On Evaluating Author's Performance by Publications: An Axiomatic Study

Mukherjee, Conan; Alam, Aftab

2016

Document Version:
Other version

Link to publication

Citation for published version (APA):
Mukherjee, C., & Alam, A. (2016). *On Evaluating Author's Performance by Publications: An Axiomatic Study*. (Working Papers; No. 2016:14).

Total number of authors:
2

Working Paper 2016:14

Department of Economics
School of Economics and Management

On Evaluating Author's Performance by Publications: An Axiomatic Study

Conan Mukherjee
Aftab Alam

June 2016
Revised: May 2017

LUND UNIVERSITY

On evaluating author's performance by publications: an axiomatic study

Conan Mukherjee[a,1] **and Aftab Alam**[b,2]

[a]Department of Humanities and Social Sciencies, Indian Institute of Technology Bombay, India; [b]Department of Physics, Indian Institute of Technology Bombay, India

This manuscript was compiled on May 9, 2017

Administrators in every academic institution across the world have to deal with the unenviable task of comparing researchers on the basis of their academic contributions. Unfortunately, however, there is no reasonably established consensus on the method of arriving at such comparisons, which typically involve trading off accomplishments in teaching, grant writing and academic publication. In this paper, we focus on the particular dimension of academic publication, and analyze this issue from a more fundamental perspective than addressed by the popular h-index (which may lead to unfair and counter-intuitive comparisons in certain situations). In particular, we undertake an axiomatic analysis of all possible ways to measure academic authorship for a given dataset of research articles and find that an egalitarian e-index is the *only* method which satisfies the axioms of anonymity, monotonicity, and efficiency. This index divides authorship of joint projects equally and sums across all publications of an author.

e-index | h-index | Anonymity | Monotonicity

In academic institutions and research laboratories, it is often required to evaluate and compare the cumulative impact and research performance of individuals. Such comparisons are mostly used to obtain rankings across authors; after which these rankings are used to undertake important professional decisions like hiring, promotion, granting tenure, awarding grants etc. However, a reasonably established consensus on the method of arriving at such comparisons which, typically, involve accomplishments in teaching, grant writing and academic publication, is lacking at present. In this paper, we focus on the particular issue of academic publication, and provide an intuitive method with strong justifications, to compare authors on the basis of their publication output.

Over last few years, h-index proposed by Hirsch (1), has emerged as the most popular and accepted measure of academic authorship. The h-index of an author is the largest number x of her publications, that have at least x citations. However, as we argue below, h-index is inappropriate to account for various crucial complexities inherent in academic research and so, is an unsatisfactory measure of academic performance.

The first major drawback of h-index is that, by definition, it cannot exceed the total number of papers written by an author. This feature of h-index penalizes ambitious researchers. This is because, in general, the most profound papers take a long time to write. And so, pursuing such a path-breaking paper may mean forgoing several less ambitious papers that are relatively low hanging fruits. For example, an author A may have 5 solo publications of $10,000$ citations each, and author B may have 6 joint author publications of 6 citations each. Then A has an h-index of 5 while B has an h-index of 6. Thus, while any serious academic will recognize A as the more successful researcher, h-index would state otherwise. This tendency to encourage quantity at the cost of quality is particularly harmful for any discipline. New ideas that revolutionize a discipline by generating drastic new knowledge must generate adequate rewards. As we show below, our e-index measures academic performance with respect to a combined evaluation of the total number of papers published and the associated citations and hence, addresses this issue in a substantive manner.

The second major drawback of h-index is its failure to distinguish between solo author and co-author publications. This feature entirely overlooks any need for proration of authorship in joint publications. However, this lack of proration can lead to rampant unjustified co-authorships, and sometimes, false authorship.* In fact, proliferation of such undesirable authorship has now forced several fields to explicitly formalize the definition of an author. For example, American Physical Society (APS) and International Council of Biomedical Journal Editors have established specific rules for authorship (see Section VII in Liebowitz (3)).† In an attempt to discourage such tendencies, Robert Berk, then an editor of a leading radiology journal, called for devaluing the impact of co-author papers on academic authorship, in Berk (5). In fact, our e-index follows this advice by suitably devaluing authorship generated out of joint papers in an ethical manner.

*Stephan (2) presents a concept of $false$ authorship that results out of an agreement between a group of authors to share credits of different papers with the understanding that the heavy lifting for different papers will be done, not by all, but a few members of this group. The unbridled increase in coauthorship does not preclude widespread prevalence of such authorship.

