



**HAL**  
open science

## A ranking method based on handicaps

Gabrielle Demange

► **To cite this version:**

| Gabrielle Demange. A ranking method based on handicaps. 2013. halshs-00687180v2

**HAL Id: halshs-00687180**

**<https://shs.hal.science/halshs-00687180v2>**

Preprint submitted on 11 Apr 2013 (v2), last revised 16 Sep 2013 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**PARIS SCHOOL OF ECONOMICS**  
ÉCOLE D'ÉCONOMIE DE PARIS

**WORKING PAPER N° 2012 – 16**

**A ranking method based on handicaps**

**Gabrielle Demange**

**JEL Codes: D71, D89**

**Keywords: Ranking ; Scores, Invariant Method ; Peers Method ; Handicap ;  
Scaling Matrix**



**PARIS-JOURDAN SCIENCES ÉCONOMIQUES**

48, Bd JOURDAN – E.N.S. – 75014 PARIS  
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10  
[www.pse.ens.fr](http://www.pse.ens.fr)

# A ranking method based on handicaps

Gabrielle DEMANGE<sup>1</sup>

April 9, 2013

## Abstract

Ranking methods are a fundamental tool in many areas. Popular methods are based on the statements of ‘experts’ and aggregate them in some way. As such, there is a variety of meaningful ranking methods, more or less adapted to the environment under consideration. This paper introduces and characterizes a new method, called the handicap-based method. The method assigns scores to the items *and* weights to the experts. Scores and weights form an equilibrium for a relationship based on the notion of handicaps. The method is, in a sense made precise in the paper, the counterpart of the counting method in environments that require intensity-invariance. Intensity-invariance is a desirable property when the intensity of the experts’ statements has to be controlled. Otherwise, both the counting and handicap-based methods satisfy a property called homogeneity, which is a desirable property when cardinal statements matter, as is the case in many applications.

**Keywords** ranking, scores, invariant method, peers’ method, handicap, scaling matrix.

**JEL** D71, D89.

## 1 Introduction

The use of rankings is becoming pervasive in many areas. In the Web environment and in academia, popular ranking methods are based on observed data such as hyperlinks

---

<sup>1</sup>PSE-EHESS, address 48 bd Jourdan, 75014 Paris, France e-mail demange@pse.ens.fr. Part of the material was in the working paper entitled ‘Collective attention and ranking methods’ WP 2010-42, PSE, which I split into two papers. I thank Oscar Volij for providing me journals’ citations data.

for ranking Web pages or citations for ranking journals. The underlying premise is that these data reflect preferences: interpreting hyperlinks toward a Web page or citations toward an article as positive votes, a ranking method aggregates these votes. Science Citation Index for example uses the *counting* method, which counts the total number of cites received by journals; the impact factors introduced by Pinski and Narin (1976) use a more complex method that counts not only the direct cites but also the indirect ones in a certain way. PageRank of Google (Brin and Page 1998) is based on a similar method, called the *invariant* method. Methods can produce different results and arguments for and against each method can be put forward. This difficulty arises in the so-called peers' settings in which the items (or agents) are ranked according to data provided by these items (or agents), as in the two above examples; it also arises in settings in which the items to be ranked differ from the experts. Viewing a ranking method as aggregating evaluations or preferences of several individuals, no method is universally good as known from social choice theory; there is a variety of reasonable methods and the choice of a method should depend on the context under consideration. The axiomatization approach, which aims to characterize methods through properties or 'axioms', provides a useful guide for choosing a method. This is the approach followed here. The paper introduces a new method, called the *handicap-based method*, and characterizes it through some properties. The method applies to general settings including the peers' one as I explain now.

I consider ranking problems described by a set of 'items' to be ranked and a set of 'experts', who provide statements on the items. A ranking method assigns to each profile of experts' statements a ranking of the items. Such a ranking<sup>2</sup> is described by the relative *scores* assigned to the items, namely a ranking is cardinal and defined up to a multiplicative scalar.

