

# Intellectual Influence: quality versus quantity

## Citation for published version (APA):

Koczy, L., Nichifor, A., & Strobel, M. (2010). *Intellectual Influence: quality versus quantity*. METEOR, Maastricht University School of Business and Economics. METEOR Research Memorandum No. 029 <https://doi.org/10.26481/umamet.2010029>

## Document status and date:

Published: 01/01/2010

## DOI:

[10.26481/umamet.2010029](https://doi.org/10.26481/umamet.2010029)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

László Á Kóczy, Alexandru Nichifor,  
Martin Strobel

**Intellectual Influence:  
Quality versus Quantity**

RM/10/029  
(RM/09/016 -revised-)

**METEOR**

Maastricht University School of Business and Economics  
Maastricht Research School of Economics  
of Technology and Organization

P.O. Box 616  
NL - 6200 MD Maastricht  
The Netherlands

# Intellectual Influence: Quality versus Quantity <sup>\*</sup>

László Á. Kóczy<sup>†</sup>    Alexandru Nichifor<sup>‡</sup>    Martin Strobel<sup>§</sup>

## Abstract

To take development and budgeting decisions for research activities the officials in charge need to constantly evaluate and assess the quality of research. Over the years a handful of scoring methods for academic journals have been proposed. Discussing the most prominent methods (de facto standards) we show that they cannot distinguish quality from quantity at article level and that they are inherently biased against journals publishing more articles. If we consider the length of a journal by the number of pages or characters, then all methods are biased against lengthier journals. The systematic bias we find is analytically tractable and implies that the methods are manipulable. We show that the strategies for successful manipulation are relatively easy to infer and implement. The implications of our findings extend beyond the evaluation of academic research, to related settings like the ranking of web domains. Non-manipulable methods for measuring intellectual influence exist.

**JEL:** A1, C8, Y1.

**Keywords:** scoring methods bias, ranking rules bias, impact factor, invariant method, LP method, invariance to article-splitting, quality and quantity in ranking academic journals

---

<sup>\*</sup>A previous version of this paper was circulated under the title “Article Length Bias in Journal Rankings”. We thank Benjamin Edelman, Çağatay Kayı, Bettina Klaus, Scott Duke Kominers, Itay Fainmesser, Lavinia Gotovan and William Thomson for helpful discussions and comments. Kóczy thanks funding from OTKA under grant NF-72610; Nichifor thanks: NWO for its support under grant VIDI-452-06-013, METEOR for its support via the ITG, and Harvard Business School for the hospitality he enjoyed while working on this paper; Strobel thanks funding from METEOR.

<sup>†</sup>Keleti Faculty of Business and Management, Óbuda University, Tavaszmező u. 15-17., H-1084 Budapest. Email: [koczy.laszlo@kgk.uni-obuda.hu](mailto:koczy.laszlo@kgk.uni-obuda.hu).

<sup>‡</sup>*Corresponding author:* Department of Economics, Maastricht University, P.O.Box 616, NL-6200MD Maastricht. Email: [a.nichifor@maastrichtuniversity.nl](mailto:a.nichifor@maastrichtuniversity.nl).

<sup>§</sup>Department of Economics, Maastricht University, P.O.Box 616, NL-6200MD Maastricht. Emails: [m.strobel@maastrichtuniversity.nl](mailto:m.strobel@maastrichtuniversity.nl).

“We are a group of economists including current and former editors of *Econometrica*, the *American Economic Review*, and the *Review of Economic Studies*, several past Presidents of the Econometric Society, the incoming President of the American Economic Association [...]. We [...] recognize that the assessment of the significance of a publication depends heavily on the reputation of the journal in which it is published. [...] *Journal of Economic Theory*’s [...] reputation is supported by numerical measures, such as the impact factor.”

– Theoretical Economics, Open Letter to Promotion Committees <sup>1</sup>

“The perception that the Research Assessment Exercise (RAE) rewards publication in journals with high impact factors is affecting decisions made by authors about where to publish. We urge HEFCE to remind RAE panels that they are obliged to assess the quality of the content of individual articles, not the reputation of the journal in which they are published.”

– The Research Assessment Exercise<sup>2</sup>

## 1 Introduction

The last decades saw an explosion in the number of academic journals. Researchers find it more and more difficult to keep up with the growing literature even in narrow fields. Libraries face higher subscription fees and must allocate budgets in an efficient way. Promotion decisions are often taken based on researchers’ publications. National organizations for scientific research steer the course of science by funding proposals based on their potential and on the publication record of the applicants. However, the quality of the publications, approximated by the containing journals’ quality indicator, is becoming increasingly difficult to evaluate and compare and there is a growing interest in finding measures, both cardinal and ordinal, that would allow for an objective assessment. To this end, various scoring methods and ranking rules have been devised. The former capture the cardinal aspect, by giving scores to each journal, and the latter capture the ordinal aspect, by establishing an order of preference among the journals.