†In spite of this drastic measure, Tarnow (4) argues that there is widespread suspicion of undesirable authorship contrary to the APS definition in the discipline of Physics.

Significance Statement

Administrators in academic institutions across the world, have to deal with the unenviable task of comparing researchers on the basis of their academic publications. Unfortunately, however, there is no widespread consensus on this issue. In this paper, we address this question by conducting an axiomatic analysis of all possible ways to measure academic authorship for a given dataset of such research papers. We find that an egalitarian e-index is the *only* method which satisfies the axioms of anonymity, monotonicity, and efficiency. This index divides authorship of joint projects equally and sums across all publications of an author.

Conan Mukherjee executed the axiomatic exercise while Aftab Alam accomlished the data exercise.

[1] I am grateful to Professors Tommy Andersson, Satya R. Chakravarty, and Manipushpak Mitra for their valuable advice. Discussions with Jens Gudmundsson and Professor D. J. Saikia are also gratefully acknowledged. A substantial part of this paper was developed while visiting the Department of Economics, Lund University. Their kind hospitality is greatly acknowledged. Also, the financial support from IIT Bombay via seed grant $13IRCCSG014$ is gratefully acknowledged.
[2] Financial support from IIT Bombay via the SEED grant project code $13IRCCSG020$ is acknowledged.

[1] E-mail: conan.mukherjee@gmail.com

Finally, the h-index always returns a natural number as output. This does not create any complications, if contribution of a paper is measured in terms of simple citations. However, often value of a publication is measured by the impact factor of the journal where it appeared. This is specially true during tenure decisions of young faculties who have not had enough time for their citations to reasonably accumulate. Note that such impact factors, even for the best journal, widely vary across disciplines. In fact, they may very well take some positive fraction value. Therefore, in the extreme, if all papers of an author are published in such journals with fractional impact factors, she will get assigned the h-index equal to 0; *irrespective* of the number of papers she may have published. Thus, quite unfairly, the mere fractional value of journal impact factor identifies the author with the least possible academic credit. In contrast, our index can easily be adapted to different measures of academic authorship without entailing such criticisms.

In this paper, we present an alternative index: the e-index which is free from all these drawbacks, and is easy to interpret. The e-index for an author is obtained by equally dividing credit for all her joint projects and then summing across all her publications in the dataset. More importantly, we show that this index is a strong embodiment of certain desirable properties or *axioms* that, in our opinion, any index evaluating authorship should satisfy. In particular, it is the *only* index that satisfies the axioms of *anonymity*, *monotonicity*, and *efficiency*.

In our model, the notion of: (i) anonymity requires that an author's rank should not depend on her own identity; (ii) monotonicity requires that increase in publication quality or quantity, should enhance the academic credit assigned to a researcher; and finally (iii) efficiency requires that sum of indices across all authors, must exhaust the total number of citations arising out of the dataset. We feel that these three properties must necessarily be satisfied by any author ranking method that can be deemed reasonable. A non-anonymous method would lead authors getting compared on individual characteristics other than their academic contributions, which would lead to formation of prejudiced author rankings. A non-monotonic method could lead to loss of academic credit upon increase in academic contribution, and thus, create perverse incentives for academic research. Finally, an inefficient index that does not exhaust sum of all citations in a dataset, would not allow sufficient degree of comparability among the authors that have no joint work together.

Model

Define a dataset as the pair (P, v), where $P = \{p_1, p_2, \dots, p_k\}$ is a finite set of papers and for all $t = 1, \dots, k$, $v(p_t) \geq 0$ denotes the worth of paper p_t. These worths could be citation numbers or some other weighted measure of contribution of these papers to existing literature. For all $p \in P$, define $A(p)$ to be the set of authors of paper p and let $N^P := \cup_{t=1}^{k} A(p_t)$ be the set of all authors corresponding to a dataset. We assume that the datasets under consideration involve at least two authors, that is, $n^P := |N^P| \geq 2$. Let $\rho(N^P)$ be the set of non-empty subsets of N^P and define a function $c : \rho(N^P) \mapsto \mathbb{R}_+$ such that for all $S \in \rho(N^P)$, $c(S) := \sum_{p \in P : A(p) = S} v(p)$.

