $\mathbf{c}$ into G disjoint subgroups, indexed by $g = 1, \ldots, G$, the citation impact of distribution $\mathbf{c}$ can be expressed as follows:

$$F(c;\theta) = \sum_g \left(\frac{n_g}{n}\right) F(c_g;\theta),$$

where n_g is the number of publications in subgroup g, and $n = \sum_g n_g$ is the number of publications in distribution $\mathbf{c}$. To illustrate the usefulness of this property, introduced in citation analysis by Albarrán, Ortuño, and Ruiz-Castillo (2011), consider the following two situations in which the indicator F is assumed to be size- and scale-independent.

A. Under our assumptions, in every field j the distributions $\mathbf{c_{ij}}$, $i = 1, \ldots, I$, constitute a partition of $\mathbf{c_j}$. If F is additively decomposable, then we can write

$$F(c_j;\theta_j) = \sum_i \left(\frac{N_{ij}}{N_j}\right) F(c_{ij};\theta_j). \tag{3}$$

This is a very natural condition, indicating that the citation impact of field j as a whole can be expressed as the weighted average of the research units' citation impact under a common θ_j.

B. Assume that country v consists of R regions, indexed by $r = 1, \ldots, R$, and assume that the R citation distributions in field j, $\boldsymbol{c_{vrj}}$, form a partition of the citation distribution of country v in that field, $\boldsymbol{c_{vj}}$. If F is additively decomposable, then we can write

$$F(c_{vj}; \theta_j) = \sum_r \left(\frac{N_{vrj}}{N_{vj}} \right) F(c_{vrj}; \theta_j), \tag{4}$$

where N_{vrj} is the number of publications in region r, so that $N_{vj} = \sum_r N_{vrj}$. Eq. (4) indicates that the citation impact of country v in field j can be expressed as the weighted average of the regions' citation impact in field j under a common θ_j.

To assess the importance of the additive decomposability property, it is essential to introduce the following weaker property. As before, consider any partition of a citation distribution $\boldsymbol{c}$ into G disjoint subgroups, indexed by $g = 1, \ldots, G$, in which case we can write $\boldsymbol{c} = \cup_g \{\boldsymbol{c_g}\}$, where $\boldsymbol{c_g}$ is the citation distribution of subgroup g. Given θ, an indicator F is said to be *subgroup consistent* if the overall citation impact of distribution $\boldsymbol{c}$, $F(\boldsymbol{c}; \theta)$, increases whenever the citation impact of one of the subgroups, say $F(\boldsymbol{c_g}; \theta)$, increases while the citation impact of all other subgroups, $F(\boldsymbol{c_{g}}'; \theta)$ for all g' different from g, remain constant. Thus, subgroup consistency merely ensures that the aggregate, or overall citation impact value does not respond perversely to changes in the level of citation impact within one subgroup while the level of the others stays constant.

From a practical point of view, consider for example situation B, where the object of study is the citation distribution of articles in a certain scientific field published by regions in a certain country. Subgroup consistency is needed to coordinate the efforts of the country's decentralization strategy toward, say, a citation impact increase in the field in question. Such a strategy may typically involve a set of policy measures targeted at specific regions. If the citation impact indicator F is not subgroup consistent, then we may be faced with a situation in which each targeted region achieves the objective of increasing its own citation impact level, and yet the country's citation impact level decreases. Subgroup consistency may therefore be viewed as an essential counterpart to a coherent citation impact policy program. Naturally, a similar justification for subgroup consistency can be given in other contexts. For example, if F is not subgroup consistent in situation A, then it may be the case that the research impact of some research units in a scientific field increases, while the research impact of the remaining units stays constant, and yet the entire field's citation impact level decreases. We conclude that subgroup consistency can be viewed as a highly desirable property for a citation impact indicator to have in many practical contexts.[2]

Of course, decomposable measures are also subgroup consistent, but not *vice versa*. However, consider the class of *weakly monotonic* indicators where a new citation of an article does not decrease the citation impact of a citation distribution. In the context of economic poverty, Foster and Shorrocks (1991) show that, under weak monotonicity, continuity, and some other reasonable conditions, subgroup consistent and decomposable measures order equally all citation distributions.[3] For those that regard additive decomposability as putting too detailed a restriction on the functional form of a citation impact index, this result justifies, from an ordinal point of view, the use of indicators satisfying this property. For, corresponding to each continuous and weakly monotonic subgroup consistent index, there is a continuous and weakly monotonic decomposable index that ranks distributions in precisely the same way.

Given a partition, if we are merely interested in comparing the subgroup citation impact levels with one another, the decomposability requirement is quite unnecessary. On the other hand, if the analysis involves comparisons between subgroup and overall levels, then decomposability can be very useful indeed. To appreciate this, note that Eq. (3) can be written as follows:

$$\sum_i \left(\frac{N_{ij}}{N_j} \right) \left[\frac{F(c_{ij}; \theta_j)}{F(c_j; \theta_j)} \right] = 1,$$

so that the value one can serve as a benchmark for evaluating the research units in the usual way. The same can be said of Eq. (4). On the other hand, as explained in Albarrán et al. (2011), decomposability can also be useful to express, say, the citation impact differences between two distributions in two periods of time as the sum of two terms involving differences in publication shares and differences in subgroups' citation impact levels. Finally, note that the two indicators introduced in expressions 1 and 2 are additively decomposable.

2.3. Counting methods

The assignment of responsibility becomes problematic when some of the N distinct articles in $\boldsymbol{C}$ are co-authored by two or more research units. Let us begin by distinguishing between the following two counting methods.

---

[2] It should be noted that subgroup consistency and additive decomposability require strong doses of independence or autonomy among subgroups in all conceivable partitions. See Sen (1992, p. 106) for criticisms of this notion in an economic context.

[3] Given a Critical Citation Line (CCL), define low- and high-impact articles as those that receive a number of citations smaller than or greater than the CCL. In turn, low- and high-impact indicators are defined over lowand high-impact articles. The original result in Foster and Shorrocks (1991) refers to poverty measures, or lowimpact indicators. This result is readily extended to high-impact indicators. To simplify the exposition, in this article we have ignored this distinction (for a full discussion, see Appendix A in Albarrán et al., 2011).

Table 1
Example involving a single field.

Distinct publications	Authors	Number of raw citations
Publication 1	Unit A	3
Publication 2	Unit A	6
Publication 3	Unit B	1
Publication 4	Units A and B	10

(i) In the fractional counting approach, each co-authored publication in $\boldsymbol{c_j}$ is fractionally assigned to each co-author. The weight with which a publication is assigned to a co-author indicates the share of the publication allocated to that co-author. The sum of the weights of all co-authors of a publication equals one. Let $\boldsymbol{c^f}_{\boldsymbol{ij}}$ be unit i's citation distribution in field j in the fractional case, and let w_{ij} be the fractional number of publications in $\boldsymbol{c^f}_{\boldsymbol{ij}}$. Of course, for all fields j, $\boldsymbol{c_j} = \cup_i \{\boldsymbol{c^f}_{\boldsymbol{ij}}\}$, and $N_j = \sum_i w_{ij}$. Consequently, $\boldsymbol{C} = \cup_j \cup_i \{\boldsymbol{c^f}_{\boldsymbol{ij}}\}$, and $N = \sum_j \sum_i w_{ij}$.