Loosely speaking, a scoring (ranking) problem can be thought of as a social choice problem where a social welfare function is used to obtain total

---

<sup>1</sup>Available online at: <http://www.econtheory.org/ProComLt.pdf>

<sup>2</sup>House of Commons - Science and Technology - Tenth Report. 2004-07-07. Available online at:  
<http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/39912.htm>

preorders on the set of alternatives, with the additional requirement that the set of agents and the set of alternatives coincide.<sup>3</sup> That is, journals are asked to express their opinions about each other and themselves. Citations made by a journal are considered to be votes about the importance of the destination journal and a scoring method is used to aggregate the information and determine a score for each journal. Each scoring method induces a ranking rule.

In practice, the predominant scoring methods used for the measurement of intellectual influence are the impact factor (Garfield, 1955), the LP method (Liebowitz and Palmer, 1984; Laband and Piette, 1994), and the invariant method (Pinski and Narin, 1976). The last two methods generated many variations of great practical importance. The best known is the PageRank algorithm (Brin and Page, 1998) which is at the core of how search engines rank web pages. Another variation is known as the DeGroot (1974) model, which is used in models of learning in social and communication networks (Golub and Jackson, 2010), physics, and computer science (Sobel, 2000).

Despite their extensive usage, the above scoring methods have only been intuitively motivated, if at all. We are familiar with two notable exceptions which present characterizations. Given the invariant method, Palacios-Huerta and Volij (2004) find a set of cardinal properties that fully characterize it. Given the PageRank algorithm, Altman and Tennenholtz (2005) find a set of ordinal properties that fully characterize it.

This paper complements the efforts made towards a better understanding of scoring methods from a normative perspective: while we do not derive a characterization of any scoring method, we formalize a sensible property, and then check if the currently most used methods satisfy it. We call our property *invariance to article-splitting*. Intuitively, this property requires a scoring method to assign the same valuation to any two journals having similar citation patterns but publishing a different number of articles. Another interpretation for our property is to say that everything else equal, publishing a number of results in distinct papers or together as a survey or a book should make no difference, that is, the scoring method should assign the same value for both the former and the latter. We discuss the case in which the scoring methods take as input the number of pages or characters rather than the number of articles (see Mirrlees, Neary, and Tirole (2003) for a good survey of works that opt for this approach), and we also interpret our property in the context of ordering web domains. Palacios-Huerta

---

<sup>3</sup>Note that despite the fact that the models are closely related, they lead to very different results: for example, Shoham and Leyton-Brown (2009, Proposition 9.5.1) show that Arrow's impossibility result holds *exclusively* in a social choice setting.

and Volij (2004) were the first to introduce invariance properties for scoring methods. However, their properties are not directly related to our invariance to article-splitting. In particular, with Theorem 1, Step 3, we show that the invariant method, which is characterized by their properties does not satisfy invariance to article-splitting.

We find that the impact factor, LP method, and the invariant method do not satisfy invariance to article-splitting. Moreover, with Theorem 1 we find a systematic bias common to all methods favoring journals with fewer articles (or pages, or characters). While invariance to article-splitting is a cardinal property, we give Example 2 where the valuations are sufficiently biased to affect the ordinal ranking.

For the classification of academic journals, our results have several implications. First, whether we control for sheer size using the number of pages or the number of articles has a profound effect on the classification results. Second, we find that quality and quantity are indistinguishable at article level. Third, it is a direct consequence of our results that the scoring methods presented are manipulable, and we suggest how in principle publishers or editors could artificially boost the scores of their journals. Finally, we discuss how our findings can reach beyond the journal setting, to settings of great practical importance like the classification of web domains.