We call c, the **citation function** *induced by the dataset*. This function measures the academic contribution of each possible group of coauthors for the given dataset. Our objective is to provide an index measuring the publishing success of each author corresponding to any dataset. Note that an arbitrarily chosen dataset (P, v) may lead to any possible citation function $c \in \mathbb{R}_+^{\rho(N^P)}$. And so, an index for any dataset (P, v) is effectively a mapping $\phi : \mathbb{R}_+^{\rho(N^P)} \mapsto \mathbb{R}^{N^P}$. For simplicity of notation, henceforth, we denote $\mathbb{R}_+^{\rho(N^P)}$ by $\mathcal{C}$.

Note that any index rule must ensure that professional repute of authors depends *only* on their academic contributions and not on their identities. Hence, an index must be **anonymous** in the following manner.

Definition 1 *$\phi(c)$ satisfies anonymity if for all bijections $\pi : N^P \mapsto N^P$ and all $i \in N^P$,*

$$\phi_{\pi(i)}(\pi c) = \phi_i(c)$$

where $\pi c(\pi(S)) = c(S)$ for all $S \in \rho(N)$.

This property requires that authors be treated without any prejudices that may arise from their respective identities.

Further, any index should encourage authors to write better papers, and so, must exhibit a minimum degree of monotonicity. Otherwise, an author may find herself with lesser index than others, even after making a greater contribution to literature; and hence, would have no incentive to pursue new scientific knowledge. Therefore, an index should be **monotonic** in the following manner,

Definition 2 *ϕ satisfies monotonicity iff for all $c, d \in \mathcal{C}$, and all $i \in N^P$,*

$$[\forall\, S \subseteq N \setminus \{i\}, c(S \cup \{i\}) \geq d(S \cup \{i\})] \Longrightarrow \phi_i(c) \geq \phi_i(d)$$

This principle embodies the idea that if across any two datasets, the academic contribution of each possible group of authors that includes academic i, is *no less* for the first than the second; then the authorship ascribed to i by any index satisfying monotonicity, should be *no less* for the first than the second.

Finally, we should be able to compare all authors in the dataset, irrespective of whether they have a joint paper or not. This comparability is best achieved when an index is **efficient** in the following manner.

Definition 3 *$\phi(.)$ satisfies efficiency iff for all $c \in \mathcal{C}$ and all $i \in N^P$,*

$$\sum_{i \in N^P} \phi_i(c) = \sum_{S \in \rho(N^P)} c(S)$$

Therefore, for an efficient index, the sum total of author index values of all authors in N^P should be equal to the total number of citations in the dataset.

Now, we define the e-index that shares authorship of every group of authors equally and then sums across the publications of each author in the dataset, to obtain her index value.

Definition 4 *ϕ is e-index iff for all $c \in \mathcal{C}$, all $i \in N^P$,*

$$\phi_i(c) = \sum_{S \subseteq N^P : i \in S} \frac{c(S)}{|S|}$$

1. Results

The following theorem is our main result. It shows that if one accepts anonymity, monotonicity and efficiency as the requisite properties that every index comparing authors must satisfy, then the only option available is e-index.

Theorem 1 *An index satisfies anonymity, monotonicity and efficiency if and only if, it is e-index.*[‡]

Proof: It can easily be checked that e-index satisfies anonymity, monotonicity and efficiency. The following proof establishes that no index other than e-index, satisfies anonymity, monotonicity and efficiency. The proof of necessity is accomplished by the technique of induction in the following manner.

Consider any citation function $c \in \mathcal{C}$ such that $c(S) = 0$ for all $S \in \rho(N^P)$. By anonymity, for all $i \neq j$, $\phi_i(c) = \phi_j(c)$, and so, efficiency implies that $\phi_i(c) = 0$ for all $i \in N^P$. Define a function $\kappa : \mathcal{C} \mapsto \mathbb{N}^{\mathbb{P}}$ such that $\kappa(c) := |\rho(N^P)| - |\{S \in \rho(N^P) : c(S) = 0\}| + 1$ for each possible citation function c. Fix a $k \in \{1, \ldots, |\rho(N^P)|\}$ and suppose that for all $c \in \mathcal{C}$ such that $\kappa(c) \leq k$, $\phi_i(c) = \sum_{S \subseteq N^P : i \in S} \frac{c(S)}{|S|}$ for all $i \in N^P$. In the following paragraphs we show how our supposition (or, the induction hypothesis); implies that for all $c \in \mathcal{C}$ such that $\kappa(c) = k+1$, $\phi_i(c) = \sum_{S \subseteq N^P : i \in S} \frac{c(S)}{|S|}$ for all i.

