Journal of Informetrics 9 (2015) 667–673

Contents lists available at ScienceDirect

Journal of Informetrics

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

Rethinking the comparison of coauthorship credit allocation schemes

Jinseok Kim[a,*], Jinmo Kim[b]

[a] Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign, 501 E. Daniel St., Champaign, IL 61820, USA
[b] Department of Bank Examination, Korea Financial Supervisory Service, 38 Yeoui-daero, Youngdeungpo-gu, Seoul 150-743, South Korea

ARTICLE INFO

Article history:
Received 6 March 2015
Received in revised form 13 July 2015
Accepted 13 July 2015

Keywords:
Coauthor order
Coauthorship credit
Author contribution
Bibliometrics
Empirical validation

ABSTRACT

This paper compares Fractional, Geometric, Arithmetic, Harmonic, and Network-Based schemes for allocating coauthorship credits. Each scheme is operationalized to be flexible in producing credit distribution by changing parameters, and to incorporate a special situation where the first and corresponding authors are assigned equal credits. For testing each scheme, empirical datasets from economics, marketing, psychology, chemistry, and medicine, were collected and errors in how each scheme approximates empirical data was measured. Results show that Harmonic scheme performs best overall, contrary to some claims of preceding studies in support of Harmonic or Network-Based models. The performance of a scheme, however, seems to heavily depend on empirical datasets and flexibility of the scheme, not on its innate feature. This study suggests that the comparison of coauthorship credit allocation schemes should be taken with care.

1. Introduction and background

Scholarly publication is not merely a vehicle of communicating research but also an indicator of a scholar's performance. Thus, the evaluation of a scholar for hiring, promotion, awards, or funding tends to involve assessment of her/his publication history (Beasley & Wright, 2003; Thomas et al., 2004). As the dependency of scholars on coauthoring for knowledge production has increased (Wuchty, Jones, & Uzzi, 2007) and each coauthor's contribution to a publication is not regarded to be equal in the majority of scientific fields (Waltman, 2012), the method of allocating authorship credit to coauthors in a paper has attracted a burgeoning interest from bibliometrics scholars. For a comprehensive review on this topic, refer to Marušić, Bošnjak, and Jerončić (2011).

Against the traditional practice of assigning a full authorship credit or an equal fraction to coauthors in a paper (Lindsey, 1980; Oppenheim, 1998; Price, 1981), some scholars have proposed unequal coauthorship credit allocation schemes. Their main assumption is that each author's contribution can be decoded from publications (Frandsen & Nicolaisen, 2010). For this, the byline order is frequently explored: e.g., (1) the first author is assumed to contribute more than the second author, and so on, or (2) the last-positioned author is regarded as a senior or head of the research (He, Ding, & Yan, 2012; Jian & Xiaoli, 2013; Wren et al., 2007). Others turn to the corresponding author information or a footnote stating that, for instance, two authors contribute equally (Hu, 2009; Jian & Xiaoli, 2013; Mattsson, Sundberg, & Laget, 2011).

* Corresponding author. Tel.: +1 217 751 2943; fax: +1 217 244 3302.
E-mail addresses: jkim362@illinois.edu (J. Kim), jinmo@fss.or.kr (J. Kim).

http://dx.doi.org/10.1016/j.joi.2015.07.005
1751-1577/

Table 1
Formulae of fixed and flexible types of coauthorship credit allocation schemes (N = number of authors per paper, R = order of an author in contribution, d = distribution parameter).

Schemes	Fixed type	Flexible type
Fractional	$1/N$	–
Geometric	$2^{N-R}/2^N-1$	$1/N(1-d)+2^{N-R}/2^N-1d$
Arithmetic	$2(N-R+1)/N(N+1)$	$1/N(1-d)+2(N-R+1)/N(N+1)d$
Harmonic	$1/R/\sum_{i=1}^{N}1/i$	$1/N(1-d)+1/R/\sum_{i=1}^{N}1/id$
Network-Based	–	$1/N+1/Nd\sum_{i=1}^{N-R}1/(N-i) \quad (R=1)$ $1/N(1-d)+1/Nd\sum_{i=1}^{N-R}1/(N-i) \quad (1<R<N)$ $1/N(1-d) \quad (R=N)$

Despite their methodological variations, the schemes of unequal coauthorship credit allocation can be categorized into fixed and flexible ones. Some scholars propose a fixed set of credit allocation that is only adjusted to the number of authors per paper (e.g., Egghe, Rousseau, & Van Hooydonk, 2000; Van Hooydonk, 1997). According to this method, the coauthorship credit an author receives is decided based on a strict formula regardless of scientific fields. Others allow their schemes to produce diverse sets of credits by changing parameters that can be decided by researchers or evaluators (e.g., Abramo, D'Angelo, & Rosati, 2013; Kim & Diesner, 2014; Liu & Fang, 2012; Trueba & Guerrero, 2004). These flexible schemes are often proposed to incorporate different credit assignment cultures across fields or specific evaluation needs, e.g., life science researchers in Italian universities (Abramo et al., 2013).