## 2 Scoring Methods

Let  $J = \{1, \dots, n\}$  denote a non-empty finite set of journals. For each  $i, j \in J$ ,  $c_{ij}$  represents the number of citations to journal  $i$  by journal  $j$ , that is, the number of references *made* by journal  $j$  to journal  $i$ . Let us consider a  $n \times n$  nonnegative *citation matrix*  $C = \{c_{ij}\}_{i,j \in J}$ . Let  $c_j = \sum_{i \in J} c_{ij}$  denote the total number of citations made by  $j$  and let  $D_C$  denote the diagonal matrix with the elements in  $\{c_j\}_{j \in J}$  along the diagonal. Let the entries of the nonnegative vector  $a$  denote the number of articles in each journal and let  $A$  be the diagonal matrix with the elements in  $\{a_j\}_{j \in J}$  along the diagonal. For each vector  $z \in \mathbb{R}^n$ , we denote the 1-norm of  $z$  by  $\|z\| = \sum_{i=1}^n |z_i|$ .

**Definition 1.** A scoring problem is a triple  $(J, a, C)$  consisting of a finite set of journals  $J$ , a vector  $a \in \mathbb{N}^n$  containing the number of published articles and a citation matrix  $C = \{c_{ij}\}_{i,j \in J}$ .

Let  $\mathcal{S}$  denote the set of scoring problems. The score for each journal in  $J$  is given by the transposed *valuations vector*  $v^T = (v_1, v_2, \dots, v_n)$ , where  $v_i$  is interpreted as the value of a representative article in journal  $i$ .

**Definition 2.** A scoring method  $\phi$  maps a scoring problem  $(J, a, C) \in \mathcal{S}$  to a valuations vector  $v \in \mathbb{R}^n$ .

A scoring method induces a weak ordering of the journals via the *ranking rule*  $i \succcurlyeq j$  if and only if for all  $i, j \in J$ ,  $v_i \geq v_j$ . Ties, i.e.,  $i \succcurlyeq j$  and  $j \succcurlyeq i$ , are allowed, but only occur if  $v_i = v_j$ .

The impact factor (Garfield, 1955) considers all citations received by an article to have the same value and measures the direct influence that a typical article in journal  $i$  has on all journals.

**Definition 3.** The impact factor gives valuations according to the unique vector  $v$  that solves

$$A^{-1}Ce = v \tag{1}$$

where  $e$  is a vector of ones of dimension  $J$  and the matrices  $A$  and  $C$  contain data for a two-year period.

The next two scoring methods assign different values to citations received based on the importance of the journal that made the citation. The importance of each journal is established endogenously and simultaneously for all journals in each method, using convergent iterative procedures. Roughly speaking, the LP method (Liebowitz and Palmer, 1984) gives valuations that reflect the influence that a typical article in journal  $i$  has on journal  $j$ . The invariant method (Pinski and Narin, 1976) is a modification of the LP method such that the valuations given are also weighted by the reference intensity (i.e., the average number of citations made by a typical article in  $j$ ).

For a formal presentation, we need an additional assumption and some extra notation. We require the citation matrix  $C$  to be *primitive*<sup>4</sup>: there should be no partition of the set of journals  $J$  in two sets  $J_1$  and  $J_2$  such that *i*) there are no inter partition citations *or ii*) all inter partition citations are unidirectional, say from journals in  $J_2$  to journals in  $J_1$ , and we should have at least one self-citing journal. This is a very natural and plausible assumption for classifying journals within the same field, and from a technical perspective it ensures that the following scoring methods are well defined.<sup>5</sup>

**Definition 4.** The LP method gives valuations according to the unique vector  $v$  that solves

$$\frac{A^{-1}Cv}{\|A^{-1}Cv\|} = v. \tag{2}$$

---

<sup>4</sup>It is well known that a sufficient condition for a matrix to be primitive is to be nonnegative and irreducible with at least a positive element on the main diagonal.

<sup>5</sup>Under this assumption the iterative procedures defining the following two methods are known to converge. We give directly the steady-state equations.

**Definition 5.** *The invariant method gives valuations according to the unique vector  $v$  that solves*

$$A^{-1}CD_C^{-1}Av = v. \quad (3)$$

### 3 Article-Splitting and Manipulability

In this section we introduce a basic desirable property of a scoring method: invariance to article-splitting. When we refer to an article’s citations, we consider both citations made (that we interpret as an article’s “consumption”) and received (that we interpret as an article’s “production” which was consumed by other articles). We require that if a paper is split into  $k$  sub-papers, then the citations are split among them. Conversely, if a set of  $k$  articles with no overlapping citations and published in the same journal are merged into a single longer paper, then the resulting publication collects all citations. Our invariance property requires that having the split ex-post or ex-ante should make no difference. That is, splitting the citations between the sections of a paper, or publishing sub-papers, leave the journals’ valuation unchanged. The above interpretation implicitly assumes that there are two separable layers of citations: *background citations*, that are used merely to position and relate an article to the relevant literature, and *active citations*, that are made (received) by an article for actively consuming (producing) knowledge. Article-splitting is then natural to assume for active citations, but less so for background citations.