Fix a $c \in \mathcal{C}$ such that $\kappa(c) = k+1$ and define the set $\hat{N^P} := \{i \in N^P : \exists\ S \subseteq N^P \setminus \{i\}$ such that $S \neq \emptyset$ and $c(S) \neq 0\}$. Therefore, $N^P \setminus \hat{N^P}$ is the set of agents i, for the given citation function c, such that the academic contribution of any group of authors not containing i is 0. We call these agents in $N^P \setminus \hat{N^P}$ as the *star* authors in c. If all authors are star authors, that is $\hat{N^P} = \emptyset$, then $c(S) = 0$ whenever $S \neq N^P$; and so, no group of authors other than the grand coalition can produce a positive academic contribution. Therefore, it easily follows that, by anonymity, $\phi_i(c) = \phi_j(c)$ for all $i \neq j \in N^P$. And so, efficiency implies that for all $i \in N^P$, $\phi_i(c) = \frac{c(N^P)}{n^P} = \sum_{S \subseteq N^P : i \in S} \frac{c(S)}{|S|}$.

Consider the other possibility where all authors are not stars, that is, $\hat{N^P} \neq \emptyset$. For any $i \in \hat{N^P}$, construct a citation function $c_i \in \mathcal{C}$ such that; $c_i(S^i) = 0 \neq c(S^i)$ for some $S^i \in \rho(N^P)$ with $i \notin S^i$, and $c_i(S) = c(S)$ for all $S \neq S^i$. By construction of $\hat{N^P}$, citation function c_i is well defined. Further, $\kappa(c_i) = k$ and so, by induction hypothesis, $\phi_i(c_i) = \sum_{S \subseteq N^P : i \in S} \frac{c_i(S)}{|S|} = \sum_{S \subseteq N^P : i \in S} \frac{c(S)}{|S|}$ for all $i \in \hat{N^P}$. Now, if $\hat{N^P} = N^P$ implying that there are no star authors, then the result follows trivially. If $\hat{N^P} \subset N^P$, that is, the set of star authors $N^P \setminus \hat{N^P}$ is non-empty; then by construction, for any $S \in \rho(N^P)$, $c(S) > 0$ only if $[N^P \setminus \hat{N^P}] \subseteq S$. By applying anonymity for all bijections $\pi : N^P \mapsto N^P$ such that $\pi(i) = i$ for all $i \in \hat{N^P}$, and $\pi(N^P \setminus \hat{N^P}) = N^P \setminus \hat{N^P}$, we get that $\phi_i(c) = \phi_j(c)$ for all $i \neq j \in N^P \setminus \hat{N^P}$. Therefore, by efficiency we get that; for any $i \in N^P \setminus \hat{N^P}$, we get that

$$
\begin{aligned}
&\phi_i(c) \\
&= \frac{1}{n^P - |\hat{N^P}|} \left\{ \sum_{S \in \rho(N^P)} c(S) - \sum_{i \in \hat{N^P}} \sum_{S \subseteq N^P : i \in S} \frac{c(S)}{|S|} \right\} \\
&= \frac{1}{n^P - |\hat{N^P}|} \sum_{i \in N^P \setminus \hat{N^P}} \sum_{S \subseteq N^P : i \in S} \frac{c(S)}{|S|}
\end{aligned}
$$

By construction, $N^P \setminus \hat{N^P} \neq \emptyset$ implies that for all $i \neq j \in N^P \setminus \hat{N^P}$, $\{S \subseteq N^P : i \in S, c(S) > 0\} = \{S \subseteq N^P : j \in S, c(S) > 0\}$ and so, the right hand side of the equation above collapses to $\sum_{S \subseteq N^P : i \in S} \frac{c(S)}{|S|}$. Thus, the result follows.

In the following example, we undertake a simple data exercise to show how our e-index compares with h-index.

Example 1 We focus on four highly accomplished physicists of same professional designation, specializing in theoretical condensed matter physics; Professors S. Das Sarma, J. Fabian, C. Nayak, R. I. Greene. To compare them using our e-index, we start with dataset consisting of the *top twenty* best cited papers of each of these authors, and then, compute the e-index values for each of them.[§] We find that e-indices for these authors are as follows: S Das Sarma, $e = 8211$; J. Fabian, $e = 3628$; C. Nayak, $e = 2379$; R. I. Greene, $e = 1630$. In contrast, the corresponding h-index values for these authors are; 105, 41, 52 and 90, respectively. Note that the values returned by both indices are somewhat different. This is not surprising because; while the h-index ignores a lot of information available in data by neglecting the citation numbers of highly ranked papers, our e-index processes all of the information available in citation data, and weights them suitably to generate index values.

