Journal of Informetrics 7 (2013) 784–791

Contents lists available at ScienceDirect

Journal of Informetrics

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

Harmonic coauthor credit: A parsimonious quantification of the byline hierarchy

Nils T. Hagen*

Faculty of Biosciences and Aquaculture, University of Nordland, N-8049 Bodø, Norway

ARTICLE INFO

Article history:
Received 27 March 2013
Received in revised form 28 May 2013
Accepted 26 June 2013

Keywords:
Byline hierarchy
Coauthor credit
Harmonic formula
Bibliometric bias
Parsimony

ABSTRACT

In this paper the accuracy of five current approaches to quantifying the byline hierarchy of a scientific paper is assessed by measuring the ability of each to explain the variation in a composite empirical dataset. Harmonic credit explained 97% of the variation by including information about the number of coauthors and their position in the byline. In contrast, fractional credit, which ignored the byline hierarchy by allocating equal credit to all coauthors, explained less than 40% of the variation in the empirical dataset. The nearly 60% discrepancy in explanatory power between fractional and harmonic credit was accounted for by equalizing bias associated with the omission of relevant information about differential coauthor contribution. Including an additional parameter to describe a continuum of intermediate formulas between fractional and harmonic provided a negligible or negative gain in predictive accuracy. By comparison, two parametric models from the bibliometric literature both had an explanatory capacity of approximately 80%. In conclusion, the results indicate that the harmonic formula provides a parsimonious solution to the problem of quantifying the byline hierarchy. Harmonic credit allocation also accommodates specific indications of departures from the basic byline hierarchy, such as footnoted information stating that some or all coauthors have contributed equally or indicating the presence of a senior author.

1. Introduction

Most scientific papers are multi-authored (Wuchty, Jones, & Uzzi, 2007), by coauthors who have not contributed equally (Waltman, 2012). Nevertheless, it is still routine practice to quantify authorship credit by dividing one credit equally among all coauthors of a paper irrespective of their actual contribution, thereby underestimating the credit of primary authors and overestimating the credit of secondary authors (Hagen, 2008). A more tenable approach is to estimate each coauthor's share of credit as accurately as possible by including all relevant information about the relative size of each contribution when calculating coauthor credit (Hagen, 2010a, 2010b).

Accurate quantification of coauthor credit requires a formulaic interpretation of the byline hierarchy which by convention lists coauthors in order of decreasing contribution (Lake, 2010; Zuckerman, 1968). It is also important that the quantification procedure accommodates all specific indications of departures from the basic byline hierarchy, for example footnoted information stating that some or all coauthors have contributed equally (Akhabue & Lautenbach, 2010; Frandsen & Nicolaisen,

* Tel.: +47 755 17 398.
E-mail address: Nils.Hagen@uin.no

1751-1577
http://dx.doi.org/10.1016/j.joi.2013.06.005

2010; Hu, 2009) or indicating the presence of a senior author (Buehring, Buehring, & Gerard, 2007; Mattsson, Sundberg, & Laget, 2011).

The harmonic formula, first proposed by Hodge and Greenberg (1981), fits these requirements by providing a straightforward quantification of the byline hierarchy that is easily modified to include specific information about the seniority or equality of some coauthors (Hagen, 2008), while demonstrating a robust fit when validated against empirical data from medicine, psychology, and chemistry (Hagen, 2010b). Recently, Liu and Fang (2012a, 2012b) proposed adding an additional parameter to the harmonic formula, in order to define a continuum of intermediate formulas ranging from fractional equal credit at one extreme to harmonic credit at the other, and suggested that a parameter value corresponding to two thirds of the distance between fractional and harmonic would be the preferred alternative.

In this paper I use the harmonic and fractional formulas as reference points when comparing the performance of Liu and Fangs's (2012a, 2012b) model with earlier parametric models from the bibliometric literature proposed by Lukovits and Vinkler (1995), and by Trueba and Guerrero (2004). I assess these formulas by quantifying their ability to explain the variation in a composite set of empirical data on perceived notions of coauthor contribution in chemistry (Vinkler, 2000), medicine (Wren et al., 2007), and psychology (Maciejovsky, Budescu, & Ariely, 2009), adapted for bibliometric analysis by Hagen (2010b). Finally, I conclude by discussing the results with reference to the problem of overfitting and the principle of parsimony.