Formally, let  $\lambda_j \in \mathbb{R}$ ,  $\lambda_j > 1$ , be a *splitting factor* and consider two ranking problems  $\{(J, a, C), (J, a', C)\} \subseteq \mathcal{S}$  where for some journal  $j \in J$ ,  $a'_j = \lambda_j a_j$  and for all other journals  $i \neq j$ ,  $a'_i = a_i$ . Let  $S = (J, a, C)$  and  $S' = (J, a', C)$ . For the two problems  $S, S' \in \mathcal{S}$  defined as above,  $S'$  is an *article-split modification* of  $S$ .

**Definition 6.** *A scoring method  $\phi$  is invariant to article-splitting if for any two problems  $S, S' \in \mathcal{S}$  such that  $S'$  is an article-split modification of  $S$ ,  $\phi_j(S) = \phi_j(S')$ .*

Observe that the citation matrix is not affected: the number of citations does not change, they are only distributed among more papers. The property thus relates scoring problems implicating journals having the same citation patterns but publishing different numbers of articles. Note that in each of the scoring methods introduced in the previous section, a journals’ opinion is given by the aggregated opinions of the articles contained in it. Hence, each component of the valuation vector produced by each method can be interpreted as a measure of the intellectual influence per typical article published



in a journal. However, there is a considerable heterogeneity in the length of a typical article even for journals within the same field. Some journals differ also in terms of page size. The applied studies of (Kalaitzidakis, Mamuneas, and Stengos, 2003; Combes and Linnemer, 2003; Coupé, 2003) correct for this by computing scores per page or per character, not per article. Reinterpreting our invariance to article-splitting in this context, the property requires that for journals with similar citation patterns and the same number of articles, the number of pages per article (or per characters) should not matter. The following example gives a more vivid representation of the situation described.

*Example 1. Invariance to article-splitting when scores are computed per page, or per character.*

Assume that the editorial board of a journal accepts a number of articles. Consider the following scenarios: *i*) the articles are published as such, or *ii*) for some of the articles the authors are requested to shorten their length, by relegating inessential details to the “web appendix”. The web appendix is available on-line but it is not part of the printed journal, whose length is taken as input by the scoring methods. Observe that scenario *ii*) leads to a shorter journal in terms of pages or characters than scenario *i*), and that the citations made (received) by an article are invariant between scenarios, as typically there are no citations made (received) on inessential details. Invariance to article-splitting requires the score of the journal to be the same in both scenarios.

Next, we define manipulability and we formalize a systematic deviation from invariance to article-splitting.

**Definition 7.** *A scoring method  $\phi$  is manipulable if a journal can increase its valuation unilaterally.*

**Definition 8.** *A scoring method  $\phi$  is biased against article-splitting if for any two problems  $S, S' \in \mathcal{S}$  such that  $S'$  is an article-split modification of  $S$ ,  $\phi_j(S) > \phi_j(S')$ .*

*Remark 1.* If a method is biased against article-splitting then it is manipulable.

## 4 Results

In the following, we show that none of the listed scoring methods is invariant to article-splitting.

**Theorem 1.** *The impact factor, the LP method and the invariant method are biased against article-splitting.*

*Proof.* Let  $S, S' \in \mathcal{S}$  be such that  $S'$  is an article-split modification of  $S$ . We now proceed in several (independent) steps.

*Step 1:* The impact factor is biased against article-splitting.

Observe that  $v'_j = \frac{a_j}{a'_j} v_j = \frac{1}{\lambda_j} v_j$ . Hence,  $v'_j < v_j$ , and for all  $i \neq j$ ,  $v'_i = v_i$ . The result is independent of the fact that the impact factor is calculated for a period of two years.

*Step 2:* The LP method is biased against article-splitting.

Adapting a technique introduced by Roy, Saberi, and Wan (2008), we show that an increase in the number of articles of a journal decreases its valuation. Let  $\Gamma = A^{-1}C$  and  $\Gamma' = A'^{-1}C$ . Then, for  $S$  and  $S'$ , the LP method gives valuations according to vectors  $v$  and  $v'$  that solve the following equations:

$$\Gamma v = \|\Gamma v\| v, \quad (4)$$

$$\Gamma' v' = \|\Gamma' v'\| v'. \quad (5)$$

Equations (4) and (5) are algebraic eigenvalue problems:  $\varrho(\Gamma) = \|\Gamma v\|$  is the spectral radius of  $\Gamma$  and  $v$  the eigenvector associated with  $\varrho(\Gamma)$ , and  $\varrho(\Gamma') = \|\Gamma' v'\|$ ,  $\Gamma'$ ,  $v'$  are similar. Since the matrix  $\Gamma$  is primitive,  $\Gamma'$  is also primitive and (5) is well defined. Since for all  $i \in J$ ,  $a'_i \geq a_i$ ,  $\Gamma'$  is weakly smaller in every entry than  $\Gamma$ . Therefore, there exists  $\delta > 0$  such that:

$$\varrho(\Gamma') = \varrho(\Gamma) - \delta. \quad (6)$$

We scale  $v'$  such that  $v'_j = v_j$  and rewrite  $v'^T$  as  $\bar{v}^T = v^T - x^T = [v_1 - x_1, \dots, v_j - x_j, \dots, v_n - x_n]$  where  $x \in \mathbb{R}^n$  such that  $x_j = 0$ . By (5),  $\Gamma' \bar{v} = \varrho(\Gamma') \bar{v}$ . Replacing  $\bar{v}$ ,  $\varrho(\Gamma')$  and using (4), we have:

$$\Gamma' v - \Gamma' x = \varrho(\Gamma') v - \varrho(\Gamma') x = \varrho(\Gamma) v - \delta v - \varrho(\Gamma') x = \Gamma v - \delta v - \varrho(\Gamma') x \quad (7)$$

Let  $v_{-j}, x_{-j} \in \mathbb{R}^{|J-\{j\}|}$ , where  $v_i, x_i > 0$ , be the valuation vectors except for journal  $j$ , and let  $\Gamma_{-j}$  and  $\Gamma'_{-j}$  be the matrices  $\Gamma$  and  $\Gamma'$  where we removed row and column  $j$ . Note that  $\Gamma_{-j} = \Gamma'_{-j}$ . Dropping the  $j$ 'th equation from the system of Equations 7, we obtain<sup>6</sup>:

$$\Gamma_{-j} x_{-j} = \delta v_{-j} + \varrho(\Gamma') x_{-j}. \quad (8)$$

---

<sup>6</sup>For clarity, we detail this step in Appendix A.

Rearranging terms:

$$(\varrho(\Gamma')I - \Gamma_{-j})x_{-j} = -\delta v_{-j}. \quad (9)$$

Define  $N = \frac{\Gamma_{-j}}{\varrho(\Gamma')}$ , and  $M = (I - N)$ . Since  $\Gamma_{-j}$  and  $\varrho(\Gamma')$  are nonnegative,  $N$  is entrywise nonnegative, i.e.,  $N \geq 0$ . Marcus and Minc (1975) show that the spectral radius of a primitive matrix is greater than the spectral radius of any of its submatrices. Hence,  $\varrho(\Gamma') > \varrho(\Gamma'_{-j}) = \varrho(\Gamma_{-j})$ . Thus, the moduli of the eigenvalues of  $N < 1$ , and consequently  $\lim_{t \rightarrow \infty} N^t = 0$ . But:

$$I - N^{k+1} = M(I + N + N^2 + \dots + N^k). \quad (10)$$

Letting  $k \rightarrow \infty$ ,  $I = M \sum_{k=0}^{\infty} N^k$ . Premultiplying by  $M^{-1}$ , we have  $M^{-1} = \sum_{k=0}^{\infty} N^k$ . Since  $N \geq 0$ ,  $M^{-1} \geq 0$ . Observing that in (9) the vector  $v_{-j}$  is positive,  $x_{-j}$  has to be negative. Hence,  $\bar{v}_j = v_j$  and for all  $i \neq j$ ,  $\bar{v}_i > v_i$ . Rescaling  $\bar{v}$  to  $v'$ , we have  $\frac{v'_j}{v'_i} < \frac{v_j}{v_i}$ . Since  $v'_j = v_j$ , for all  $i \neq j$ ,  $v'_i > v_i$ .

*Step 3:* The invariant method is biased against article-splitting.

Observe that  $v'$ , defined as  $v'_j = \frac{1}{\lambda_j} v_j$  and  $v'_i = v_i$  for  $i \neq j$ , is the solution of:

$$A'^{-1}CD_C^{-1}A'v' = v'. \quad (11)$$

In order to see this, premultiply (3) by  $A$ , and (11) by  $A'$ . Then, note that  $A'v' = Av$ . Finally observe that as  $\lambda_j > 1$ ,  $v'_j < v_j$ , while the valuations of other journals have not changed.  $\square$

Note that for the impact factor and for the invariant method the valuation of a journal  $j$  whose articles are split into  $\lambda_j$  articles decreases by a factor of  $\frac{1}{\lambda_j}$ . For an appropriate choice of  $\lambda_j$ , the decrease can be arbitrarily low. In particular, it can be lower than the valuation of the journal ranked next, thus changing also the ranking of the journals. Similarly, an increase in the number of articles of journal  $j$  decreases its relative valuation given by the LP method, which may also affect journal  $j$ 's ranking.