Interestingly, several scholars have recently argued that specific schemes approximate real-world credit allocations better than others. For example, Hagen (2010, 2013) tested how closely various allocation schemes can approximate coauthorship credit allocations that are obtained from surveys in chemistry, biomedicine, and psychology, and found that Harmonic scheme shows the best performance. In contrast, Kim and Diesner (2014) proposed the Network-Based model and argued that their scheme performs mostly better than others including Harmonic scheme.

Although each study has claimed that their scheme, when tested against empirical datasets, is superior to others in terms of producing approximate allocation, we argue such a claim has come with biased selection of data and operationalization of schemes. Therefore, this paper aims to compare the performance of several commonly discussed coauthorship credit allocation schemes under the same conditions. Specifically, such a motivation comes from our observations as follows. (1) In Kim and Diesner (2014), the proposed Network-Based scheme, which is a flexible type, was compared to other fixed-type schemes. As the authors mentioned in their paper (Kim & Diesner, 2014, p. 600), the good performance of their model may be due to its flexibility to produce various sets of scores to fit to empirical data. (2) In Hagen (2010, 2013), the scheme modification for equal contribution of the first and last authors was applied only to the proposed Harmonic scheme, not to other comparative schemes. (3) Moreover, in aforementioned studies, the empirical datasets were not identical, although some were shared.

2. Method

This paper attempts to validate the superiority of a scheme over others as suggested in Hagen (2010, 2013) and Kim and Diesner (2014). For this, we explain how this study reduces biases in scheme operationalization and data acquisition in previous studies to conduct a more robust performance test under the same conditions.

2.1. Selection of test schemes

This paper tests five coauthorship credit allocation schemes that were tested both in Hagen (2010) and Kim and Diesner (2014). Characteristics of each scheme are thoroughly discussed in those studies. For more detailed information, we refer readers to them[1]. Hereafter, a paper is assumed to have a unit value of one and authors in a byline are ordered in a decreasing order of contribution. Fractional counting and its variations (Carbone, 2011; Lindsey, 1980; Oppenheim, 1998; Price, 1981) are commonly used in practice and often serve as a baseline to compare the performance of a newly proposed scheme. Moreover, it was included both in Hagen (2010) and Kim and Diesner (2014) for model performance tests. Thus, we include Fractional scheme in our study although it is not a scheme of unequal allocation. Allocation formulae of Fractional, Geometric (Egghe et al., 2000), Arithmetic (Van Hooydonk, 1997), Harmonic (Hagen, 2010), and Network-Based (Kim & Diesner, 2014) schemes are shown in Table 1. Each formula describes how much credit an R-th author in an N-authored paper receives according to the scheme. Please note that Network-Based scheme has three different calculation methods for the first, last, and middle authors, respectively.

[1] Drawbacks of some models are worth noting from Hagen (2010). Fractional scheme does not consider the reality of relative contribution of coauthors, thus inflating the contribution of secondary authors. According to Geometric scheme, a few authors high in rank in bylines get most of credits, while others receive negligible credits. Arithmetic model fails to provide an allocation with fixed proportion if the number of authors is larger than two.

Table 2
Coauthorship credit allocation per parameter value for geometric scheme (N = number of authors per paper, R = order of an author in contribution, d = distribution parameter with the increment of 0.1 between zero and one).