2. Material and methods

The empirical data consists of three independent samples from the scientific subfields of chemistry (Vinkler, 2000), medicine (Wren et al., 2007), and psychology (Maciejovsky et al., 2009), that were extracted from the original publications as described by Hagen (2010b).

The consolidated data set describes perceived notions of coauthor credit for papers with up to 6 coauthors (Table 1). The data for medicine (Wren et al., 2007), imply the presence of a senior last author and support the assumption that the senior and first authors have contributed equally (Hagen, 2010b).

3. Theory/calculation

Harmonic authorship credit for the ith author of a publication with N coauthors is calculated according to the following formula (Hagen, 2008; Hodge & Greenberg, 1981):

$$\text{Harmonic}\, i\text{th author credit} = \frac{1/i}{[1 + 1/2 + \cdots + 1/N]} \quad (1)$$

Evidence of senior authorship in the scientific subfield of medicine (Table 1), was included in the calculation as described by Hagen (2008), by assuming that the senior author and the first author had contributed equally (Hagen, 2010b).

Fractional credit is calculated as follows:

$$\text{Fractional}\, i\text{th author credit} = \frac{1}{N} \quad (2)$$

Liu and Fang's (2012a, 2012b) model for coauthor credit is identical to fractional credit when the tuning parameter $q = 0$, identical to harmonic credit when $q = 1$, and provides a continuous range of potential formulas for $0 < q < 1$.

$$\text{Liu and Fang's (2012a, 2012b)}\, i\text{th author credit} = \frac{i^{-q}}{\sum_{j=1}^{N} j^{-q}} \quad (3)$$

Evidence of senior authorship in the scientific subfield of medicine (Table 1), was included in the calculation of Liu and Fang's (2012a, 2012b) model as described above for harmonic credit.

According to Lukovits and Vinkler's (1995) model, coauthor credit is calculated as follows:

$$\text{1st author credit} = \frac{N+1}{2NF}, \text{ and} \quad (4)$$

$$i\text{th author credit} = \frac{i+T}{2iFT}, \text{ for } i = 2, \ldots, N, \text{ where} \quad (5)$$

$$F = \frac{1}{2}\left[\frac{1}{N} + \frac{N+1}{T} + \sum_{j=1}^{N} \frac{1}{j}\right], \text{ and } T = \frac{100}{H}.$$

The tuning parameter H is the percentage value of their contribution threshold. To facilitate comparison among the different tuning parameters, H is expressed as a fraction in Fig. 3.

According to Trueba and Guerrero's (2004) model, coauthor credit is calculated as follows:

$$\text{1st author credit} = \frac{2N+1}{N \cdot (N+1)} \cdot \frac{2}{3} \cdot (1-f) + c_1 \cdot f \quad (6)$$

Table 1
Empirical[a] data and calculated[b] model estimates of coauthor credit for scientific papers with N coauthors, whose byline rank i corresponds to their contribution to the paper. The coefficient of determination R^2 measures how well a model explains the variation in the empirical data set.