The following example shows that the bias against article-splitting of the scoring methods above may also induce changes in the ranking of the journals:

*Example 2.* Let  $J = \{j_1, j_2, j_3\}$ ,  $a = (2, 2, 3)$ ,  $a' = (4, 2, 3)$  and define  $C$  as:

$$C = \begin{pmatrix} 12 & 8 & 4 \\ 6 & 10 & 2 \\ 3 & 3 & 9 \end{pmatrix}$$

Let  $S = (J, a, C)$  and  $S' = (J, a', C)$  be two ranking problems and note that the only difference between  $S$  and  $S'$  is that  $a'_1 = \lambda_1 a_1$ , with  $\lambda_1 = 2$ . For each problem  $S$  and  $S'$ , the following table presents the normalized vector of valuations (so that the entries add up to 1) produced by each scoring method.

	$\phi_{IF}(S)$	$\phi_{IF}(S')$	$\phi_{IM}(S)$	$\phi_{IM}(S')$	$\phi_{LP}(S)$	$\phi_{LP}(S')$
$j_1$	0.46	0.30	0.48	0.31	0.50	0.29
$j_2$	0.35	0.45	0.33	0.44	0.37	0.51
$j_3$	0.19	0.25	0.19	0.25	0.13	0.20

Table 1: Article-splitting bias in scoring methods inducing changes in the ranking of the journals.

Note that for  $S$ , all scoring methods induce the ranking  $j_1 \succ j_2 \succ j_3$ , while for  $S'$ , all scoring methods induce the ranking  $j_2 \succ j_1 \succ j_3$ .

Theorem 1 clearly indicates how to *manipulate* scoring methods. Should the methods take as input the number of articles, then publishers or editors might opt to implement a policy of publishing a small number of articles. Alternatively, if the number of pages or characters is taken as input, preference can be given to publishing briefer communications. Recall Example 1. In principle, editors requests of discarding inessential details by relegating them to the web appendix are desirable, as we don't want to waste valuable resources such as journal pages and reader's time for irrelevant details. However, the same requests might easily be abused to game the scoring methods. Assume that the editors ask for a significant shortening of the paper, and now even relevant details, as opposed to only irrelevant ones, are relegated to the web appendix. This is unlikely to affect the citations made/received by the article.<sup>7</sup> As there is definitely a gray zone in what should be printed and what is inessential, such manipulation is easy, its detection is hard, and by Example 2 it is clear that in principle a manipulating editor is even able to quantify the effectiveness of his action.

Interestingly, essentially the same methods are used by search engines for obtaining the ranking of web pages. The following example aims to clarify this analogy.

*Example 3. The ranking of web domains.*

Professor X acquires the domain "www.ProfessorX.net" and aims to make the following types of information available online: research, teaching, and

---

<sup>7</sup>Typically there are no citations made in discarded parts. In terms of received citations, when needed, readers will acquire the full details from the web (either from the author's page providing a full copy of the article or from the journal's web appendix), but always cite the article in the journal.

contact details. There are two natural options: *i*) to put all available information on one page, each type in a separate section, or *ii*) to create one distinct web page for each type. Note that ProfessorX is only interested in the ranking of his domain and that the links that his domain typically makes/receives are invariant under *i*), *ii*). In this case, search engines like Google or Bing use essentially the same methods as the one used for the ranking of academic journals: their algorithm relies on the LP-method where the left eigenvector is computed, i.e., the weight for each page is given by the components of vector  $v$  that solves  $\frac{vA^{-1}C^T}{\|vA^{-1}C^T\|} = v$ , where each entry  $c_{ij}$  in the transposed matrix  $C^T$  can be thought of as the number of links made by domain  $i$  to domain  $j$  out of the total links made by  $i$ , and  $a_i$  as the number of web pages per domain.<sup>8</sup> By Theorem 1 it is a dominant strategy for ProfessorX to aggregate all information on a single page.<sup>9</sup>

More generally, owners of web domains have incentives to manipulate for economic profit: a higher placement in Google results drives more internet traffic which in turn yields higher sales or advertisement revenues. Reinterpreting Theorem 1 in this context reveals that a domain owner can improve the score for his domain by jamming all his data on a single omniscient sheet.