(ii) In the full counting approach, each co-authored publication in $\boldsymbol{c_j}$ is fully assigned to each co-author. Let $\boldsymbol{c^{FC}}_{\boldsymbol{ij}}$ be unit i's citation distribution in field j in the full counting case, and let $N^{FC}{}_{ij}$ be the number of publications in $\boldsymbol{c^{FC}}_{\boldsymbol{ij}}$. The citation distribution in each field in the full counting case, $\cup_i \{\boldsymbol{c^{FC}}_{\boldsymbol{ij}}\}$, does not coincide with $\boldsymbol{c_j}$, and the sum of publications, $N^{FC}{}_j = \sum_i N^{FC}{}_{ij}$, is typically larger than N_j. Consequently, $\boldsymbol{C}$ is not equal to the union $\cup_j \cup_i \{\boldsymbol{c^{FC}}_{\boldsymbol{ij}}\}$, and $N^{FC} = \sum_j \sum_i N^{FC}{}_{ij}$, is typically larger than N.

To illustrate the problem with full counting, WVE find it convenient to distinguish between two field normalization concepts. "*Weak field normalization requires the average of the normalized citation scores of all publications in a field to be equal to one. Strong field normalization is more demanding. It requires the weighted average of the MNCS of all countries active in a field to be equal to one, where the weight of a country is given by its number of publications in the field*." (p. 15). As shown by WVE's examples, full counting is in agreement with the idea of weak normalization, but it violates the idea of strong field normalization.

In our view, this proposal warrants the following two comments. Firstly, the weak field-normalization condition is only satisfied in the case studied by WVE, namely the standard field-normalization procedure in which field mean citations are used as normalization factors. However, this condition need not be satisfied in any other normalization context. For example, it is not satisfied in four of the field-normalization procedures studied in Li, Castellano, Radicchi, and Ruiz-Castillo (2013). Secondly, quite independently of the previous point, we will presently establish that, for exhibiting one key full counting shortcoming, it suffices to examine the situation in a single field prior to any normalization of raw citations scores in this or any other field. The reason is as follows. Consider any additively decomposable indicator F. Except in the trivial case where all publications have the same number of citations, as long as $\boldsymbol{c_j} \neq \cup_i \{\boldsymbol{c^{FC}}_{\boldsymbol{ij}}\}$, so that the citation distributions $\boldsymbol{c_{ij}}$, $i = 1, \ldots, I$, do not constitute a partition of $\boldsymbol{c_j}$, we have:

$$\sum_i \left(\frac{N_{ji}^{FC}}{N_{ji}^{FC}}\right) [F(c_{ij}^{FC};\theta_j) \neq F(c_j;\theta_j)],$$

or

$$\sum_i \left(\frac{N_{ji}^{FC}}{N_{ji}^{FC}}\right) \left[\frac{F(c_{ij}^{FC};\theta_j)}{F(c_j;\theta_j)}\right] \neq 1.$$

In WVE's terminology, the difference $\sum_i (N^{FC}{}_{ij}/N^{FC}{}_j)[F(\boldsymbol{c^{FC}}_{\boldsymbol{ij}};\theta_j)/F(\boldsymbol{c_j};\theta_j)] - 1$ is the full counting bonus.[4] Naturally, in the general case with several scientific fields, the appearance of a set of full counting bonus of different size in each field only worsens the situation.

We will illustrate this flaw of the full counting method using the example in Table 1 (taken from Table 6 in WVE, p. 11). We do this in three steps. Firstly, we describe how each of the three counting methods organizes the data. Secondly, as WVE, we apply the *MNCS* indicator to illustrate the full counting method flaw. Thirdly, to demonstrate that, as long as the indicator is additively decomposable, the problem is independent of the citation indicator we care to use, we apply the *Top 50%* indicator.

The organization of the data according to the three approaches

- Under fractional counting, the units distributions are $\boldsymbol{c^f}_{\boldsymbol{A}} = (3, 6, ½ \text{ of } 10)$, and $\boldsymbol{c^m}_{\boldsymbol{B}} = (1, ½ \text{ of } 10)$, so that $\boldsymbol{C} = \boldsymbol{c^f}_{\boldsymbol{A}} \cup \boldsymbol{c^f}_{\boldsymbol{B}} = (1, 3, 6, 10)$, whose mean is $\mu(\boldsymbol{C}) = 20/4 = 5$.
- Under full counting, $\boldsymbol{c^{FC}}_{\boldsymbol{A}} = (3, 6, 10)$, and $\boldsymbol{c^{FC}}_{\boldsymbol{B}} = (1, 10)$, but $\boldsymbol{C}$ is still used at the aggregate level.

[4] In practice, as pointed out in WVE, the full counting bonus is typically positive independently of the citation impact indicator we use.

- In the multiplicative approach, like in the full counting case, each co-authored publication in the dataset is fully assigned to each co-author. Therefore, like in the full counting case, $\boldsymbol{c^{FC}}_{\boldsymbol{i}}$ is unit *i*'s citation distribution in the multiplicative case. The difference is that the overall citation distribution is made equal to the union of the units' citation distributions, that is, $\boldsymbol{C^m} = \boldsymbol{c^{FC}}_{\boldsymbol{A}} \cup \boldsymbol{c^{FC}}_{\boldsymbol{B}} = (1, 3, 6, 10, 10)$, whose mean is $\mu(\boldsymbol{C^m}) = 30/5 = 6$.

Results using the MNCS

- Under fractional counting (equations 3, 4, and 5 in WVE):

$$MNCS_A^f = 1.12, \quad MNCS_B^f = 0.80,$$

so that

$$\left(\frac{2.5MNCS_A^f + 1.5MNCS_B^f}{4}\right) = 1.$$

However, under full counting, using $\mu(\boldsymbol{C}) = 5$ in the computation of the *MNCS* values one obtains (equations 1, 2, and 6 in WVE):

$$MNCS_A^{FC} = 1.27, \quad MNCS_B^{FC} = 1.10,$$

so that

$$\left(\frac{3MNCS_A^{FC} + 2MNCS_B^{FC}}{5}\right) = 1.20.$$

- Under multiplicative counting, using $\mu(\boldsymbol{C^m}) = 6$ in the computation of the *MNCS* values one obtains:

$$MNCS_A^m = \left(\frac{1}{3}\right)\left(\frac{19}{6}\right) = \frac{19}{18} = 1.05,$$

and

$$MNCS_B^m = \left(\frac{1}{2}\right)\left(\frac{11}{6}\right) = \frac{11}{12} = 0.92,$$

so that

$$\left(\frac{3MNCS_A^m + 2MNCS_B^m}{5}\right) = 1.$$

Note that, $MNCS^k{}_A > MNCS^k{}_B$, $k = f, FC, m$. That is, when we use the *MNCS* indicator, unit *A* performs better than unit *B* in the three approaches.