Scientific subfield	N	i	Empirical coauthor credit	Fractional coauthor credit	Harmonic coauthor credit	Liu and Fang	Lukovits and Vinkler	Trueba and Guerrero
R^2				0.3863	0.9676	0.9483	0.7942	0.8169
Chemistry								
	2	1	0.65	0.50	0.6667	0.6135	0.7143	0.5704
		2	0.35	0.50	0.3333	0.3865	0.2857	0.3630
	3	1	0.55	0.33	0.5455	0.4738	0.5634	0.4593
		2	0.25	0.33	0.2727	0.2985	0.2535	0.2889
		3	0.20	0.33	0.1818	0.2278	0.1831	0.1852
	4	1	0.50	0.25	0.4800	0.3988	0.4747	0.4000
		2	0.25	0.25	0.2400	0.2512	0.2278	0.2444
		3	0.15	0.25	0.1600	0.1917	0.1646	0.1556
		4	0.10	0.25	0.1200	0.1583	0.1329	0.1333
	5	1	0.40	0.20	0.4380	0.3509	0.4162	0.3630
		2	0.25	0.20	0.2190	0.2211	0.2081	0.2148
		3	0.15	0.20	0.1460	0.1687	0.1503	0.1333
		4	0.10	0.20	0.1095	0.1393	0.1214	0.1185
		5	0.10	0.20	0.0876	0.1200	0.1040	0.1037
	6	1	0.35	0.17	0.4082	0.3172	0.3743	0.3376
		2	0.25	0.17	0.2041	0.1998	0.1925	0.1937
		3	0.15	0.17	0.1361	0.1525	0.1390	0.1164
		4	0.10	0.17	0.1020	0.1259	0.1123	0.1058
		5	0.10	0.17	0.0816	0.1085	0.0963	0.0952
		6	0.05	0.17	0.0680	0.0961	0.0856	0.0847
Medicine								
	3	1	0.42	0.33	0.4091	0.3861	0.5634	0.4593
		2	0.17	0.33	0.1818	0.2278	0.2535	0.2889
		3	0.41	0.33	0.4091	0.3861	0.1831	0.1852
	5	1	0.34	0.20	0.3285	0.2860	0.4162	0.3630
		2	0.12	0.20	0.1460	0.1687	0.2081	0.2148
		3	0.08	0.20	0.1095	0.1393	0.1503	0.1333
		4	0.07	0.20	0.0876	0.1200	0.1214	0.1185
		5	0.38	0.20	0.3285	0.2860	0.1040	0.1037
Psycology								
	2	1	0.60	0.50	0.6667	0.6135	0.7143	0.5704
		2	0.39	0.50	0.3333	0.3865	0.2857	0.3630
	3	1	0.49	0.33	0.5455	0.4738	0.5634	0.4593
		2	0.29	0.33	0.2727	0.2985	0.2535	0.2889
		3	0.22	0.33	0.1818	0.2278	0.1831	0.1852
	4	1	0.42	0.25	0.4800	0.3988	0.4747	0.4000
		2	0.24	0.25	0.2400	0.2512	0.2278	0.2444
		3	0.19	0.25	0.1600	0.1917	0.1646	0.1556
		4	0.15	0.25	0.1200	0.1583	0.1329	0.1333

[a] The empirical coauthor credit data are adapted from three independent data sets on perceived coauthor contribution from the scientific subfields of chemistry (Vinkler, 2000), medicine (Wren et al., 2007), and psychology (Maciejovsky et al., 2009), as detailed by Hagen (2010b).

[b] The fractional model divides credit equally among N coauthors and ignores information about differential coauthor contribution. The harmonic model allocates credit according to coauthor rank i, on the assumption that the rank corresponds to the size of a coauthor's contribution to the paper. The harmonic model accommodates evidence suggesting the presence of a senior last author in the subfield of medicine, and that the senior and first authors contributed equally (Hagen, 2010b). Liu and Fang's (2012a, 2012b) model includes a tuning parameter q, which specifies the fractional and harmonic models as extreme cases at $q = 0$ and $q = 1$. The tabulated values for Liu and Fang's model are calculated with their proposed value of $q = 2/3$, and also assuming the presence of a senior author in the subfield of medicine. Lukovits and Vinkler's (1995) model is used with their proposed parameter value $H = 10\%$. Trueba and Guerrero's (2004) model is used with their proposed parameter values $f = 1/3$, $c_1 = 0.6$, and $c_2 = c_3 = 0.2$.

$$\text{2nd author credit} = \frac{2}{(N+1)} \cdot \frac{2}{3} \cdot (1-f) + c_2 \cdot f \tag{7}$$

$$N\text{th author credit} = \frac{N+2}{N \cdot (N+1)} \cdot \frac{2}{3} \cdot (1-f) + c_3 f \tag{8}$$

$$i\text{th author credit} = \frac{2N-i+2}{N \cdot (N+1)} \cdot \frac{2}{3} \cdot (1-f), \quad \text{for } i = 3, \ldots, (N-1) \tag{9}$$

The tuning parameter f determines the proportion of extra credit to be shared among the first, second and last authors, according to the coefficients $c_1 + c_2 + c_3 = 1$ and $c_1 \geq c_2 \geq c_3$.

All credit values and the coefficient of determination (R^2) were calculated using *R* (http://www.r-project.org/). The R^2 values were calculated for each of the tuning parameters *q*, *H*/100, and *f*, in 0.01 increments throughout their entire range from 0 to 1. For each increment of *f*, the coefficients c_1, c_2, and c_3 were systematically varied to maximize the corresponding R^2 value. The figures were made using the *R* lattice package (Sarkar, 2008).