For example, in a ranking of journals based on citations, journals are both the items to be ranked and the experts; a statement of a journal (as an expert) is derived from the number of citations by its articles toward articles published in all journals. Similarly, in a ranking of Web pages based on the link structure, the statement of a page (as an expert) is given by its links toward other pages. A different type of

---

<sup>2</sup>The term 'ranking' might not be appropriate to a cardinal setting but this is a rather standard terminology.

example arises when representatives state their preferences over the distribution of a budget to various issues; here the items are the issues and the experts are the representatives. A method describes how the final budget will be distributed as a function of the representatives' statements.

To understand how the handicap-based method is built, it is useful to view a method as assigning simultaneously scores to the items and weights to the experts. This encompasses two properties. First, the ranking writes as a weighted combination of the experts' statements and second, the scores and the weights form an equilibrium relationship. Current methods are built that way, though not always explicitly. The counting method, which ranks items according to their received totals, simply assigns identical weights to the experts whatever their statements. The invariant method alluded to above determines which Web pages are influential on the basis that a page is influential if it is heavily cited by other influential pages. This generates a loopback between the scores and the weights. By its very definition, the invariant method looks for scores and weights that form an equilibrium of the process of equalization between scores and weights.

The method introduced in this paper is supported by an equilibrium relationship built on handicaps. Observe that scores and handicaps are strongly related, specifically a handicap can be seen as the inverse of a score. Since the purpose of handicaps is to adjust the items' marks so as to equalize their 'strength', the handicap of  $i$  is half that of  $\ell$  if  $i$  can be said to be twice as good as  $\ell$ , that is, if  $i$ 's score is the double of that of  $\ell$ . In a handicap-based ranking, handicaps equalize the strengths between items and experts' weights equalize the handicap points assigned by the experts. These relationships define a unique ranking under mild conditions (Proposition 1). The handicaps can be viewed as a tool for building the method; I present now a justification based on the properties enjoyed by the method.

The handicap-based method is characterized by three properties. A main property is *intensity-invariance*. The property has been introduced in the environments in which the 'intensity' of statements is not controlled, as in the first two examples above, where neither the number of citations per article nor the number of links from a page are restricted *a priori*. In the Web environment for example, PageRank deflates a link from a site by the total number of links from that site. Factoring out the intensity of outward links avoids pages to increase their score by inflating

the number of links. Note that intensity-invariance is automatically satisfied when statements are controlled and represent shares, as in the budget allocation problem. A second important property is called *homogeneity*. In our framework, the (relative) values taken by the scores, and not only the order, matter. Saying that the score of  $i$  is twice the score of  $k$  should mean that  $i$  is 'twice as good' as  $k$ . In this interpretation, start with some statements and multiply the statement for  $i$  relative to other items by an identical factor for each expert. The homogeneity property requires the relative score of  $i$  to be improved by that factor. The third property is *uniformity*,<sup>3</sup> according to which items are not discriminated when they all receive the same totals: the method assigns equal scores when the counting method does.

The handicap-based method is the only method that is intensity-invariant, homogeneous, and uniform. Furthermore, the handicap-based method can be seen as the counterpart of the counting method in environments that require intensity-invariance. Indeed, the counting method, which is not intensity-invariant, is characterized by homogeneity and uniformity on the set of statements whose intensity has not been factored out. To summarize, the handicap-based method provides a tool for aggregating shares; it is appropriate when experts statements are not controlled and intensity-invariance is perceived as a prerequisite, or when experts statements are controlled and represent shares (since then intensity-invariance is automatically satisfied).

To illustrate the method, Section 3.4 reports the invariant and handicap-based rankings of 37 economic journals, based on the same data as in Palacios-Huerta and Volij( 2004). There are differences in the rankings, though rather moderate but for some exceptions. An interesting lesson from the exercise is the information carried out by the weights of the handicap-based ranking, which reflect significant differences in the reference behavior across journals.