Results using the Top 50%

- The top 50% publications in distribution $\boldsymbol{C}$ are (6, 10). Therefore, under fractional counting: $T^f{}_A = 1.5/2.5$, and $T^f{}_B = 0.5/1.5$, so that

$$\left(\frac{2.5}{4}\right)\left(\frac{1.5}{2.5}\right) + \left(\frac{1.5}{4}\right)\left(\frac{0.5}{1.5}\right) = \frac{2}{4} = 0.5.$$

- In our interpretation, under full counting the *Top 50%* values are $T^{FC}{}_A = 2/3$, and $T^{FC}{}_B = 1/2$, so that $\left(\frac{3}{5}\right)\left(\frac{2}{3}\right)+\left(\frac{2}{5}\right)\left(\frac{1}{2}\right)=\frac{3}{5}=0.6\neq 0.5$.
- Under multiplicative counting, the top 50% publications in $\boldsymbol{C^m}$ are (1/2 of 6, 10, 10). Therefore, $T^m{}_A = 1.5/3 = 1/2$, and $T^m{}_B = 1/2$, so that

$$\left(\frac{3}{5}\right)\left(\frac{1}{2}\right)+\left(\frac{2}{5}\right)\left(\frac{1}{2}\right)=\frac{1}{2}=0.5.$$

Note that, $T^k{}_A > T^k{}_B$, $k = f$, FC. That is, according to the *Top 50%* indicator unit *A* performs better than unit *B* under the fractional and full counting approaches. But this is only because the full counting approach works as if there are three publications in the top 50% when there are only two top 50% publications in distribution *C*. When the logic of the full counting approach is taken to the end, so that the set *C* is extended to C^m, then there are 2.5 top 50% publications, and $T^m{}_A = T^m{}_B$.

Note that the example serves three purposes. Firstly, it shows that – given a counting method – research units may receive different evaluations when we use different citation impact indicators. In particular, under the multiplicative approach, unit *A* performs better than unit *B* according to the *MNCS* indicator, but the performance of both units is the same under the *Top 50%* indicator. Thus, in spite of the fact that the *MNCS* and the *Top 50%* indicators seem to be very similar, the example shows that there are instances in which they lead to different research units' evaluations. To enhance the robustness of our results, in the empirical part of the paper we use two citation impact indicators. Secondly, the example illustrates that, given a citation impact indicator, such as the *Top 50%*, research units may be ranked differently depending on whether we use the fractional or the multiplicative approach. Using a large dataset, we will investigate this issue systematically below in Section IV.3. Thirdly, the example shows that the full counting method fails to satisfy the additively decomposability property regardless of the citation impact indicator we care to use. Instead, the fractional and the multiplicative approaches do not have this problem in either case.

In brief, full counting is incompatible with evaluating research units using additively decomposable citation impact indicators. In addition, as demonstrated by WVE, properly field-normalized results cannot be obtained with full counting. Moreover, as also shown by WVE, full counting can seriously bias the empirical results at high aggregate levels. Consequently, in our view, full counting should not be used in practice. However, those interested in fully assigning co-authored publications to each co-author can maintain this practice within the multiplicative approach. Nevertheless, adopting the multiplicative approach entails abandoning the idea that the overall citation distribution is maintained equal to the original set of distinct articles. Instead, under this approach the overall citation distribution is considerably enlarged because it is made equal to union of the units' extended citation distributions.

Note that other counting methods different from the fractional and the multiplicative ones require additional information.[5] Quite apart from the fact that we do not have this information, the remaining of this paper focuses solely on a comparison between these two readily applicable approaches. However, we should not ignore the strong limitations of these simple methods. The fractional approach does not take into account the fact that authors' contributions are never equal and hence dilutes the credit of the intellectual leader, while the multiplicative approach is biased toward researchers with multiple collaborations or large teams, customary in experimental particle physics or genomics (Shen & Barabási, 2014).

3. Data, and characteristics of university distributions under the fractional and multiplicative counting methods

3.1. The data and descriptive statistics

Our dataset results from the application of the publication-level methodology introduced in Waltman and Van Eck (2012) to 9,446,622 distinct articles published in 2003–2012. Publications in local journals, as well as popular magazines and trade journals have been excluded. We work with journals in the sciences, the social sciences, and the arts and humanities, although many arts and humanities journals are excluded because they are of a local nature. In this paper, we use the classification system recommended in Ruiz-Castillo and Waltman (2015), consisting of 5119 clusters. We focus on the set of 3,614,447 distinct articles published in 2005–2008. In terms of the notation introduced in Section II.1, we have $\boldsymbol{C} = \cup_j\{\boldsymbol{c_j}\} = (c_1, \ldots, c_N)$ with $J = 5119$, and $N = 3{,}614{,}447$. Citation distributions refer to the citations received by these articles over a five-year citation window for each year in that period. To save space, descriptive statistics of this dataset are available in Ruiz-Castillo and Waltman (2015).

[5] Consider, for example, the following three alternatives in which the scientific credit is allocated (a) according to the order in which the authors appear in the publication's byline (Abramo et al., 2013; Hagen, 2008; Stallings et al., 2013; Zhang, 2009), (b) solely to the corresponding author (Moya-Anegón et al., 2013), and (c) according to the author's contribution as perceived by the scientific community (Shen & Barabási, 2014). WVE also study first author and correspondent author counting.

Table 2
Distribution by number of address lines and mean normalized citations of the total number and the top 10% of distinct articles.

Address lines	All distinct articles			Top 10% distinct articles		
	Articles (1)	% (2)	Mean citation (3)	Articles (4)	% (5)	Mean citation (6)
1	725,608	30.0	8.2	65,403	27.0	37.1
2	742,510	30.7	8.7	69,050	28.5	38.0
3	462,539	19.1	9.7	44,388	18.3	40.6
4	238,882	9.9	11.4	25,155	10.4	46.3
5	115,454	4.8	13.1	13,925	5.8	49.8
6	57,340	2.4	15.0	7,900	3.3	54.8
7	29,649	1.2	17.3	4,766	2.0	58.7
8	16,208	0.7	19.2	2,955	1.2	61.9
≥9	31,864	1.3	113.6	8,463	3.5	185.8
Total	2,420,054	100	106.5	242,006	100	175.3

Let us focus on the 2,420,054 distinct articles – or 67% of the 3.6 million articles published in 2005–2008 – with at least one address line belonging to an LR university. The distribution of this total by the number of address lines, as well as the evolution of mean normalized citations as we increase the number of address lines is in columns 1 to 3 in Table 2. For later reference, the same information for the top 10% of most cited articles is in columns 4 to 6 in Table 2. Two points need to be emphasized. Firstly, the percentage of articles with a single address line, *solo articles* hereafter, is 30% of the total. Interestingly, this percentage is only slightly lower in the top 10% of the distribution. Secondly, as expected, mean normalized citations steadily increase with the number of address lines in both distributions, but at a very small rate.

3.2. Counting methods

Of course, each solo article is fully assigned to the corresponding LR university in both counting methods. The number of solo articles in each university is in column 1 in Table A in the SMS. The key question in this paper is how to assign the remaining 70% of the 2.4 million distinct articles that are co-authored by two or more institutions.