4. Results

Harmonic credit explained nearly 97% of the variation in the empirical data ($R^2 = 0.9676$, Table 1). This level of accuracy is achieved by allocating credit according to a straightforward interpretation of the byline hierarchy which assumes that coauthors, in the absence of specific indications to the contrary, are listed in decreasing order of contribution (Hagen, 2008; Hodge & Greenberg, 1981). Consequently, the only information required to calculate basic harmonic credit is the number and order of a paper's coauthors. Furthermore, harmonic credit uses an evidence-based approach to the inclusion of supplementary information which may specify or imply seniority or equality of certain coauthors. Here, in the present analysis, equal harmonic credit was allocated to the first and last authors in the subfield of medicine, since the data for medicine implied the presence of a senior last author and supported the assumption that the senior and the first author had contributed approximately equally (Table 1).

By comparison, the models of Liu and Fang (2012a, 2012b), Lukovits and Vinkler (1995), and Trueba and Guerrero (2004), all included additional parameters but explained less of the variation in the empirical data when credit was calculated in accordance with the respective authors' proposed parameter values (Fig. 1, Table 1).

Liu and Fang's (2012a, 2012b) model uses the tuning parameter *q*, which defaults to fractional credit at one extreme of its range ($q = 0$), and harmonic credit at the other ($q = 1$). Fractional credit ignores differential coauthor contribution by dividing one credit equally among a paper's authors, and explained only 38.63% of the variation in the empirical data set (Table 1). Increasing the value of *q* incrementally improved the models performance visibly as the calculated credit values approached the harmonic formula at $q = 1$ (Fig. 2). A more detailed incremental analysis revealed that the explanatory power of Liu and Fang's (2012a, 2012b) model peaked at $q = 0.89$ with $R^2 = 0.9707$, which is 2.16% better than their proposed value of $q = 2/3$, and 0.31% better than harmonic credit at $q = 1$ (Fig. 3).

Similarly, the explanatory power of Lukovits and Vinkler's (1995) model peaked at $H = 21\%$ with $R^2 = 0.8039$, which is only marginally (0.96%) better than $R^2 = 0.7942$ at their proposed value of $H = 10\%$ (Table 1, Fig. 3).

Trueba and Guerrero's (2004) model is more complex than the other models. Its primary parameter *f* determines a fixed proportion of extra credit for the first, second, and last authors, and three additional coefficients c_1, c_2 and c_3, determine how to divide this extra credit among the specified authors subject to the restriction that $c_1 \geq c_2 \geq c_3$. The explanatory power of their model peaked in the vicinity of the proposed value of $f = 1/3$, but with $c_1 = 0.7$ and $c_2 = c_3 = 0.15$ rather than the proposed values of $c_1 = 0.6$ and $c_2 = c_3 = 0.2$. The peak value of $R^2 = 0.8328$ is 1.5% better than $R^2 = 0.8189$ at the proposed values (Table 1, Fig. 3).

In summary, these results indicate that fractional credit explained less than 40% of the variation in the empirical data; the parametric models of Lukovits and Vinkler (1995) and Trueba and Guerrero (2004), although a great improvement over fractional, still only explained approximately 80% of the variation; while harmonic credit explained almost 97%. Moreover, the addition of a tuning parameter (Liu & Fang, 2012a, 2012b) had a negative or negligible effect on the excellent performance of the harmonic formula.

5. Discussion

Harmonic coauthor credit, which includes information about the relative size of each coauthor's contribution, explained nearly 97% of the variation in the empirical dataset, as opposed to fractional coauthor credit which explained less than 40% of the variation. The difference between fractional and harmonic credit is a measure of the equalizing bias that is generated when equal fractional credit is misallocated to coauthors who have not contributed equally (Hagen, 2008). This result implies that equalizing bias accounted for approximately 60% of the total variation in the empirical data set.

The maximum explanatory power of the parametric models of Lukovits and Vinkler (1995) and Trueba and Guerrero (2004) although approximately 40% better than fractional credit, was nearly 20% below harmonic credit (Fig. 3). Contrary to Trueba and Guerrero's (2004) claim of superior complexity-induced performance, their model exhibited only intermediate predictive accuracy despite the added complexity of three additional coefficients c_1, c_2 and c_3, and did not provide any substantial improvement in explanatory capacity over the earlier single-parameter model of Lukovits and Vinkler (1995).

Adding an additional tuning parameter to the harmonic formula, as proposed by Liu and Fang (2012a, 2012b), with their suggested value of $q = 2/3$, reduced the explanatory power of the harmonic formula by nearly 2%. Moreover, the maximum potential improvement in explanatory power derived from the additional parameter was a mere 0.31% (Fig. 3). It would seem, therefore, that the added complexity of the additional parameter *q* is a case of overfitting, as the added parameter provided a negligible or negative gain in predictive accuracy, and no substantial new insight into the underlying causality of the byline hierarchy.