N	2		3			4			
R	1	2	1	2	3	1	2	3	4
d = 0.0	0.5000	0.5000	0.3333	0.3333	0.3333	0.2500	0.2500	0.2500	0.2500
d = 0.1	0.5167	0.4833	0.3571	0.3286	0.3143	0.2783	0.2517	0.2383	0.2317
d = 0.2	0.5333	0.4667	0.3810	0.3238	0.2952	0.3067	0.2533	0.2267	0.2133
d = 0.3	0.5500	0.4500	0.4048	0.3190	0.2762	0.3350	0.2550	0.2150	0.1950
d = 0.4	0.5667	0.4333	0.4286	0.3143	0.2571	0.3633	0.2567	0.2033	0.1767
d = 0.5	0.5833	0.4167	0.4524	0.3095	0.2381	0.3917	0.2583	0.1917	0.1583
d = 0.6	0.6000	0.4000	0.4762	0.3048	0.2190	0.4200	0.2600	0.1800	0.1400
d = 0.7	0.6167	0.3833	0.5000	0.3000	0.2000	0.4483	0.2617	0.1683	0.1217
d = 0.8	0.6333	0.3667	0.5238	0.2952	0.1810	0.4767	0.2633	0.1567	0.1033
d = 0.9	0.6500	0.3500	0.5476	0.2905	0.1619	0.5050	0.2650	0.1450	0.0850
d = 1.0	0.6667	0.3333	0.5714	0.2857	0.1429	0.5333	0.2667	0.1333	0.0667

For the purpose of comparison, this study generates flexible schemes for Geometric, Arithmetic, and Harmonic schemes. The motivation for this operationalization is to make these models to be comparable to the Network-Based model. When Kim and Diesner (2014) showed that Network-Based scheme performed better than other schemes in approximating empirical data, only their model was allowed to provide its best performing score sets by adjusting a distribution parameter (as a flexible type of scheme), while other schemes provided single sets of scores (as fixed types). A concern is that the good performance of Network-Based model might be due largely to its flexibility: if competing models are flexible to produce their own best fitting score sets against test data in the same way as Network-Based model is, they may outperform it.

The basic idea of Network-Based model is that coauthors keep an equal amount of partial credit to themselves and distribute the rest to others unequally based on their position in a paper's byline (for more detail, see Kim & Diesner, 2014, p.590). Here, the amount of distributable credit per author is decided by a distribution parameter, which is set by a model user. For example, if each coauthor in a three-authored paper gets 1/3 of a unit credit and is set by a model user to distribute a half of their shares (here, the distribution parameter is 0.5 or 50%), they will distribute 1/6 of the unit credit ($1/3 \times 0.5$) to other coauthors. The same operationalization is applied to other schemes so that (1) a partial credit is allocated to each author in an equal amount regardless of byline position and (2) the remaining share will be unequally distributed to other coauthors according to a distribution parameter (see Table 1 Flexible Type)[2].

Table 2 illustrates the possible credit allocation generated by the flexible type of Geometric scheme. Here, a distribution parameter d decides the ratio of credit share of each author that gets allocated to other coauthors. The parameter can be any real number between zero and one. We can see that, when d is set to be zero, the scheme equals the Fractional scheme. In contrast, if d is set to be one, the allocation has the same distribution as that of the fixed type of Geometric scheme. As the parameter increases, the allocation gets more unequal as the difference of credit assigned to each author also increases. These features all apply to other flexible-type schemes.

2.2. Handling equal contributors

In coauthorship credit allocation studies, the first author is usually assumed to contribute most to a paper and the order of authors in a byline represents their relative contributions in a decreasing order (Marušić et al., 2011). Sometimes, however, two authors may be indicated to contribute equally as joint first authors. This convention was found to increase in some fields (Hu, 2009). Another problematic situation is when the corresponding author is not the first author. Despite a controversy over the contribution of corresponding authors, many of previous coauthorship credit studies have assumed that the first and corresponding authors contribute equally (e.g., Hagen, 2010, 2013; Jian & Xiaoli, 2013; Kim & Diesner, 2014). Thus, this second problem becomes the same as the first one where the issue is how to operationalize equal contributors into an unequal allocation scheme.

The common strategy of aforementioned comparative studies is to get an average of the first and second largest credits and assign the averaged credit to each of joint first authors or to each of the first and corresponding authors. This method is favored because it does not affect the credits allocated to other remaining coauthors (Hagen, 2010; Kim & Diesner, 2014). An interesting observation is, however, that those studies applied such an operationalization only to their proposed models but not to other comparing ones. The problem is that the fitting performance of Harmonic and Network models with this operationalization was used as evidence of their superiority over other models. This paper applies the operationalization

[2] A reviewer raised a concern about overfitting of a credit allocation scheme against test data by using a parameter. Overfitting can be an issue if a model user attempts to find a parameter value best fitting a flexible allocation scheme against an empirical dataset and then use the value to fit the scheme to other datasets. In our study, however, this may not be a problem because each flexible type of scheme is supposed to find multiple distribution parameters fitting best (with a lowest Lack of Fit) against each of test datasets, not across all datasets.