## 5 Conclusion

Applied studies have long suggested that the scoring methods used for ranking academic journals may not be immune to various biases. We derive analytically that one such bias exists. For journals with similar citation patterns, the journals publishing fewer articles are privileged to the ones publishing more articles. If we account for the length of a journal based on the number of pages or characters, then the current scoring methods give extra credit to shorter more formal journals and punish those that make an effort of keeping good English in their articles. Consequently, if the journals have the same number of articles, more credit is given to the ones publishing shorter articles. Thus, the currently most used scoring methods share a common drawback: they cannot distinguish quality from quantity at article level. Furthermore, observe that for a journal with relatively numerous articles but few pages, measuring its influence by taking the number of pages or articles as input will make a crucial difference. One must therefore use

---

<sup>8</sup>The exact algorithms used by search engines are both a moving target and a black box, but the characteristics to which we make the analogy to here are known to be relatively stable (see for example Langville and Meyer (2006)).

<sup>9</sup>The same technique used in Step 2 of Theorem 1 can be easily adapted for the left eigenvector and leads to the same qualitative conclusions.

and interpret the valuations and the induced rankings with care when estimating the *quality* of journals. Hence, our findings support the suggestion in the Research Assessment Exercise to be cautious about the current de facto practices in economics emphasized by the Open Letter to Promotion Committees in Theoretical Economics, and call for putting the traditional economic tools at work to obtain a better understanding of the properties satisfied by various “numerical measures.”

A consequence of our results is that the current methods for measuring influence are manipulable, and that the strategies that increase the payoffs are relatively easy to infer and implement. This is a concern for the evaluation of research and also in closely related settings like raking web domains.

It is an open question to find endogenous bounds for what variations can be allowed for the journals or users to input in the scoring methods without inducing changes in the associated rankings. It would be also interesting to find bounds for what is maximally achievable in terms of both score and rank by manipulating.

Finally, we note that not all rankings are subject to such bias. In the case of the H-index (Hirsch, 2004; Braun, Glänzel, and Schubert, 2006) it is easy to see that article-splitting has an ambiguous effect, while the tournament method (Kóczy and Strobel, 2007) is by definition invariant to article-splitting.

## A Appendix

Writing equality 7 in detail for the left most and right most terms of the equality, we obtain the following system of equations:

$$\begin{cases} \frac{1}{a_1}[(v_1c_{11} + \cdots + v_nc_{1n}) - (x_1c_{11} + \cdots + x_{j-1}c_{1j-1} + x_{j+1}c_{1j+1} + \cdots + x_nc_{1n})] = \frac{1}{a_1}(v_1c_{11} + \cdots + v_nc_{1n}) - \delta v_1 - \varrho(\Gamma')x_1 \\ \vdots \\ \frac{1}{a_j}[(v_1c_{j1} + \cdots + v_nc_{jn}) - (x_1c_{j1} + \cdots + x_{j-1}c_{jj-1} + x_{j+1}c_{jj+1} + \cdots + x_nc_{jn})] = \frac{1}{a_j}(v_1c_{j1} + \cdots + v_nc_{jn}) - \delta v_j - \varrho(\Gamma')x_j \\ \vdots \\ \frac{1}{a_n}[(v_1c_{n1} + \cdots + v_nc_{nn}) - (x_1c_{n1} + \cdots + x_{j-1}c_{nj-1} + x_{j+1}c_{nj+1} + \cdots + x_nc_{nn})] = \frac{1}{a_n}(v_1c_{n1} + \cdots + v_nc_{nn}) - \delta v_n - \varrho(\Gamma')x_n \end{cases}$$

After canceling terms and dropping the  $j$ 'th row from the system of equations above, we obtain:

13

$$\begin{cases} \frac{1}{a_1}(x_1c_{11} + \cdots + x_{j-1}c_{1j-1} + x_{j+1}c_{1j+1} + \cdots + x_nc_{1n}) = \delta v_1 + \varrho(\Gamma')x_1 \\ \vdots \\ \frac{1}{a_{j-1}}(x_1c_{j-11} + \cdots + x_{j-1}c_{j-1j-1} + x_{j+1}c_{j-1j+1} + \cdots + x_nc_{j-1n}) = \delta v_{j-1} + \varrho(\Gamma')x_{j-1} \\ \frac{1}{a_{j+1}}(x_1c_{j+11} + \cdots + x_{j-1}c_{j+1j-1} + x_{j+1}c_{j+1j+1} + \cdots + x_nc_{j+1n}) = \delta v_{j+1} + \varrho(\Gamma')x_{j+1} \\ \vdots \\ \frac{1}{a_n}(x_1c_{n1} + \cdots + x_{j-1}c_{nj-1} + x_{j+1}c_{nj+1} + \cdots + x_nc_{nn}) = \delta v_n + \varrho(\Gamma')x_n \end{cases}$$

Rewriting the above system of equations using vector and matrix notation yields equation 8.