The opposite problem of underfitting arises when an important explanatory variable is missing from the predicitive model. Here, the fractional formula provides an example of the serious consequences of underfitting, by demonstrating the acute loss

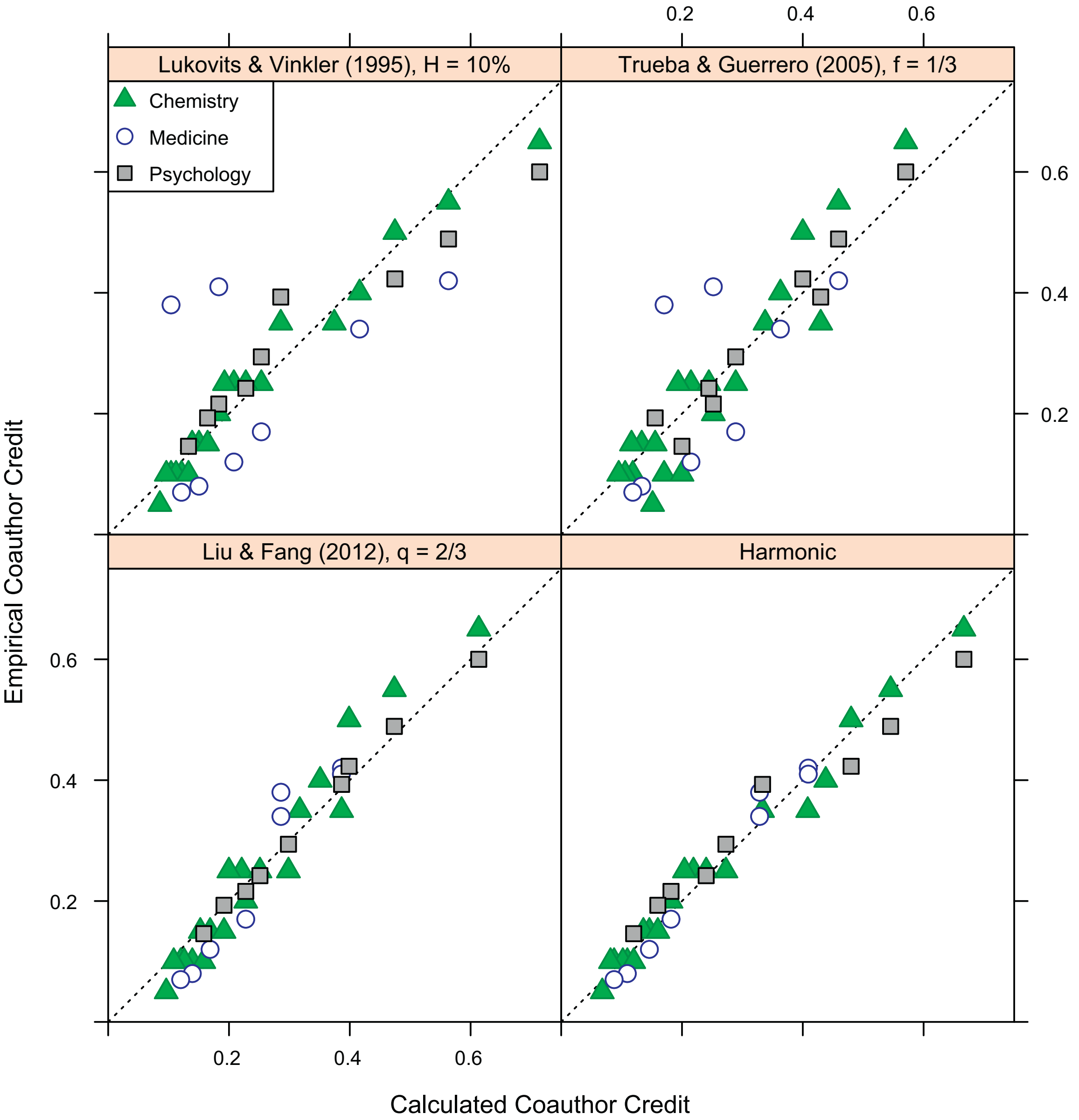

Fig. 1. Calculated versus empirical estimates of coauthor credit. Harmonic credit is allocated according to a simple model of the byline hierarchy which assumes that coauthors, in the absence of specific indications to the contrary, are listed in decreasing order of contribution (Hagen, 2008; Hodge & Greenberg, 1981). Credit estimates from the models of Liu and Fang (2012a, 2012b), Lukovits and Vinkler (1995), and Trueba and Guerrero (2004), are calculated in accordance with the respective authors' proposed parameter values. The dotted diagonal reference lines indicate a perfect fit between empirical and calculated coauthor credit. The empirical coauthor credit is adapted from three independent data sets from the scientific subfields of chemistry (Vinkler, 2000), medicine (Wren et al., 2007), and psychology (Maciejovsky et al., 2009), assuming the presence of a senior last author in the subfield of medicine, and that the senior and first authors contributed equally (Hagen, 2010b). $N = 37$ observations.

of explanatory power associated with omission of relevant information about the ranking order of the byline hierarchy. In an underfitted model the missing information generates bias which underestimates some values and overestimates others. The specific bias associated with the fractional model is known as equalizing bias because by allocating equal credit to coauthors from a paper with a hierarchical byline, the fractional model underestimates the credit of primary coauthors and overestimates the credit of secondary coauthors (Hagen, 2008).

To strike a balance between underfitting and overfitting, the principle of parsimony dictates that superfluous complexity should be eliminated by adherence to the simplest satisfactory solution to a modeling problem (e.g. Forster, 2000). Accordingly, the harmonic formula offers a parsimonious solution to the quest for a formulaic quantification of the byline hierarchy because it explains nearly all of the variation in the empirical dataset simply by including information about the number of