Table 3
Coauthorship credit scores from empirical data (N = number of authors per paper, R = order of an author in contribution).

		Maciejovsky et al. (2008)			Vinkler (1993, 2000)		Wren et al. (2007)	
N	R	Economics	Marketing	Psychology	Chemistry (1993)	Chemistry (2000)	Medicine	
2	1	0.6224	0.5864	0.6054	0.71	0.65		
	2	0.3776	0.4136	0.3946	0.29	0.35		
3	1	0.4429	0.4557	0.4915	0.61	0.55	0.42	
	2	0.3253	0.3116	0.2924	0.26	0.25	0.41	
	3	0.2318	0.2328	0.2160	0.13	0.20	0.17	
4	1	0.3635	0.3493	0.4230	0.54	0.50		
	2	0.2586	0.2832	0.2411	0.31	0.25		
	3	0.2245	0.2302	0.1911	0.09	0.15		
	4	0.1534	0.1373	0.1448	0.06	0.10		
5	1				0.34	0.40	0.38	0.34
	2				0.24	0.25	0.34	0.29
	3				0.17	0.15	0.12	0.16
	4				0.14	0.10	0.08	0.12
	5				0.11	0.10	0.07	0.08
6	1					0.35		
	2					0.25		
	3					0.15		
	4					0.10		
	5					0.10		
	6					0.05		

of equal contributors to all comparative schemes because such an operationalization is an external decision that is not an innate feature of any specific formula.

2.3. Empirical data for validation

The same empirical datasets used in Hagen (2010, 2013) and Kim and Diesner (2014) were obtained for this study, as shown in Table 3. A total of six datasets were collected for five academic fields. Maciejovsky, Budescu, and Ariely (2008) was the data source for economics, marketing, and psychology. In this study, participants from economics (n = 45), marketing (n = 150), and psychology (n = 52) were requested to assign relative contribution (in percentage) of coauthors according to their byline position in 7234 synthetic author lists[3]. For chemistry, two sources were searched. First, Vinkler (1993, Table 5, Total Contribution Factors, p.222) included a list of credit allocation based on a survey of 68 researchers in a Hungarian chemical research institute. This dataset was used in Kim and Diesner (2014). Also, Vinkler (2000) provided credit shares based on the position and number of authors in a paper (Table 4, p. 608)[4]. These scores were used in Hagen (2010, 2013). For medicine, we re-used the dataset in Wren et al. (2007, Table 1) that reports a survey of 87 promotion committee members in U.S. medical schools. The respondents were asked to allocate authorship credit based on an author's byline position.

Although most datasets used in Hagen (2010, 2013) and Kim and Diesner (2014) came from the same sources (Maciejovsky et al., 2008; Vinkler, 1993, 2000; Wren et al., 2007), some datasets appeared in one study and did not in another. For example, the datasets for economics and marketing from Maciejovsky et al. (2008) were excluded in Hagen (2010, 2013), while they were used in Kim and Diesner (2014). From Vinkler (1993, 2000), Hagen (2010, 2013) used Table 4 in Vinkler (2000), while Kim and Diesner (2014) chose Table 5 in Vinkler (1993). From Wren et al. (2007), Hagen (2010, 2013) also excluded a dataset for a five-authored paper case where the third author is designated as a corresponding author. To fairly validate the performance of each scheme in Table 1, this study used all the available datasets. In Table 3, some credit values are rearranged in a decreasing order: for example, in Medicine, corresponding authors who appear in the last or in the middle but their contributions are the first or second largest are moved to the first or second author position accordingly.

[3] In Maciejovsky et al. (2008), two versions of empirical datasets for economics, marketing, and psychology were reported. One came from surveys of respondents in those fields on alphabetically-ordered name sets, while the other from surveys on non-alphabetically-ordered name sets. As we test contribution-based ordering of authors (Waltman, 2012), we chose the latter version (Maciejovsky et al., 2008, Fig. A2), following the choice of Hagen (2010, 2013) and Kim and Diesner (2014).

[4] It is not clear how the scores in Table 4, referred as a "scoring system" by Vinkler (2000, p. 608), were generated. According to Vinkler (2000), an algorithm was used to allocate credit shares among coauthors and "empirical (Vinkler, 1993) and theoretical (Lukovits and Vinkler, 1995) investigations give support to the use of the distribution shares applied." This seems to suggest that the score table, unlike other datasets in our study, is a synthetic data, not an empirical one.