This paper is related to recent studies that have characterized ranking methods based on citations. In the peers' context, several axiomatizations of the invariant method have been provided (Palacios-Huerta and Volij 2004, Slutzki and Volij 2006, and Altman and Tennenholtz 2005 for its ordinal version). Du, Lehrer and Pautner (2012) propose a 'market' approach to rank Web pages and obtain a variety of methods including the invariant method. These methods are discussed

---

<sup>3</sup>Intensity-invariance and uniformity have been considered in the literature by various authors, as those cited at the end of this introduction. Homogeneity in the sense used here seems to be new.

in Section 2.3. In a different context, Woeginger (2008) provides an axiomatization of the  $h$ -index, a much used method for ranking researchers. The method relies on the number of citations received by each researchers' paper, independently of the citations origin. The paper is also related, albeit loosely, to the concern on incentive compatibility, which has been recently investigated in the peers' context. Altman and Tennenholtz (2008) provide an impossibility theorem. Clippel, Moulin, and Tideman (2008) introduce a requirement in the spirit of non-manipulability (the score of an entity must not be affected by its own citations) and display a family of methods that satisfy it. This paper does not consider incentive compatibility, though intensity-invariance prevents a specific form of manipulation.

The paper is organized as follows. Section 2 presents ranking methods, defines some properties, and describes the invariant and the HITS methods. Section 3 introduces the handicap-based method under the assumption of strictly positive statements and shows how to compute the ranking can be computed using a well-known procedure of matrix scaling, called RAS method or iterative proportional fitting procedure. It provides characterizations for both the handicap-based and counting methods. Comparisons between methods are drawn and rankings of economic journals presented. Section 4 investigates settings in which statements can be nil or binary, a typical situation in the Web environment in which statements are limited to the presence or absence of a link. Most proofs are given in Section 6.

## 2 Ranking methods

### 2.1 The framework

Let  $N$  be the set of items to be ranked. Items can be individuals, or journals, or political parties, etc.. Let  $M$  be the set of experts. Experts can be pundits, or journals, or voters, etc.; in the judgment by peers, an expert is an item and the two sets  $N$  and  $M$  coincide. Examples are provided below. In the following, an element of  $N$  is called an *item* and an element of  $M$  an *expert*, keeping in mind the different interpretations.

Experts provide statements on which the ranking of the items will be based. A statement assigns a *valuation* on each item and is described by a column vector. Let

$n$  denote the cardinality of  $N$  and  $m$  the cardinality of  $M$ . Experts' statements are described by a  $n \times m$  *statement matrix*  $\boldsymbol{\pi} = (\pi_{i,j})$ , in which  $j$ 's column represents  $j$ 's statement over  $N$ . Row  $i$  thus represents all the experts' valuations on  $i$ . Given these statements, one seeks for a ranking that assigns to each item  $i$  a non-negative number  $r_i$ , called the *score* of  $i$ . The aim of the ranking is to provide the relative strength of the  $n$  items. This means that not only the order of the scores matters but also their values up to a multiplicative constant. Normalizing the sum of the scores to 1, a ranking of  $N$  is given by a vector  $\mathbf{r}$  in the simplex  $\Delta_N$ :  $\Delta_N = \{\mathbf{r} = (r_i) \in \mathbb{R}^n, r_i \geq 0, \sum_i r_i = 1\}$ .

Statements are restricted to be non-negative,  $\pi_{i,j} \geq 0$  for each  $i, j$ ; additional restrictions will be considered. A matrix is said to be feasible if each experts' statement is feasible. A method assigns a ranking to each feasible statement matrix. Formally,

**Definition 1** *Given  $N, M$ , and the set  $\mathcal{S}$  of feasible statements matrices, a ranking method  $F$  assigns to each matrix  $\boldsymbol{\pi}$  in  $\mathcal{S}$  a ranking  $\mathbf{r} = F(\boldsymbol{\pi})$  in  $\Delta_N$ .*

For example, the **counting** method assigns scores proportional to the received totals:

$$r_i = \frac{\pi_{i+}}{\sum_{\ell} \pi_{\ell+}} \text{ for each } i \text{ where } \pi_{i+} = \sum_{j \in M} \pi_{i,j}. \quad (1)$$

More sophisticated methods are introduced in Section 2.3. Let us illustrate the problem in different contexts.