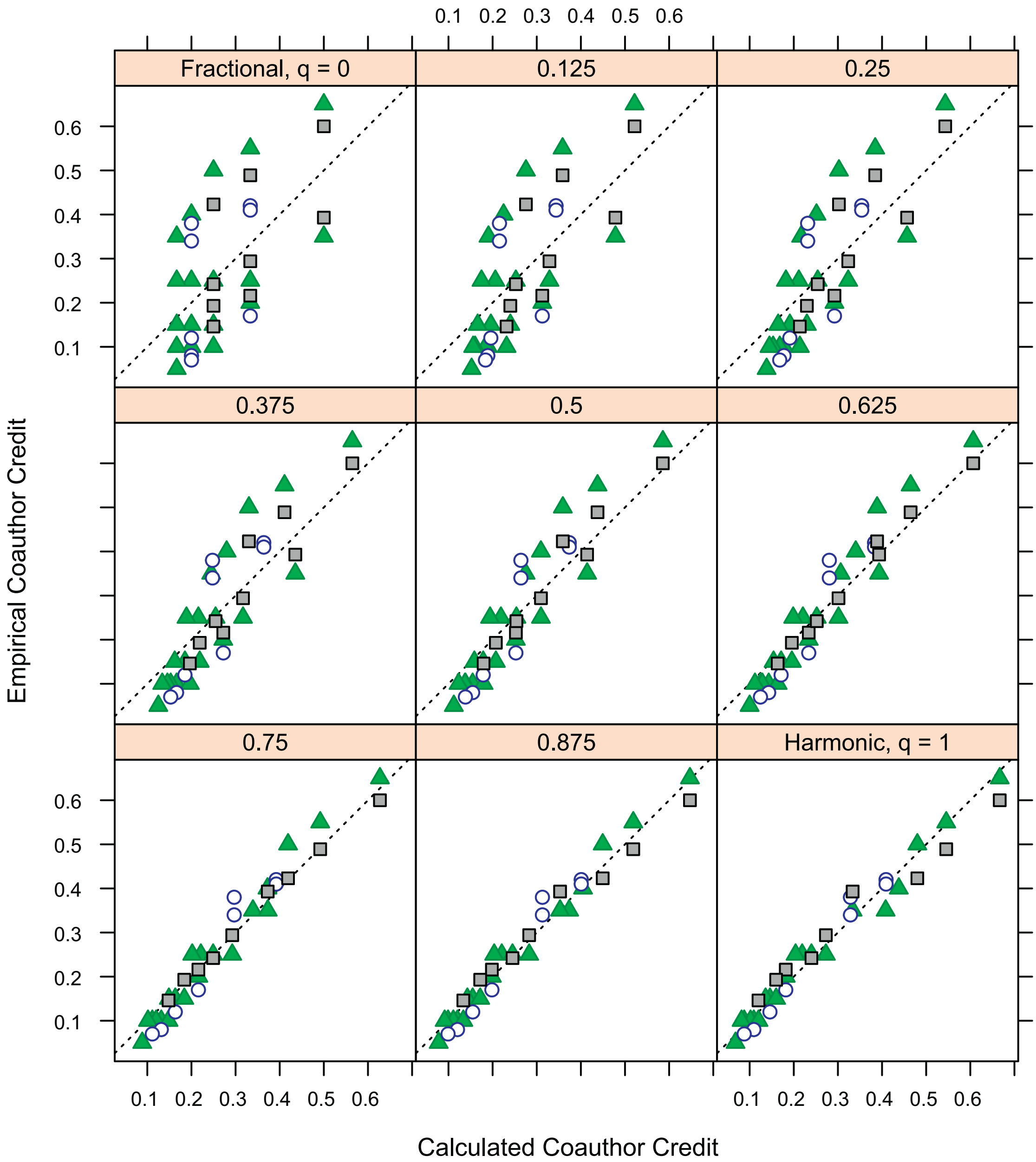

Fig. 2. Effect of q. Empirical data of perceived coauthor credit from three different subfields of science are plotted against calculated estimates of coauthor credit for a range of different q-values, as indicated in the strip above each panel. Fractional credit at $q = 0$ is allocated equally to all coauthors of a scientific paper, irrespective of their actual contribution. Harmonic credit at $q = 1$ is a parsimonious formula where q is redundant. The key to the plot symbols is included in Fig. 1. The dotted diagonal reference lines indicate a perfect fit between empirical and calculated coauthor credit.

coauthors and their position in the byline. The transparent simplicity of the harmonic formula arises from its compliance with three explicit ethical criteria capturing the essential features of the byline hierarchy (Hagen, 2010b):

(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 (ith + 1)th author.
(3) The greater the number of authors, the less credit per author.