Table 4
Performance comparison of fixed and flexible-type schemes against empirical data (LOF = lack of fit, *d* = distribution parameter with the lowest LOF).

Dataset	Scheme type	Performance rank			
		1st	2nd	3rd	4th
Economics	Fixed (LOF)	Arithmetic (0.0102)	Harmonic (0.0146)	Fractional (0.0273)	Geometric (0.0410)
	Flexible (LOF, *d*)	Arithmetic (0.0005, 0.66)	Harmonic (0.00180, 0.59)	Network (0.00184, 0.25)	Geometric (0.0019, 0.49)
Marketing	Fixed (LOF)	Arithmetic (0.0108)	Harmonic (0.0175)	Fractional (0.0253)	Geometric (0.0401)
	Flexible (LOF, *d*)	Arithmetic (0.0008, 0.64)	Geometric (0.0021, 0.47)	Harmonic (0.0030, 0.55)	Network (0.0040, 0.23)
Psychology	Fixed (LOF)	Harmonic (0.0060)	Arithmetic (0.0087)	Geometric (0.0259)	Fractional (0.0430)
	Flexible (LOF, *d*)	Harmonic (0.0003, 0.73)	Geometric (0.0004, 0.60)	Network (0.0024, 0.30)	Arithmetic (0.0037, 0.76)
Chemistry (1993)	Fixed (LOF)	Harmonic (0.0125)	Arithmetic (0.0168)	Geometric (0.0297)	Fractional (0.1010)
	Flexible (LOF, *d*)	Network (0.0110, 0.51)	Harmonic (0.0125, 1.00)	Geometric (0.0161, 0.79)	Arithmetic (0.0168, 1.00)
Chemistry (2000)	Fixed (LOF)	Harmonic (0.0026)	Arithmetic (0.0089)	Geometric (0.0282)	Fractional (0.0732)
	Flexible (LOF, *d*)	Harmonic (0.0026, 0.97)	Geometric (0.0031, 0.70)	Network (0.0060, 0.46)	Arithmetic (0.0088, 1.00)
Medicine	Fixed (LOF)	Harmonic (0.0028)	Arithmetic (0.0082)	Geometric (0.0180)	Fractional (0.0683)
	Flexible (LOF, *d*)	Geometric (0.0025, 0.76)	Harmonic (0.0028, 1.00)	Network (0.0034, 0.59)	Arithmetic (0.0082, 1.00)
All Datasets	Fixed (LOF)	Harmonic (0.0075)	Arithmetic (0.0099)	Geometric (0.0274)	Fractional (0.0587)
	Flexible (LOF, *d*)	Geometric (0.0061, 0.68)	Harmonic (0.0066, 0.88)	Arithmetic (0.0093, 0.92)	Network (0.0099, 0.39)

2.4. Measurement

A lack of fit (*LOF*, hereafter; Browne, Cudeck, Bollen, & Long, 1993) is used to measure the error of a scheme-generated allocation against empirical data. This is a common measurement used in Hagen (2010) and Kim and Diesner (2014). Calculation of *LOF* follows the formula below.

$$\mathrm{LOF} = \frac{1}{(n-1)} \sum_{i=1}^{n} \frac{(E_i - C_i)^2}{C_i}$$

here n is the total of observations in an empirical dataset, E represents the set of coauthorship credit scores for each author in the empirical dataset, and C refers to the set of coauthorship credit scores for each author generated by a scheme. We can decide which scheme performs better than others by comparing LOF values of schemes. In other words, the low LOF of a scheme means that the scheme generates scores approximating those in empirical data better than other schemes that show higher LOFs.

In addition, measuring LOF can be used to find the optimal distribution parameter for each flexible-type scheme. Given a set of scores in empirical data, one can generate various sets of coauthorship credit scores by applying one parameter after another to the test scheme. In this study, parameters with the increment of 0.01 between zero (exclusive) and one (inclusive) are tried for each flexible scheme. Among those parameters, one can pick up the parameter that produces a score set resulting in the lowest LOF against empirical data. In this way, all flexible schemes can find their lowest LOFs against the same empirical dataset and then be compared to decide which scheme performs best.