We know the total number of address lines of every publication, but we have information about the number of address lines of specific institutions only for the 500 LR that have at least 500 publications in the 2005–2008 period. Therefore, we cannot identify small-sized universities and, more importantly, key non-teaching research institutes in many countries of the world. In other words, we are restricted to working with 500 institutions, which is a number well below I – the total number of research units in the notation introduced in Section 2.1. As explained in WVE, the reason is that performing a comprehensive unification of organization names is extremely time consuming and, therefore, not feasible. Consequently, it is not possible to use the organization-level fractional (or multiplicative) counting method.[6]

The address-level fractional counting method works as follows. If a publication is co-authored by two or more LR universities, then it is assigned fractionally to all of them in proportion to the number of address lines in each case. For example, if the address list of an article contains five addresses and two of them belong to a particular university, then 0.4 of the article is assigned to this university, and only 0.2 of the article is assigned to each of the other three universities. Finally, consider a publication co-authored by an LR university and an unknown number of other institutions outside the Leiden Ranking. Assume, for example, that the publication has four address lines, two of which correspond to the LR university. In this case, only 0.5 of the article will be assigned to the LR university. The total number of articles in the LR universities according to the address-level fractional counting method is 1,886,106.1, or 77.9% of the 2.4 million articles with at least one address line belonging to a LR university, and 52.2% of the 3.6 million articles published in 2005–2008. Consequently, the percentage of co-authored articles decreases to 61.5%. The distribution of the 1.9 million articles among the 500 universities, as well as the percentage of co-authored articles, or the co-authorship rate, is in columns 2, and 3 in Table A in the SMS, where universities are ordered by the co-authorship rate in column 2.

Next, we turn to the address-level multiplicative counting method. In the presence of co-authorship, each LR university with v_u address lines is assigned v_u articles. If the article has a total of address lines, v, greater than or equal to the sum $\sum_u v_u$ over the LR universities, then the article is multiplied v times. The total number of articles in the LR universities according to the address-level multiplicative counting method is 4,351,584, or 179.8% of the 2.4 million articles with at least one address line belonging to a LR university, and 120.4% of the 3.6 million articles published in 2005–2008. Consequently, the percentage of co-authored articles increases to 83.3%. The distribution of the 4.3 million articles among the 500 universities, as well as the co-authorship rate are in columns 4 and 5 in Table A in the SMS.

[6] For an overview of fractional counting methods, see Section 2.1 in WVE. For an empirical comparison at the level of countries between full counting, address-level fractional counting, and three other fractional counting methods, see Section 5.2 in WVE. The results of the comparison between full counting and fractional counting are very robust to the type of fractional counting method used. In this paper, as in Section 5.1 in WVE at the LR university level, we restrict ourselves to the address-level counting method.

Table 3A
Re-rankings between universities classified by co-authorship rate when moving from the fractional to the multiplicative approach.

	First 100 universities	Remaining 400 universities	Total	%
>50 positions	21	90	111	22.2
26–50	32	60	92	18.4
16–25	23	119	142	28.4
6–15	20	93	113	22.6
≤5 positions	2	20	22	4.4
No change	2	18	20	4.0
Total	100	400	500	100.0

Table 3B
Changes in university co-authorship rate when moving from the fractional to the multiplicative approach.

	First 100 universities	Remaining 400 universities	Total	%
>0.20	0	3	3	0.6
>0.10 and ≤0.2	8	33	41	8.2
>0.05 and ≤0.1	24	118	142	28.4
≤0.05	68	246	314	62.8
Total	100	400	500	100.0

3.3. Changes in co-authorship patterns

For any university i, let S_i, CF_i, and CM_i be the number of solo articles, and the number of co-authored articles in the fractional and the multiplicative case, so that $TF_i = S_i + CF_i$, and $TM_i = S_i + CM_i$ are the total number of articles under the two approaches in columns 2 and 4 in Table A in the SMS. In spite of the fact that the total number of articles in the LR universities in the multiplicative case is 2.3 times greater than this number in the fractional case, the Pearson correlation coefficient between the two measures of university size is 0.98.

In the above notation, $CRF_i = CF_i/TF_i$ and $CRM_i = CM_i/TM_i$ are the co-authorship rates under the two approaches in columns 3 and 5 in Table A in the SMS. The Pearson correlation coefficient between the two variables is 0.97. However, this does not preclude that the move from the fractional to the multiplicative approach generates important differences in university co-authorship rates. We will consider two aspects of the changes induced by this move. Firstly, we study the re-rankings when universities are ordered according to the two co-authorship rates, CRF_i and CRM_i. The results are in Table 3A. Secondly, we study the cardinal differences between these rates, namely, the variable $\Delta CR_i = CRM_i - CRF_i$. The results are in Table 3B.

Differences in co-authorship rates are generally small: 314 universities, or 62.8% of the total, experience differences smaller than or equal to 0.05 percentage points. However, independently of the magnitude being evaluated, in the presence of a long list of research units small cardinal differences may have strong re-ranking effects. This is exactly what we observe in Table 3A. In the move from the fractional to the multiplicative approach, only 42 universities, or 8.4% of the total, experience small re-rankings of less than or equal to five positions. Most universities experience intermediate re-rankings of between 6 and 25 positions (51.0%), or large re-rankings of greater than 26 positions (40.6%). The last two quantities are even larger for the 100 universities with the largest co-authorship rates. In brief, small differences in co-authorship rates generate relatively large re-rankings.

4. The citation impact consequences of adopting the two counting methods

4.1. Citation impact indicators in the all-sciences case

Changes in co-authorship rates are both expected and worth monitoring in the move from the fractional to the multiplicative counting method. However, a complete evaluation of this move requires studying its effect on citation impact. In this Section, we consider the citation performance of the LR universities in what we call the all-sciences case. We use two citation impact indicators.

Firstly, the *MNCS* for university i, M_i, defined as

$$M_i = \left(\frac{1}{N_i}\right) \sum_j \sum_k \frac{c_{ijk}}{\mu_j},$$

where $N_i = \sum_j N_{ij}$ is the total number of articles in university i, and c_{ijk} is the number of citations received by article k in field j in university i. Of course, for every i, M_i is the weighted average of the RCR_{ij} for every field j introduced in Section 2.1:

$$\sum_j \left(\frac{N_{ij}}{N_i}\right) RCR_{ij} = \sum_j \left(\frac{N_{ij}}{N_i}\right) \left(\frac{\mu_{ij}}{\mu_j}\right) = \left(\frac{1}{N_i}\right) \sum_j \sum_k \frac{c_{ijk}}{\mu_j} = M_i.$$

Table 4A
University ranking differences according to the *MNCS* indicator in the move from the fractional to the multiplicative approach.

	First 100 universities	Remaining 400 universities	Total	%
>50 positions	4	61	65	13.0
26–50	9	68	77	15.4
16–25	28	141	169	33.8
6–15	52	96	148	29.6
≤5 positions	7	12	19	3.8
No change	0	22	22	4.4
Total	100	400	500	100
Median	9			
Mean	14.5			
SD	15.6			
CV	1.08			
Max	101			

The M_i values for the 500 universities according to the fractional and the multiplicative counting methods, denoted by $M^F{}_i$ and $M^M{}_i$, respectively, are in columns 1 and 2 in Table B in the SMS, where universities are ordered according to the $M^F{}_i$ values.