These criteria are not met by the fractional formula, neither are they consistently met by other simple solutions from the bibliometric literature, such as the arithmetic and geometric formulas (Hagen, 2010b). The underlying structure of Liu and Fang's (2012a, 2012b), Lukovits and Vinkler's (1995), and Trueba and Guerrero's (2004) parametric models is consistent

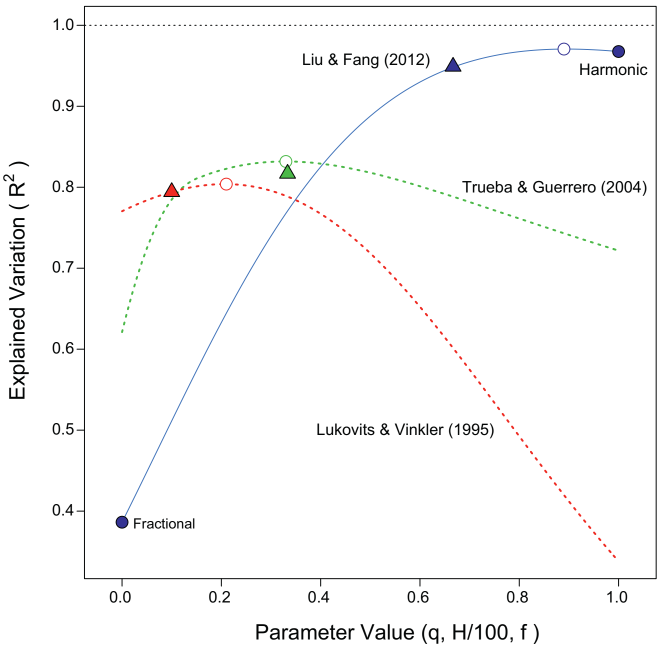

Fig. 3. Explanatory capacity for three formulaic quantifications of the byline hierarchy expressed as a function of parameter value. The coefficient of determination (R^2) is plotted for the full range of values from zero to one for the parameters q (Liu & Fang, 2012a, 2012b), H/100 (Lukovits & Vinkler, 1995), and f (Trueba & Guerrero, 2004). Fractional credit is indicated at q = 0, and harmonic credit at q = 1. Triangles indicate the respective authors' proposed parameter values of q = 2/3, H = 10%, and f = 1/3. Open circles indicate the maximal explanatory capacity of each formula. The dotted horizontal reference line at R^2 = 1.0 indicates a perfect fit explaining 100% of the variation in the empirical dataset.

with the criteria, but these authors do not recognize additional byline information specifying or implying the equality of some or all coauthor contributions, and their approach to seniority is rule-based rather than evidence-based.