1. Ranking journals based on citations. The two sets of items and experts  $N$  and  $M$  coincide, both given by the set of journals under consideration, and statements are the number of citations by articles in journal  $j$  toward articles published in journal  $i$ . To be more precise, let  $C_{i,j}$  be the total number of cites from  $j$  to  $i$  in a relevant period. Cites are normalized to account for the total number  $n_j$  of articles in  $j$ : the matrix  $\boldsymbol{\pi}$  is given by  $\pi_{i,j} = \frac{C_{i,j}}{n_j}$  where  $\pi_{i,j}$  is the average number of citations per article in  $j$  to articles in  $i$ . The total of column  $j$  is the average total number of citations in an article of  $j$ . This total may differ across journals due to different practice. The problem is to assign a ranking based on  $\boldsymbol{\pi}$ .
2. Representation problem. The problem is to assign voting weights to categories based on the votes of electoral bodies. In the political domain for example, a category is a political party and an electoral body represents a constituency,



say a district. In the context of a scientific association, a category represents a specialty and the electoral bodies are defined by geographical areas. So  $N$  is a set of categories,  $M$  is a set of electoral bodies and statements are the number of votes from  $j$  to  $i$ . The counting method treats all the votes similarly without distinguishing from which electoral body they have been cast: it assigns voting weights to categories in proportion to their votes' totals. Other methods distinguish the votes according to the electoral body (see more on the subject in next section).

3. Budget allocation. The problem is to allocate a budget to different issues (transport, sanitation, education,...) as a result of the desiderata of citizens or representatives. Here the items to be ranked are the issues, the experts are the representatives, and a representative's statement is given by the *shares* of the budget she would like to allocate to each issue (possibly not knowing the budget total).<sup>4</sup> Each expert's statement can be represented by a vector that sums to one and will be called *relative*. Here, a method describes a rule that distributes the final budget as a function of the representatives' statements.
4. Ranking Web pages based on the link structure. The two sets of items and experts  $N$  and  $M$  coincide, both given by the set of 'relevant' pages. A method defines a ranking of the pages that is based on the links within  $N$ , as performed by PageRank using the invariant method or by the HITS method described in Section 2.3. Hence the statement  $\pi$  is the (transpose of) adjacency matrix of the Web network:  $\pi_{i,j}$  is equal to 1 if page  $j$  points to  $i$  and 0 otherwise. One could account for the number of times a page links toward another page. By doing so, the statement matrix has zeros, possibly many, but is not restricted to be 0-1; this is however not the standard practice. Such a binary representation

---

<sup>4</sup>This is in the spirit of the Participatory Budgeting procedures implemented in Porto Alegre (see e.g. Souza 1998). The system is more complex since it involves two levels of voting, first at the neighborhood regional (the city is divided into regions, themselves divided into neighborhoods) or thematic level and second at the regional level. All inhabitants in a neighborhood can vote both to formulate priority demands over investment spendings and to elect representatives at the region level. These representatives aggregate the neighborhoods' demands to set regional priorities. Finally all these regions' priorities are used by the city to derive a general priority ranking over issues and an allocation to regions.

also arises in approval voting, in which an expert is asked to name the items he finds acceptable (without being allowed to state intensity). Here items and experts can differ. An expert's statement is described by the vector of 1 and 0 indicator of the set of items he approves.

The setting where the matrix has only zeros and ones will be called the 0-1 setting.

In all these examples but the last one, the statements are cardinal and the precise relative values stated by the experts have a meaning. It is thus natural to assign a cardinal ranking. In the 0-1 setting, though statements are not cardinal, a cardinal ranking still makes sense. For example, the proportion of the links that points to each site or the proportion of the approval votes received by each candidate are relevant (these proportions are produced by the counting method). But some axioms valid for cardinal statements may have no sense in this 0-1 setting. This will be the case for the homogeneity axiom introduced in the next section.