Secondly, the *Top 10%* for university i, T_i, defined as

$$T_i = \sum_j \frac{n_{ij}}{N_i}.$$

For[7] every i, T_i is equal to the weighted average of the Top 10% indicators T_{ij} for every field j introduced in Section 2.1, that is,

$$\sum_j \left(\frac{N_{ij}}{N_i} T_{ij}\right) = \sum_j \left(\frac{N_{ij}}{N_i}\right)\left(\frac{n_{ij}}{N_{ij}}\right) = \sum_j \left(\frac{n_{ij}}{N_i}\right) = T_i.$$

The T_i values for the 500 universities according to the fractional and the multiplicative counting methods, denoted by $T^F{}_i$ and $T^M{}_i$, respectively, are in columns 4 and 6 in Table B in the SMS.[8]

4.2. The comparison of university rankings

We begin with the case in which citation impact is measured in terms of the *MNCS*. Both the Pearson and the Spearman correlation coefficients between university values are 0.99. However, high correlations between university values and ranks do not preclude important differences for individual universities. In analyzing the consequences of going from the fractional to the multiplicative approach we take two aspects into account: the re-rankings that take place in such a move (from the left-hand column to column 3 in Table B in the SMS), and the differences between the university values themselves (columns 1 and 2 in Table B). The results for both aspects are in Table 4.

Fortunately, we have a relevant instance with which to compare our results: the differences found in Table 5 in Ruiz-Castillo and Waltman (2015) in going from the university rankings according to the *MNCS* indicator using the Web of Science classification system with 236 journal subject categories, or sub-fields, and the classification system we are using in this paper with 5119 clusters.

Only 41 universities or 8.2% of the total – among which seven belong to the first 100 – experience small re-rankings of less than or equal to five positions in the move from the fractional to the multiplicative approach. These quantities are considerably smaller than in the move from the WoS classification system to our dataset. At the other extreme, 142 universities, or 28.4% of the total, experience re-rankings of greater than 25 positions, while 168 universities, or 33.6% of the total, are in this situation in the change between classification systems.

As far as cardinal changes are concerned, differences are more or less negligible: 86.6% of universities have differences in *MNCS* values smaller than or equal to 0.05 in the change of counting methods. This percentage is 73% among the first 100 universities. In the change between classification systems, these figures are smaller: 67.8% and 56%, respectively.

[7] This is the definition actually used in the Leiden Ranking itself (Waltman et al., 2012), as well as in the SCImago ranking (Bornmann et al., 2012), and in the InCites software in the Web of Science (see 'percentile in subject area' in http://incites.isiknowledge.com/common/help/h_glossary.html).

[8] We note that, as in Ruiz-Castillo and Waltman (2015), in the calculation of the university M_i and T_i values we have normalized only for field, not for publication year. This is different from the way in which *MNCS* and *Top 10%* calculations are usually performed in the CWTS Leiden Ranking and elsewhere. However, since in this paper we work with a fixed-length instead of a variable-length citation window, normalization for publication year may be considered less important.

Table 4B
University differences in *MNCS* values in the move from the fractional to the multiplicative approach.

	First 100 universities	Remaining 400 universities	Total	%
>0.20	0	0	0	0.0
>0.10 and ≤0.2	6	1	7	1.4
>0.05 and ≤0.1	21	39	60	12.0
≤0.05	73	360	433	86.6
Total	100	400	500	100
Median	0.02			
Mean	0.03			
SD	0.02			
CV	0.83			
Max	0.19			

Table 5A
University ranking differences according to the *Top 10%* indicator in the move from the fractional to the multiplicative approach.

	First 100 universities	Remaining 400 universities	Total	%
>50 positions	16	96	112	22.4
26–50	13	79	92	18.4
16–25	24	115	139	27.8
6–15	39	65	104	20.8
≤5 positions	7	7	14	2.8
No change	1	38	39	7.8
Total	100	400	500	100
Median	15			
Mean	20.0			
SD	18.0			
CV	0.90			
Max	93			

Table 5B
University differences in *Top 10%* values in the move from the fractional to the multiplicative approach.

	First 100 universities	Remaining 400 universities	Total	%
>0.20	1	1	2	0.4
>0.10 and ≤0.2	7	18	25	5.0
>0.05 and ≤0.1	18	88	106	21.2
≤0.05	74	293	367	73.4
Total	100	400	500	100
Median	0.03			
Mean	0.04			
SD	0.03			
CV	0.82			
Max	0.21			

In brief, relative to the move from the WoS classification system to our dataset, differences in *MNCS* values when moving from the fractional to the multiplicative approach are small. However, these small differences give rise to rather important re-rankings of an intermediate size: almost two thirds of universities experience re-rankings of greater than five and smaller than 26 positions, a quantity equal to 44.2% in the case of the change of classification systems. The situation is illustrated in Fig. 1.

Next, we consider the case in which citation impact is measured in terms of the *Top 10%* indicator. Again, both the Pearson and the Spearman correlation coefficients between university values are very high indeed: 0.99. However, in order to analyze the consequences of going from the fractional to the multiplicative approach for individual universities we take into account the re-rankings that take place in such a move (columns 5 and 7 in Table B in the SMS), and the differences between the university values themselves (columns 4 and 6 in Table B in the SMS). The results for both aspects are in Table 5 (to save space, Figure 2 in the SMS illustrates the situation).

The situation is very similar to what we have seen when citation impact is measured in terms of the *MNCS* indicator. On the one hand, 73.4% universities have differences in *Top 10%* values that are smaller than or equal to 0.05 (*versus* 86.6% in the previous case). On the other hand, these small cardinal differences give rise to important re-rankings: 48.6% and 40.8% universities experience intermediate or large re-rankings between 6 and 25 positions or greater than 25 positions (*versus* 63.4% and 28.4% in the previous case).