3. Result

While Hagen (2010, 2013) tested a scheme's performance against the whole set of empirical data, Kim and Diesner (2014) tested a scheme against an individual set of empirical data per field. Thus, we followed both approaches. In Table 4, the performance of fixed and flexible-type schemes is compared. For example, among fixed-type schemes, Arithmetic scheme performed best against the economics dataset with the lowest LOF (=0.0102), followed by Harmonic, Fractional, and Geometric schemes. Among flexible schemes tested against economics, Arithmetic scheme again excelled others with the lowest LOF (=0.0005) when the distribution parameter was set as d = 0.66.

When tested against individual datasets, Harmonic scheme was found to be the best performer for fixed types, while Arithmetic and Harmonic schemes tied for flexible types, if we count the frequency of each scheme being ranked as the top. Harmonic scheme also showed the best performance against the whole set of empirical data among fixed types, while Geometric scheme topped the list of flexible types. An interesting observation is that if we only consider the performance of fixed types for psychology, chemistry (2000), and medicine, Harmonic scheme was the best. These three datasets are almost identical to those used in Hagen (2010, 2013)[5].

Another noticeable observation is that Network scheme performed generally better than the fixed-type schemes (in terms of LOF) with a few exceptions in the medicine dataset and 'All Datasets,' but its performance was not impressive when compared to other flexible-types: it was ranked as the top only once for the chemistry (1993) dataset. This confirms the

[5] In Hagen (2010, 2013), a medicine dataset for a five-authored case, where the third author is designated as a corresponding author, was not used. But it was included in our medicine data. With that included or not, Harmonic scheme showed the best performance.

concern of Kim and Diesner (2014) that the performance of Network scheme seems to be mostly due to its flexibility as other comparative schemes were all fixed types in their study.

From the result, we can see that some schemes are not consistent in performance. For example, Geometric scheme did not show a good performance among fixed-type schemes. But it showed an impressive performance as a flexible type: it was ranked as the top when tested against medicine and all datasets, and as the second against marketing, psychology, and chemistry (1993). This illustrates that, depending on characteristics of empirical data (e.g., evenness of distributed scores) and the operationalization of scheme (e.g., assignment of flexibility or incorporation of equal contributors), the performance of a scheme can change. For example, if coauthorship scores in empirical data are distributed with minimal differences (i.e., almost even), Fractional scheme would be the best performer.

4. Conclusion and discussion

This paper attempted to compare coauthorship credit allocation schemes against empirical data. The main finding is that there is no absolute winner scheme in approximating real-world empirical data. Although we cannot specify the number of datasets enough to validate the superiority of an authorship credit scheme over others, we need to be cautious about an argument that a scheme is superior because it performs better than others in fitting against a few available/selective datasets[6] or under unequal conditions. This paper, however, does not attempt to discuss strengths and weaknesses of comparative models, to claim the superiority of one scheme over others, or to refute any previous research: rather, its implication is that the performance of a coauthorship credit scheme does not seem to rely much on its innate feature (e.g., a sophisticated formula) but on external conditions (e.g., test datasets and operationalization of constraints). For example, although Harmonic scheme showed the best performance among fixed types of models for five out of seven datasets in our study, it may show less than best performance against new empirical data[7].

Another consideration is that the practical implication of ranking schemes according to their fit to empirical data may not be that useful. In many cases, LOFs between the top and other schemes were close and their difference does not seem to be much discriminatory. It is also unclear what level of LOF is acceptable or not in order to recognize a scheme as a good performer.

Such a dependency of a scheme's performance on external conditions leads us to rethink the purpose of various allocation schemes. Fixed-type schemes may be used to propose a set of coauthorship credit allocation that can serve as a preferred or normative standard across fields as Hagen (2010, 2013) did[8]. Flexible types are more suitable to reflect assignment practices in specific fields: a researcher or an academic committee may use them to evaluate a field or a scholar's research output, adjusting one of flexible schemes to the tradition of a field (e.g., the last author is assumed to contribute twice more than the first author) or a special requirement (e.g., a graduate student as the first author receives a half of its due credit).

As such, we suggest that each scheme has a unique purpose and strengths in the context where it is proposed and, thus, that it would be inadequate to directly compare different schemes based on available/selective data or a purpose well-suited for a particular one. For example, Harmonic scheme, which has been shown to quite closely represent credit allocation in some fields, needs more empirical validation against data from diverse fields before it claims its utility as a "parsimonious solution to the problem of quantifying the byline hierarchy" (Hagen, 2013, p.790). Network-Based scheme has been recently applied to a study of the coauthorship network of a psychology journal to detect a hierarchical structure of collaboration (e.g., top scholars in terms of coauthorship credits tend to collaborate with other top scholars) (Kim & Diesner, 2015). Here, the contribution of Network model is that it successfully incorporates a credit allocation scheme into a coauthorship network, where the order and number of coauthors have been traditionally neglected, not that it is better than other schemes in approximating empirical data under some favorable conditions.