Statement matrices are restricted in most of the paper to be positive,  $\pi_{i,j} > 0$  for each  $i, j$ , as can be reasonably assumed in the three first examples. Two feasible sets will be considered: the full set of positive matrices denoted by  $\mathcal{P}$  representing absolute statements (the journal and representation examples) and the set of positive matrices for which each column sums to 1, denoted by  $\mathcal{R}$ , representing relative statements (the budget example). The possibility of zeros and the 0-1 setting will be analyzed in Section 2.2.

**Notation**  $\mathbf{1}_N$  denotes the vector in  $\mathfrak{R}^N$  whose components are equal to 1, and  $\mathbf{e}_N = \frac{1}{N}\mathbf{1}_N$  the ranking that assigns equal scores to items.

Given a finite set  $I$  and a vector  $\mathbf{x}$  in  $\mathfrak{R}^I$ ,  $dg(\mathbf{x})$  denotes the diagonal  $I \times I$  matrix with  $x_i$  as the  $i$ -th element on the diagonal.

Given a matrix  $\boldsymbol{\pi} = (\pi_{i,j})$ ,  $\pi_{i+}$  denotes the total in row  $i$ ,  $\pi_{i+} = \sum_{j \in M} \pi_{i,j}$ , and  $\pi_{+j}$  the total in column  $j$ ,  $\pi_{+j} = \sum_{i \in N} \pi_{i,j}$ .

## 2.2 Some properties

Let us start by describing natural properties that one may want a method to satisfy. Intensity-invariance, uniformity, and exactness appear in the literature under various

names.

**Intensity-invariance** Intensity-invariance requires the ranking not to be affected by a multiplicative scaling of a column. Justifications in different contexts are provided below. Formally, let  $\boldsymbol{\pi}'$  be the matrix obtained from  $\boldsymbol{\pi}$  by multiplying a column, say column  $j$ , by a positive scalar  $\mu_j$ . Intensity-invariance of  $F$  requires the ranking of  $\boldsymbol{\pi}$  and  $\boldsymbol{\pi}'$  to be identical; this writes in matrix form  $F(\boldsymbol{\pi}dg(\boldsymbol{\mu})) = F(\boldsymbol{\pi})$  where  $\boldsymbol{\mu}$  is the vector with  $j$ -th component equal to  $\mu_j$  and all others equal to 1. The property is required for each column  $j$ ; thus iteration yields the following equivalent definition.

**Definition 2** A method  $F$  defined on  $\mathcal{P}$  is **intensity-invariant** if  $F(\boldsymbol{\pi}') = F(\boldsymbol{\pi})$  for  $\boldsymbol{\pi}'$  the matrix obtained from  $\boldsymbol{\pi}$  by multiplying column  $j$  by a positive  $\mu_j$ , for each  $j$  in  $M$ :

$$F(\boldsymbol{\pi}dg(\boldsymbol{\mu})) = F(\boldsymbol{\pi}) \text{ for each positive } m\text{-vector } \boldsymbol{\mu} = (\mu_j), \text{ each } \boldsymbol{\pi} \text{ in } \mathcal{P}. \quad (2)$$

Basically, an intensity-invariant method is characterized by its values on matrices whose column totals are fixed. To see this, let a positive value for each column total,  $c_j$  for column  $j$ . Given a matrix  $\boldsymbol{\pi}$ , scale each of its column so that it meets the required total. Such a matrix is equal to  $\boldsymbol{\pi}dg(\boldsymbol{\mu})$  where  $\mu_j$  satisfies  $\pi_{+j}\mu_j = c_j$ . Intensity-invariance of  $F$  implies  $F(\boldsymbol{\pi}) = F(\boldsymbol{\pi}dg(\boldsymbol{\mu}))$ :  $F$  is fully determined by its restriction on the set of matrices whose column totals are fixed by  $\mathbf{c}$ .