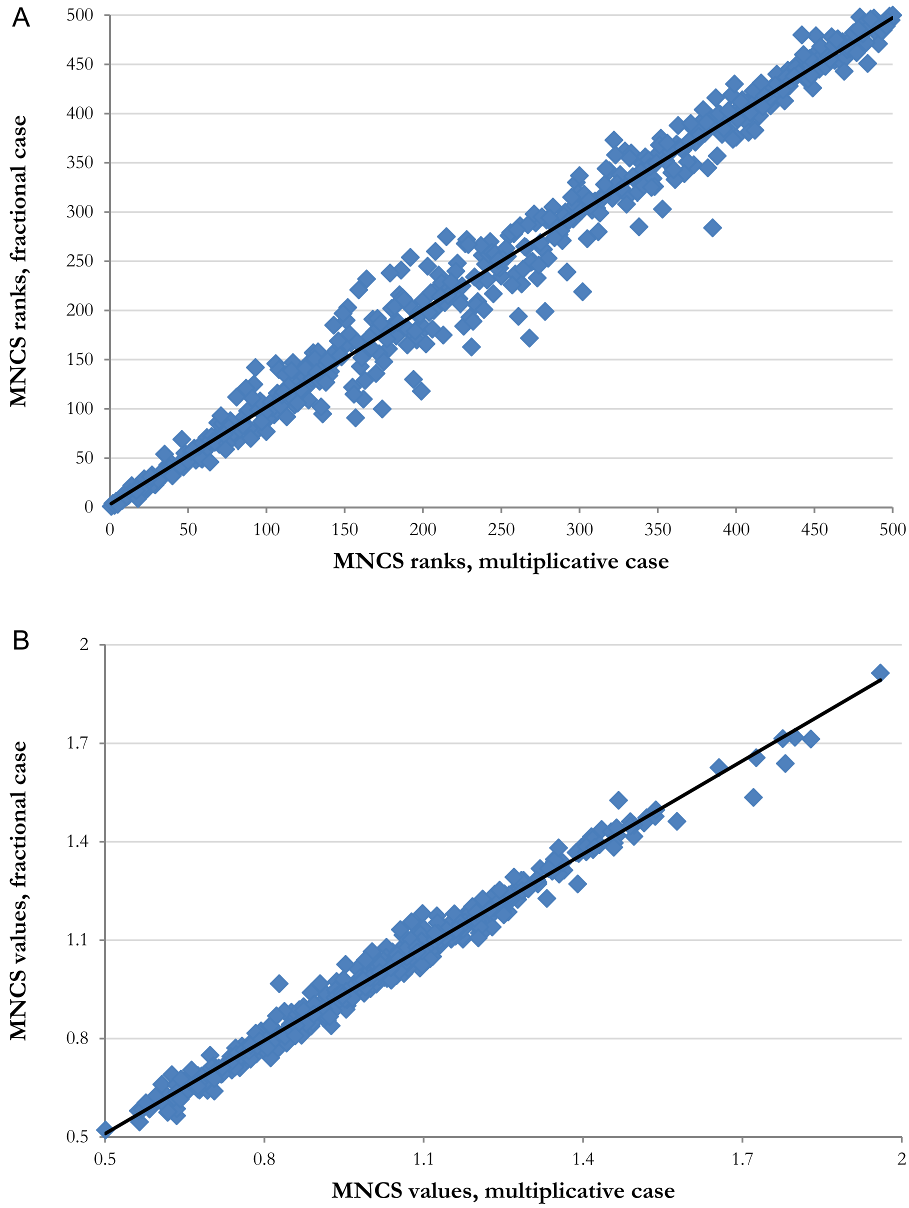

Fig. 1. (A) Scatter plot of the 500 LR universities' MNCS ranks according to the fractional and the multiplicative approaches. (B) Scatter plot of the 500 LR universities' MNCS values according to the fractional and the multiplicative approaches.

4.3. *Regression analysis*

Depending on the issue at hand, different analysts may legitimately disagree on whether the changes just analyzed are large or small. Perhaps, a majority may find these changes large enough to recommend applying both approaches in order to study the robustness of any ranking in practical applications. Be that as it may, we are interested in learning what type of university is more likely to become a gainer or a loser in the move from the fractional to the multiplicative approach in our dataset. Our attempt to answer this question in this sub-section relies on multiple regression methods.

The dependent variable is the difference in *MNCS* values, namely $\Delta M_i = M^M{}_i - M^F{}_i$, and the difference in *Top 10%* values, namely $\Delta T_i = T^M{}_i - T^F{}_i$, $i = 1, \ldots, 500$. We study the effect on the dependent variables of the following four explanatory variables.

1. The move from the fractional to the multiplicative approach typically entails increases in co-authorship rates. As we saw in Section 3.3, small differences between these rates generate considerable re-rankings when universities are ordered by the two rates CRF_i and CRM_i. Given the high correlation between the two variables, in order to study the effect of

Table 6A
Descriptive statistics.

	Sample mean	Std. deviation
ΔM_i	−0.0172	0.0321
ΔT_i	0.0127	0.0487
CRM_i	0.8174	0.0644
$(CRM_i)^2$	0.6722	0.1015
$CRM^T{}_i$	0.8422	0.0719
$(CRM^T{}_i)^2$	0.7144	0.1146
DH_i	0.3040	0.4604
DL_i	0.3120	0 .4638
DTH_i	0.3020	0 .4596
DTL_i	0.2880	0 .4533
S_i	1451.3	1022.9

co-authorship on citation impact differences, we include in the regressions the co-authorship rate in the multiplicative case, CRM_i. In so far as mean field-normalized citations steadily increase with the number of address lines (column 3 in Table 2), it is possible that the regression coefficient of CRM_i is positive.

2. On the other hand, like for other units of analysis, university citation distributions are typically highly skewed (Perianes-Rodriguez & Ruiz-Castillo, 2014). Therefore, we expect universities' citation impact – however measured – to be heavily dependent on what takes place in the upper tail of their citation distributions. As observed in column 6 in Table 2, co-authored articles have greater mean field-normalized citations than solo articles. Thus, the conjecture is that universities with a large share of highly cited co-authored articles are the ones that most benefit from a change from the fractional to the multiplicative approach. Given the high correlation between co-authorship rates among the top 10% of most cited articles (correlation coefficient equal to 0.96), we include in the analysis the rate in the multiplicative case, $CRM^T{}_i$.
3. Consider, for example, the case in which citation impact is measured in terms of the *MNCS*. An interesting question is whether the best (worse) universities according, for example, to $M^F{}_i$, are the most benefited by the move from the fractional to the multiplicative approach. To study this question, we will discretize $M^F{}_i$ by defining two dummy variables identifying high and low ranked universities. After some experimentation, we find it useful to define the following two variables:

$$\begin{aligned} DH_i = \; & 1 \quad \text{if}\, M_i^F \geq 1.11 \\ & 0 \quad \text{otherwise;} \\ DL_i = \; & 1 \quad \text{if}\, M_i^F < 0.90 \\ & 0 \quad \text{otherwise.} \end{aligned}$$

In this case, there are 150 and 157 universities with $DH_i = 1$ and $DL_i = 1$, respectively. The remaining 193 universities with $DH_i = 0$ and $DL_i = 0$ have intermediate $M^F{}_i$ values. Similarly, when citation impact is measured in terms of the Top 10% indicator we define

$$\begin{aligned} DH_i = \; & 1 \quad \text{if}\, M_i^F \geq 1.15 \\ & 0 \quad \text{otherwise;} \\ DL_i = \; & 1 \quad \text{if}\, M_i^F < 0.81 \\ & 0 \quad \text{otherwise.} \end{aligned}$$

In this case, there are 147 universities in the best and worse groups, whereas the remaining 206 universities have intermediate $M^F{}_i$ values.
4. Finally, we would like to investigate whether large or small universities benefit the most from the move from the fractional to the multiplicative approach. Since the total number of articles depends very much on the counting method used, an alternative is to focus on the number of solo articles S_i, which is a variable of interest in its own right, and whose correlation coefficients with TF_i and TM_i are 0.95 and 0.86, respectively.

In order to test for non-linearities, we include a pair of quadratic terms $(CRM_i)^2$ and $(CRM^T{}_i)^2$. The final regressions are:

$$\Delta Mi = \alpha + \beta_1 CRM_i + \beta_2 (CRM_i)^2 + \beta_3 CRM_i^T + \beta_4 (CRM_i^T)^2 + \beta_5 DH_i + \beta_6 DL_i + \beta_7 S_i,$$
$$\Delta T_i = \alpha' + \beta'_1 CRM_i + \beta'_2 (CRM_i)^2 + \beta'_3 CRM_i^T + \beta'_4 (CRM_i^T)^2 + \beta'_5 DH_i + \beta'_6 DL_i + \beta'_7 S_i,$$

Descriptive statistics, and regression results are presented in Table 6. They warrant the following four comments.

Table 6B
Regression results.