Therefore, the best way to compare these two indices is to look at the rankings that they imply among the set of authors that we have focussed on. This comparison is presented in the following table.

Author	e	h
S. DAS SARMA	1	1
J. FLABIAN	2	4
C. NAYAK	3	3
R. I. GREENE	4	2

As can be seen from the table, both indices disagree in their ranking for the second and fourth positions. The h-index ranks R. I. Greene above J. Fabian, while the e-index ranks them in an exact opposite manner. This is presumably because; in comparison to Professor Flabian who started publishing from 1993, Professor Greene has been publishing since 1960, and so, has had greater opportunity to build larger network of coauthors. Our e index, in contrast, normalizes for the coauthorship citations, and delivers a more balanced comparison between these two authors.

Thus, in this example, we find that e index delivers substantially different rankings in comparison to that from h-index (implying a rank correlation of only 0.2), even when we focus on comparing *only* four star theoretical condensed matter physicists. We feel that as one increases the number of authors, the rank correlation coefficient would go down further. Also, we believe that our e-index would continue to deliver such balanced rankings when used to compare academics from other disciplines.

Discussion

In this paper we treat authorship as a cardinal variable. That is, we believe that differences in index numbers across authors, have intuitive meaning. This allows us to conceptualize proration of academic credit for joint papers. In contrast,

‡ The result continues to hold if we focus *only* on indices that return non-negative values.

§ We use the data available on the open access Google citation page on 17th April, 2017 one may argue that academic authorship is too abstract a notion to allow for any sort of proration. However, as noted in Liebowitz(3), without sufficient proration to identify individual contribution in joint papers, there could be excessive coauthorship.[¶] Therefore, we believe that it is better for any discipline to treat authorship as a cardinal variable that can be prorated.

Further, our analysis ranks authors only on basis of the citation number (or some other relevant measure of contribution) of their publications, accommodatingly suitably for the fact whenever they are multiple author projects. However, for many such joint author papers, the order in which names of coauthors appear in print; often is *publicly* believed to be the agreed order of contribution to the paper. This is true, particularly, for papers in basic sciences like Physics. It may be argued that this additional information should affect the method of proration of authorship in joint author papers, and thus, affect any index that ranks authors.

We do not disagree with the merit of this argument. Ideally, a method of ranking authors should depend on citation numbers as well as the order in which names of authors are displayed. However, this is a complicated problem as one would have combine a set of cardinal numbers with a set of ordinal ranks, to obtain an index or ranking method. Further, there are many disciplines like social sciences where journals are publicly known to report the names of authors in an alphabetical manner, thereby leading to a situation where the aforementioned information simply does not exist. In fact this is the case for some top particle physics journals too. More importantly, Hirsch (1) himself ignores this issue of accounting for the displayed order of author contributions altogether. Therefore, considering the popularity that h-index has enjoyed among administrators over years, we feel that this issue is not a major hindrance to practical use of our index, at least as long as a better index accomplishing the aforementioned exercise is not made available to the academic community.

Conclusion

In this paper, we characterize a new method of ranking authors for a given dataset of publications. Our characterization involves three basic properties: anonymity, monotonicity and efficiency. We find the only index that satisfies these properties is the intuitive e-index.

This paper identifies several questions that the academic community must arrive at a consensus, both for the sake of objectivity, as well as adoption of fair authorship practices in academia. Two such questions would be (i) the issue of accommodating publicly available order of author contributions in evaluating academic credit, and (ii) ranking authors after combining academic accomplishment in different dimensions, like teaching, proposal application, publication etc. We leave such questions for future research.

1. Hirsch JE (2005) An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America* 102:16569–16572.
2. Stephan PE (1996) The economics of science. *Journal of Economic literature* 34:1199–1235.
3. Liebowitz SJ (2014) Willful blindness: the inefficient reward structure in academic research. *Economic Inquiry* 52:1267–1283.
4. Tarnow E (2002) Coauthorship in physics. *Science and engineering ethics* 8:175–190.
5. Berk RN (1989) Irresponsible coauthorship. *American Journal of Roentgenology* 152:719–720.

[¶] Liebowitz (3) demonstrates by a statistical exercise; how increase in coauthorship over the years can be better explained by incomplete proration than by increased specialization.