Liu and Fang (2012a) propose to elevate the rank of all corresponding authors, which may create inconsistencies if the corresponding author occupies an intermediate rank as a signal of adherence to the conventional byline hierarchy (Hagen, 2010a). Trueba and Guerrero's (2004) parametric approach to senior authorship generates inconsistency if their parameter $c_3 \geq 0$ and the last author is not a senior author, and Lukovits and Vinkler's (1995) model does not recognize senior authorship.

In sum, the harmonic formula offers a verifiable combination of parsimony and accuracy to the ubiquitous bibliometric task of equitably quantifying the byline hierarchy. To ensure maximum accuracy harmonic credit allocation also accommodates specific indications of departures from the basic byline hierarchy (Hagen, 2008, 2010a), such as footnoted information stating that some or all coauthors have contributed equally (Akhabue & Lautenbach, 2010; Frandsen & Nicolaisen, 2010; Hu, 2009) or indicating the presence of a senior author (Buehring et al., 2007; Mattsson et al., 2011).

6. Conclusions

(1) The harmonic formula explained nearly 97% of the variation in the empirical dataset.
(2) Fractional credit explained less than 40% of the variation.
(3) Equalizing bias accounted for approximately 60% of the variation.
(4) The effect of including an extra parameter to the harmonic formula was a negligible or negative gain in explanatory power.
(5) The harmonic formula provides a parsimonious solution to the problem of quantifying the byline hierarchy.
(6) To ensure maximum accuracy harmonic credit allocation also accommodates specific indications of departures from the basic byline hierarchy, such as footnoted information stating that some or all coauthors have contributed equally or indicating the presence of a senior author.

Acknowledgements

Thanks to H.K. Marshall for improving the style and flow of the manuscript, to the reviewers for their positive encouragement and for suggesting comparison with parametric models from the bibliometric literature, and to the University of Nordland for facilitating data analysis and manuscript preparation.

References

Akhabue, E., & Lautenbach, E. (2010). Equal contributions and credit: An emerging trend in the characterization of authorship. *Annals of Epidemiology*, *20*(11), 868–871.

Buehring, G. C., Buehring, J. E., & Gerard, P. D. (2007). Lost in citation: Vanishing visibility of senior authors. *Scientometrics*, *72*(3), 459–468.

Forster, M. R. (2000). Key concepts in model selection: Performance and generalizability. *Journal of Mathematical Psychology*, *44*(1), 205–231.

Frandsen, T. F., & Nicolaisen, J. (2010). What is in a name? Credit assignment practices in different disciplines. *Journal of Informetrics*, *4*(4), 608–617.

Hagen, N. T. (2008). Harmonic allocation of authorship credit: Source-level correction of bibliometric bias assures accurate publication and citation analysis. *PLoS ONE*, *3*(12), e4021.

Hagen, N. T. (2010a). Deconstructing doctoral dissertations: How many papers does it take to make a PhD? *Scientometrics*, *85*(2), 567–579.

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

Hodge, S. E., & Greenberg, D. A. (1981). Publication credit. *Science*, *213*(4511), 950.

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.

Lake, D. A. (2010). Who's on first? Listing authors by relative contribution trumps the alphabet. *PS: Political Science & Politics*, *43*(1), 43–47.

Liu, X. Z., & Fang, H. (2012a). Fairly sharing the credit of multi-authored papers and its application in the modification of *h*-index and *g*-index. *Scientometrics*, *91*(1), 37–49.

Liu, X. Z., & Fang, H. (2012b). Modifying *h*-index by allocating credit of multi-authored papers whose author names rank based on contribution. *Journal of Informetrics*, *6*(4), 557–565.

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

Maciejovsky, B., Budescu, D. V., & Ariely, D. (2009). The researcher as a consumer of scientific publications: How do name-ordering conventions affect inferences about contribution credits? *Marketing Science*, *28*(3), 589–598.

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*, 99–105.

Sarkar, D. (2008). *Lattice: Multivariate data visualization with R*. New York: Springer.

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

Vinkler, P. (2000). Evaluation of the publication activity of research teams by means of scientometric indicators. *Current Science*, *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.

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.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036–1039.

Zuckerman, H. A. (1968). Patterns of name ordering among authors of scientific papers: A study of social symbolism and its ambiguity. *American Journal of Sociology*, *74*(3), 276–291.