Expl. variables	Dependent variable = ΔM_i			Dependent variable = ΔT_i		
	Coefficient	Std. error	t-value	Coefficient	Std. error	t-value
CRM_i	0.9868	0.4607	2.1*	1.5170	0.6383	2.4*
$(CRM_i)^2$	−1.0246	0.2906	−3.5*	−1.6328	0.4026	−4.1*
$CRM^T{}_i$	−0.5624	0.3499	−1.6	−1.1176	0.4844	−2.3*
$(CRM^T{}_i)^2$	0.7173	0.2196	3.3*	1.3684	0.304	4.5*
DH_i	−0.0154	0.0029	−5.3*	−0.0214	0.004	−5.4*
DL_i	0.0091	0.0027	3.3*	0.009	0.0038	2.4*
S_i	−1.53E−06	1.20E−06	−1.3	−4.21E−06	1.70E−06	−2.5*
Constant	−0.1699	0.0999	−1.7	−0.1562	0.1377	−1.1
N	500			500		
Adjusted R^2	0.392			0.492		

* Significant regression coefficients.

Table 6C
Marginal effects on ΔM_i and ΔT_i caused by the variables CRM_i and $CRM^T{}_i$.

CRM_i	$CRM^T{}_i$
$\partial\Delta M_i/\partial CRM_i = \beta_1 + 2\beta_2 CRM_i = -0.6882$	$\partial\Delta M_i/\partial CRM^T{}_i = \beta_3 + 2\beta_4 CRM^T{}_i = 0.6458$
$\partial\Delta T_i/\partial CRM_i = \beta'_1 + 2\beta'_2 CRM_i = -1.0993$	$\partial\Delta T_i/\partial CRM^T{}_i = \beta'_3 + 2\beta'_4 CRM^T{}_i = 1.1873$

1. It is observed that 12 out of 14 regression coefficients for the seven explanatory variables are statistically significant. Furthermore, the adjusted R^2 coefficients for the two regressions, 0.39 and 0.49, indicate that the goodness of fit for the two models is acceptable.
2. The marginal effects of the variables CRM_i and $CRM^T{}_i$, evaluated at the corresponding sample means, are presented in Table 6C. The results are very interesting. On the one hand, the co-authorship rate CRM_i has a negative effect on both dependent variables. This means that the move from the fractional to the multiplicative approach penalizes universities with a high co-authorship rate for the distribution as a whole. On the contrary, this move benefits universities with a high co-authorship rate $CRM^T{}_i$ in the upper tail of the citation distribution.
3. Universities have been classified into three groups according to their *MNCS* and *Top 10%* values. Our results clearly indicate that the worse the citation performance of a university is, the greater is the benefit for this university of a move from the fractional to the multiplicative counting method.[9] Of course, it could be argued that partitioning the set of universities into these specific three groups is a useful but arbitrary procedure. This is particularly the case in a situation in which universities' citation impact values are very close to each other, so that it is difficult to assert, for example, that one university is among the best and the next one in the ranking is among the intermediate ones. Fortunately, it is possible to study the appropriateness of the above definitions in a sensitivity analysis that accentuates the differences between the three groups by eliminating a number of intermediate universities which are close to the best ones, as well as a number of intermediate universities which are close to the worst ones.[10] To save space, the regression results are presented in Table C in the SMS. It suffices to say that, except for their smaller statistical significance in the *Top 10%* case, the regression coefficients for all variables and, specifically for the variables DH_i and DL_i, for the remaining 424 and 410 universities are very close to what we obtained for the 500 universities. This establishes the robustness of the effect of the university citation impact on both dependent variables.
4. The S_i variable has a negative regression coefficient in both regressions, but this coefficient is only significant in the *Top 10%* case. This indicates that the greater the number of solo articles is, the smaller is the probability that a university has a greater citation impact in the multiplicative case. It should be noted that, judging from the size of regression coefficients, this effect is small.

5. Summary and discussion

The attribution of responsibility for co-authored publications poses a severe evaluation problem at all levels of analysis: authors; organizations, such as research groups, university departments, or the corresponding divisions in research institutes, and geographical areas, such as regions, countries, or wider aggregates such as the European Union. In this paper, we have

[9] Recall that in the example introduced in Section 2.3, the relative ranking of unit B according to the *Top 50%* indicator improves when moving from the fractional to the multiplicative approach. Although in the regression analysis we use the *MNCS* and the *Top 10%* indicator, the re-ranking in the example is consistent with the finding that, *ceteris paribus*, the worse the citation performance of a research unit is, the greater is the benefit for this unit of a move from the fractional to the multiplicative counting method.

[10] In the *MNCS* case, we eliminate 38 universities with $M^F{}_i$ in the interval [1.07, 1.11), and another 38 with $M^F{}_i$ in the interval [0.90, 0.95), whereas in the *Top 10%* case, we eliminate 46 universities with $M^F{}_i$ in the interval [1.08, 1.15), and 44 with $M^F{}_i$ in the interval [0.81, 0.89).

restricted the attention to three methods that require the minimum of information on the co-authorship phenomenon, namely, the number of addresses in the byline of scientific publications.

At a conceptual level, this paper has clarified that there is a difficulty with the full counting method that is quite independent of the field-normalization issue. The problem appears, even in a single scientific field, as long as we use citation impact indicators that are additively decomposable. This is important, since additive decomposability is a desirable property in many practical contexts. This problem adds to the critique in WVE, which is based on the argument that properly field-normalized results cannot be obtained with full counting. Consequently, full counting creates an unfair advantage to fields with a lot of co-authorship and with a strong correlation between co-authorship and citations. As shown in WVE, this may lead to seriously biased results for the evaluation of research units at high aggregate levels such as universities or countries.

In our opinion, these results justify abandoning the use of full counting at high aggregation levels. However, we need not abandon the main idea behind this approach. Co-authored publications can be fully assigned to each co-author as long as the overall citation distribution in each field is made equal to the union of the extended citation distributions of the research units working in the field in question, as it is done in the multiplicative counting method. The multiplicative approach is compatible with the use of additively decomposable indicators, and yields properly field-normalized results. Therefore, in this paper we focus on the two valid alternatives that are readily applicable with a minimum of information: the fractional and the multiplicative counting methods.

A preliminary question should be clarified at the outset. It is obvious that, relative to solo publications, the adoption of the multiplicative approach inflates the scientific impact of publications with multiple authors. But adopting the fractional approach diminishes the scientific impact of such publications. In WVE's words "*the disadvantage of multiplicative counting is that publications do not have the same weight in the calculation of field-normalized indicators*" (p. 42). Others will claim that the disadvantage of fractional counting is that it penalizes co-authored publications in field-normalized calculations. *A priori*, we do not find reasons to prefer one bias to the other before or after field-normalization.

Using a large WoS dataset, the rest of the paper has investigated the empirical consequences of adopting the two approaches in the particular case of 500 LR universities. Among other possible alternatives, the available data only allows us to use the fractional and multiplicative counting methods of the address lines variety. Nevertheless, it is hoped that a better understanding of changes in co-authorship and citation impact patterns, as well as the type of universities most affected by a change in counting methods might prove helpful for practitioners who must choose between the two alternative methodologies.

Of course, co-authorship and citation impact patterns are changed when we move from a fractional to a multiplicative approach (and *vice versa*). The question is whether these changes are large or small. Our first finding is that cardinal differences between co-authorship rates, *MNCS* values, and *Top 10%* values do not cause dramatic differences in co-authorship and citation impact pattern. However, these small cardinal differences generate considerable re-rankings between universities.

Nevertheless, the choice between the two approaches may well depend on which universities end up being gainers or losers in this move. Our second finding is that, *ceteris paribus*, the gainers with a move from the fractional to the multiplicative approach are characterized by (i) a low co-authorship rate for citation distributions as a whole, but a high co-authorship rate in the upper tail of these distributions, (ii) a low citation impact performance, and (iii) a small number of solo articles. Do we want to benefit or to penalize universities with these characteristics? In the former case, we should use the multiplicative approach, whereas in the latter case we should use the fractional approach. On our part, we do not have a clear preference in this respect.