## References

- ALTMAN, A., AND M. TENNENHOLTZ (2005): “Ranking Systems: The PageRank Axioms,” in *Proceedings of the 6th ACM conference on Electronic commerce (EC-05)*, pp. 1–8, New York. ACM, ACM Press.
- BRAUN, T., W. GLÄNZEL, AND A. SCHUBERT (2006): “A Hirsch-Type Index for Journals,” *Scientometrics*, 69(1), 169–173.
- BRIN, S., AND L. PAGE (1998): “The Anatomy of a Large-Scale Hypertextual Web Search Engine,” *Computer Networks and ISDN Systems*, 30(1–2), 107–117.
- COMBES, P.-P., AND L. LINNEMER (2003): “Where are the Economists Who Publish? Publication Concentration and Rankings in Europe Based on Cumulative Publications,” *Journal of the European Economic Association*, 1(6), 1250–1308.
- COUPÉ, T. (2003): “Revealed Performances: Worldwide Rankings of Economists and Economics Departments, 1990–2000,” *Journal of the European Economic Association*, 1(6), 1309–1345.
- DEGROOT, M. H. (1974): “Reaching a Consensus,” *Journal of the American Statistical Association*, 69(345), 118–121.
- GARFIELD, E. (1955): “Citation Indexes to Science: A New Dimension in Documentation through Association of Ideas,” *Science*, 122(3159), 108–111.
- GOLUB, B., AND M. O. JACKSON (2010): “Naive Learning in Social Networks and the Wisdom of Crowds,” *American Economic Journal: Microeconomics*, Forthcoming.
- HIRSCH, J. E. (2004): “An Index to Quantify an Individual’s Scientific Research Output,” *Proceedings of the National Academy of Sciences*, 102(46), 16569–16572.
- KALAITZIDAKIS, P., T. P. MAMUNEAS, AND T. STENGOS (2003): “Rankings of Academic Journals and Institutions in Economics,” *Journal of the European Economic Association*, 1(6), 1346–1366.
- KÓCZY, L. Á., AND M. STROBEL (2007): “Ranking Academic Journals by Tournament Methods,” Mimeo.



- LABAND, D. N., AND M. J. PIETTE (1994): “The Relative Impacts of Economics Journals: 1970-1990,” *Journal of Economic Literature*, 32, 640–666.
- LANGVILLE, A. N., AND C. D. MEYER (eds.) (2006): *Google’s PageRank and beyond: the science of search engine rankings*. Princeton University Press, Princeton, New Jersey.
- LIEBOWITZ, S. J., AND J. C. PALMER (1984): “Assessing the Relative Impacts of Economics Journals,” *Journal of Economic Literature*, 22(1), 77–88.
- MARCUS, M., AND M. MINC (1975): “On Two Theorems of Frobenius,” *Pacific Journal of Mathematics*, 60(2), 149–151.
- MIRPLEES, J. A., P. J. NEARY, AND J. TIROLE (2003): “Evaluating Economics Research in Europe: an Introduction,” *Journal of the European Economic Association*, 1(6), 1239–1249.
- PALACIOS-HUERTA, I., AND O. VOLIJ (2004): “The Measurement of Intellectual Influence,” *Econometrica*, 72(3), 963–977.
- PINSKI, G., AND F. NARIN (1976): “Citation Influence for Journal Aggregates of Scientific Publications: Theory, with Application to the Literature of Physics,” *Information Processing & Management*, 12(5), 297–312.
- ROY, S., A. SABERI, AND Y. WAN (2008): “Majorization for the Dominant Eigenvector of a Nonnegative Matrix,” *American Control Conference, Westin Seattle Hotel, Seattle, Washington, USA*, pp. 1965–1966.
- SHOHAM, Y., AND K. LEYTON-BROWN (eds.) (2009): *Multiagent Systems. Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, Cambridge.
- SOBEL, J. (2000): “Economists’ Models of Learning,” *Journal of Economic Theory*, 94(1), 241–261.