This paper is not without limitations. It could not cover other well-designed fixed or flexible schemes that have been proposed in, for example, Liu and Fang (2012) or Trueba and Guerrero (2004). In addition, choice of a different measurement for errors, instead of LOF, might reveal other insights. Also, it would be interesting to see what if we include cases where number of coauthors is large, e.g., six or more, for testing model performance. If we include more comparative schemes, measurements, and larger authorship cases, we may obtain more detailed, nuanced observations on the performance of

[6] In addition to data availability or biased selection, data quality should be considered as a confounding factor in measuring the performance of schemes. National or cultural difference might affect credit allocation. For example, Vinkler (1993) drew a list of credit shares from chemists in a Hungarian institute. If the study of Wren et al. (2007) was conducted in China, where the first author or a few co-first-authors get the most of credit (Jian and Xiaoli, 2013), respondents might report more skewed allocation. Survey design also matters. Respondents in Maciejovsky et al. (2008) and Wren et al. (2007) might respond differently to computer-generated or hypothetical byline lists and to actual author lists from journals. Wren et al. (2007) might produce different score sets if the study included the case where the first author is the corresponding author, which seems to be dominant in top medicine journals.

[7] In Harmonic scheme, the ratio of credit distributed to an author (r) and its immediate follower author ($r+1$) is fixed at ($r+1$): r (Hagen, 2010). Such a scheme may not represent well a real-world case where one or a few authors get most of the coauthorship credit (Geometric scheme or any flexible type with a high d value may perform better here) or all authors get unequal credits with marginal differences (Fractional scheme or any flexible type with a low d value would fit well).

[8] Hagen (2010, 2013) suggests that a credit allocation scheme should reflect three "ethical"conditions: "(1) one publication credit is shared among all coauthors, (2) the first author gets the most credit, and in general the ith author receives more credit than the ($i+1$)th author, and (3) the greater the number of authors, the less credit per author." According to the studies, Harmonic scheme satisfies these requirements while other schemes (Fractional, Arithmetic, and Geometric) fail to do so.

credit allocation schemes. In terms of illustrating that the performance of a scheme can differ depending on test conditions, however, we believe this paper can work as a nudge to attract attention from scholars interested in this topic and motivate them to use caution when proposing, testing, and comparing coauthorship credit allocation schemes.

Acknowledgement

The authors are grateful to Professor Boris Maciejovsky for providing original datasets and anonymous reviewers for constructive comments.

References

Abramo, G., D'Angelo, C. A., & Rosati, F. (2013). The importance of accounting for the number of co-authors and their order when assessing research performance at the individual level in the life sciences. *Journal of Informetrics*, *7*(1), 198–208. http://dx.doi.org/10.1016/j.joi.2012.11.003

Beasley, B. W., & Wright, S. M. (2003). Looking forward to promotion: Characteristics of participants in the prospective study of promotion in academia. *Journal of General Internal Medicine*, *18*(9), 705–710. http://dx.doi.org/10.1046/j.1525-1497.2003.20639.x

Browne, M. W., Cudeck, R., Bollen, K. A., & Long, J. S. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (A Sage Focus ed., pp. 136–1136). Newbury Park, CA: Sage Publications.

Carbone, V. (2011). *Fractional counting of authorship to quantify scientific research output*. , arXiv preprint arXiv:1106.0114.

Egghe, L., Rousseau, R., & Van Hooydonk, G. (2000). Methods for accrediting publications to authors or countries: Consequences for evaluation studies. *Journal of the American Society for Information Science*, *51*(2), 145–157.

Frandsen, T. F., & Nicolaisen, J. (2010). What is in a name? Credit assignment practices in different disciplines. *Journal of Informetrics*, *4*(4), 608–617. http://dx.doi.org/10.1016/j.joi.2010.06.010

Hagen, N. T. (2010). Harmonic publication and citation counting: Sharing authorship credit equitably–not equally, geometrically or arithmetically. *Scientometrics*, *84*(3), 785–793.

Hagen, N. T. (2013). Harmonic coauthor credit: A parsimonious quantification of the byline hierarchy. *Journal of Informetrics*, *7*(4), 784–791.