Of course, it would be very convenient to extend the methods of analysis used in this paper to other datasets, other types of research units, as well as other variants of counting methods. However, if forced to choose between the two approaches at this point, then we do prefer the multiplicative alternative on the following grounds. As pointed out eloquently by WVE, at a low level of aggregation multiplicative counting is in agreement with the intuitive idea that "*all publications of a researcher or a research group should be considered of equal importance*" (p. 41). At a high aggregate level, such as countries or organizations, WVE consider absolutely essential to use fractional counting because, as they exhibit in their examples, "*At this level, there is a serious risk of misinterpretation of full counting results*" (p. 40). However, as we have established in this paper, no such risk affects in the least the multiplicative approach. Thus, in our opinion, all publications of a university or a country may be considered of equal importance regardless of the number of co-authors.

Naturally, others may think differently. Therefore, in practical applications at every aggregation level it seems sensible to study the robustness of the results achieved with both approaches. For example, this is exactly what we do when studying the entire university citation distributions using this same dataset (Perianes-Rodriguez & Ruiz-Castillo, 2014). Interestingly enough, we find that the key result concerning the high skewness and the strong similarity between university citation distributions is independent of the counting method used.

Before finishing, we wish to make three remarks. The first comment is that having additional information concerning the "true" responsibility of each unit in co-authored publications would not necessarily solve the problem we have faced in this paper. Consider the possibility that all journals demand from the authors of each publication a statement indicating who did what in the manner actually done in *PLoS ONE*. Assume, for example, that we have information on who had the idea, who did the work, who did the analysis, and who wrote each paper. Assume that, on the basis of that information, it can be established that one of two authors is responsible for 2/3 of the article. Under a fractional approach the solution is to assign 2/3 to one author and 1/3 to the other. But this is only if we decide to treat each co-authored publication, independently of the number of authors, as equal to one solo publication. Given the recent decrease in the percentage of solo articles or other reasons, we may accept counting differently solo and co-authored publications. In this case, from a multiplicative point of view it would be possible to count fully two publications, assigning 4/3 to one author, and 2/3 to the other, or to assign one publication to the first author and only 1/2 to the second author. The conclusion is that we must address two issues: how to assign the responsibility of a co-authored publication to its authors, and how to establish the relationship between one solo article and a co-authored publication with a given number of authors. Beyond a pure evaluation perspective, the second issue is linked to policy considerations.

The second remark arises also from the distinction between the evaluation of scientific publications' citation impact and the research policy perspective as practiced, for example, in the European Union (EU hereafter), with programs clearly favoring co-authorship between nationals from different EU countries. Given the relation between number of co-authors and mean citations exhibited in Table 2, this policy might be justified for the incentives it provides for achieving a greater citation impact. Another possible justification is the desire to strengthen the EU cohesion by providing incentives for collaborating across EU countries. Be that as it may, in so far as this policy favors co-authorship, it is consistent with an evaluation strategy that uses the multiplicative counting method.

Finally, given the strong limitations of the fractional and the multiplicative counting methods analyzed in this paper, we must conclude that the issue of assigning responsibility in a fair way to co-authored publications is far from being closed.

Acknowledgements

This research project builds on earlier work started by Javier Ruiz-Castillo during a research visit to the Centre for Science and Technology Studies (CWTS) of Leiden University. The authors gratefully acknowledge CWTS for the use of its data. Ruiz-Castillo acknowledges financial support from the Spanish MEC through grant ECO2014-55953-P. Conversations with Ricardo Mora, and Ludo Waltman are deeply appreciated. Two referee reports greatly contributed to an improvement of the original version of the paper. All remaining shortcomings are the authors' sole responsibility.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.joi.2015.10.002.

References

Abramo, G., D'Angelo, C. A., & Rosati, F. (2013). Measuring institutional research productivity for the life sciences: The importance of accounting for the order of authors in the byline. *Scientometrics*, *97*, 779–795.

Albarrán, P., Crespo, J., Ortuño, I., & Ruiz-Castillo, J. (2010). A comparison of the scientific performance of the U.S. and Europe at the turn of the 21st century. *Scientometrics*, *85*, 329–344.

Albarrán, P., Ortuño, I., & Ruiz-Castillo, J. (2011). The measurement of low- and high-impact in citation distributions: Technical results. *Journal of Informetrics*, *5*, 48–63.

Anderson, J., Collins, P., Irvine, J., Isard, P., Martin, B., Narin, F., et al. (1988). On-line approaches to measuring national scientific output: A cautionary tale. *Science and Public Policy*, *15*, 153–161.

Bornmann, L., De Moya Anegón, F., & Leydesdorff, L. (2012). The new excellence indicator in the world report of the SCImago Institutions Rankings 2011. *Journal of Informetrics*, *6*, 333–335.

Foster, J. E., Greeer, J., & Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica*, *52*, 761–766.

Foster, J. E., & Shorrocks, A. (1991). Subgroup consistent poverty indices. *Econometrica*, *59*, 687–709.

Hagen, N. T. (2008). Harmonic allocation of authorship credit: Source level correction of bibliometric bias assures accurate publication and citation analysis. *PLoS ONE*, *3*, e4021.

Li, Y., Castellano, C., Radicchi, F., & Ruiz-Castillo, J. (2013). Quantitative evaluation of alternative field normalization procedures. *Journal of Informetrics*, *7*, 746–755.

Moya-Anegón, F., Guerrero-Bote, V., Bornmann, L., & Moed, H. (2013). The research guarantors of scientific papers and the output counting: A promising new approach. *Scientometrics*, *97*, 421–434.

Perianes-Rodriguez, A., & Ruiz-Castillo, J. (2014). *University Citation Distributions. Forthcoming in Journal of the American Society for Information Science and Technology. Available as Working Paper 14–26, Departamento de Economía, Universidad Carlos III (http://hdl.handle.net/10016/19811).*

Ruiz-Castillo, J., & Waltman, L. (2015). Field-normalized citation impact indicators using algorithmically constructed classification systems of science. *Journal of Informetrics*, *9*, 102–117.

Sen, A. (1992). *Inequality reexamined*. Cambridge: Harvard University Press.

Shen, H.-W., & Barabási, A.-L. (2014). Collective credit allocation in science. *Proceedings of the National Academy of Sciences*, *111*, 12325–12330.

Stallings, J., Vance, E., Yang, J., Liang, J., Pang, L., Dai, L., et al. (2013). Determining scientific impact using a collaboration index. *Proceedings of the National Academy of Sciences*, *110*, 9680–9685.

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E. C. M., Tijssen, R. J. W., Van Eck, N. J., et al. (2012). The Leiden Ranking 2011/2012: Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology*, *63*, 2419–2432.

Waltman, L., & Van Eck, N. J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology*, *63*, 2378–2392.

Waltman, L., & Van Eck, N. J. (2015). Field normalized citation impact indicators and the choice of an appropriate counting method. *Journal of Informetrics*, *9*, 872–894.

Zhang, C. T. (2009). A proposal for calculating weighted citations based on author rank. *EMBO Reports*, *10*, 416–417.