He, B., Ding, Y., & Yan, E. J. (2012). Mining patterns of author orders in scientific publications. *Journal of Informetrics*, *6*(3), 359–367. http://dx.doi.org/10.1016/j.joi.2012.01.001

Hu, X. (2009). Loads of special authorship functions: Linear growth in the percentage of "equal first authors" and corresponding authors. *Journal of the American Society for Information Science and Technology*, *60*(11), 2378–2381.

Jian, D., & Xiaoli, T. (2013). Perceptions of author order versus contribution among researchers with different professional ranks and the potential of harmonic counts for encouraging ethical co-authorship practices. *Scientometrics*, *96*(1), 277–295. http://dx.doi.org/10.1007/s11192-012-0905-4

Kim, J., & Diesner, J. (2014). A network-based approach to coauthorship credit allocation. *Scientometrics*, *101*(1), 587–602. http://dx.doi.org/10.1007/s11192-014-1253-3

Kim, J., & Diesner, J. (2015). Coauthorship networks: A directed network approach considering the order and number of coauthors. *Journal of the Association for Information Science and Technology*, http://dx.doi.org/10.1002/asi.23361

Lindsey, D. (1980). Production and citation measures in the sociology of science: The problem of multiple authorship. *Social Studies of Science*, *10*(2), 145–162. http://dx.doi.org/10.1177/030631278001000202

Liu, X. Z., & Fang, H. (2012). Modifying *h*-index by allocating credit of multi-authored papers whose author names rank based on contribution. *Journal of Informetrics*, *6*(4), 557–565. http://dx.doi.org/10.1016/j.joi.2012.05.002

Lukovits, I., & Vinkler, P. (1995). Correct credit distribution: A model for sharing credit among coauthors. *Social Indicators Research*, *36*(1), 91–98.

Maciejovsky, B., Budescu, D. V., & Ariely, D. (2008). The researcher as a consumer of scientific publication: How do name-ordering conventions affect inferences about contribution credits? *Marketing Science*, *28*(3), 589–598. http://dx.doi.org/10.1287/mksc.1080.0406

Marušić, A., Bošnjak, L., & Jerončić, A. (2011). A systematic review of research on the meaning, ethics and practices of authorship across scholarly disciplines. *PLoS ONE*, *6*(9), e23477.

Mattsson, P., Sundberg, C. J., & Laget, P. (2011). Is correspondence reflected in the author position? A bibliometric study of the relation between corresponding author and byline position. *Scientometrics*, *87*(1), 99–105. http://dx.doi.org/10.1007/s11192-010-0310-9

Oppenheim, C. (1998). Fractional counting of multiauthored publications. *Journal of the American Society for Information Science*, *49*(5), 482–1482. http://dx.doi.org/10.1002/(Sici)1097-4571(19980415)49:5<482::Aid-Asi11>3.0.Co;2-H

Price, D. D. (1981). Multiple Authorship. *Science*, *212*(4498), 986.

Thomas, P. A., Diener-West, M., Canto, M. I., Martin, D. R., Post, W. S., & Streiff, M. B. (2004). Results of an academic promotion and career path survey of faculty at the Johns Hopkins University School of Medicine. *Academic Medicine*, *79*(3), 258–264. http://dx.doi.org/10.1097/00001888-200403000-00013

Trueba, F. J., & Guerrero, H. (2004). A robust formula to credit authors for their publications. *Scientometrics*, *60*(2), 181–204.

Van Hooydonk, G. (1997). Fractional counting of multiauthored publications: Consequences for the impact of authors. *Journal of the American Society for Information Science*, *48*(10), 944–945. http://dx.doi.org/10.1002/(Sici)1097-4571(199710)48:10<944::Aid-Asi8>3.0.Co;2-1

Vinkler, P. (1993). Research contribution, authorship and team cooperativeness. *Scientometrics*, *26*(1), 213–230.

Vinkler, P. (2000). Evaluation of the publication activity of research teams by means of scientometric indicators. *Current Science—Bangalore*, *79*(5), 602–612.

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

Wren, J. D., Kozak, K. Z., Johnson, K. R., Deakyne, S. J., Schilling, L. M., & Dellavalle, R. P. (2007). The write position: A survey of perceived contributions to papers based on byline position and number of authors. *Embo Reports*, *8*(11), 988–991. http://dx.doi.org/10.1038/sj.embor.7401095

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036–1039. http://dx.doi.org/10.1126/science.1136099