This suggests a way to build intensity-invariant methods from a non intensity-invariant method  $F$  by extending properly the restriction of  $F$  to matrices with given column totals  $\mathbf{c}$ : given a matrix, scale each of its column so that it meets the total  $\mathbf{c}$  and apply  $F$  to the scaled matrix. Formally, given columns totals  $\mathbf{c} = (c_j)$ , define  $[F]_{\mathbf{c}}$  by: for each  $\boldsymbol{\pi}$ ,  $[F]_{\mathbf{c}}(\boldsymbol{\pi}) = F(\boldsymbol{\pi}dg(\boldsymbol{\alpha}))$  where  $\pi_{+j}\alpha_j = c_j$ .  $[F]_{\mathbf{c}}$  is intensity-invariant since the scaled matrix of  $\boldsymbol{\pi}dg(\boldsymbol{\mu})$  is the same as that of  $\boldsymbol{\pi}$ . The methods  $[F]_{\mathbf{c}}$  vary with  $\mathbf{c}$  provided  $F$  is not intensity-invariant (they all coincide with  $F$  for  $F$  intensity-invariant).

This construction shows that intensity-invariance is not related with any notion of fairness. Take for instance the method  $F$  to be anonymous with respect to the experts, namely, the ranking to be independent of a permutation of the experts' indices. In that case, choosing equal totals for each expert amounts to treat them equally *a priori*,

and choosing distinct totals treats them differentially. Let us illustrate intensity-invariance in our examples.

For ranking journals, intensity-invariance means that the ranking depends only on the proportions of the citations by journals to other journals, i.e. on  $\frac{C_{i,j}}{C_{+j}}$  where  $C_{i,j}$  is the number of cites made by articles in  $j$  to articles in  $i$  (recall that  $\pi_{i,j}$  is defined as  $\frac{C_{i,j}}{n_j}$  the average number of references of an article from  $j$  to  $i$ ). As a result, a proportional increase in the number of citations per article in a journal  $j$ , keeping the shares received by each journal unchanged, has no impact on the ranking. In particular, the ranking is not influenced by distinct citations' practice across journals or fields.<sup>5</sup>

In the case of Web pages, intensity-invariance implies that a link from a page is divided by the number of links from that page. Intensity is 'factored out'. A justification, relevant for the invariant method described in the next section, is that factoring out intensity avoids that a page improves its score simply by increasing the number of pages it points to.

In the representation problem, an expert is the electoral body of a district whose statements are given by the number of votes to the parties in that district. Intensity-invariance requires the final representation to be independent of the turnout in the districts. Using the construction described above, an intensity-invariant method is obtained by assigning totals to columns, that is a total to each electoral body. The totals are not necessarily proportional to the sizes of the electoral bodies. This allows, for example, to 'over'-represent a constituency with small population and 'under'-represent one with large population.

Finally, in the budget example, intensity-invariance is automatically satisfied since representatives are asked to provide distributions of the budget.

In the sequel, we work with matrices whose column totals equal to 1, namely matrices in  $\mathcal{R}$ , and with the associated intensity-invariant version,  $[F]_{\mathbf{1}}$ .<sup>6</sup> Specifically

---

<sup>5</sup>Indeed the average number of cites per article varies across fields. See Palacios-Huerta and Volij (2004) for an analysis of the impact of cite intensities on the ranking of economic journals.

<sup>6</sup>By doing so, no intensity-invariant methods are neglected. Let  $[F]_{\mathbf{c}}$  be the intensity-invariant version of  $F$  associated with columns totals  $\mathbf{c}$ ; we show that  $[F]_{\mathbf{c}}$  is the intensity-invariant version of a method associated with columns totals  $\mathbf{1}$ :  $[F]_{\mathbf{c}} = [F_{\mathbf{c}}]_{\mathbf{1}}$  where the method  $F_{\mathbf{c}}$  is defined by  $F_{\mathbf{c}}(\boldsymbol{\pi}) = F(\boldsymbol{\pi}dg(\mathbf{c}))$ . Observe that the matrix  $\boldsymbol{\pi}dg(\boldsymbol{\mu})$  with columns totals  $c_j$ ,  $\mu_j\pi_{+j} = c_j$  can be written as  $[\boldsymbol{\pi}]dg(\mathbf{c})$ . Hence  $[F]_{\mathbf{c}}(\boldsymbol{\pi}) = F([\boldsymbol{\pi}]dg(\mathbf{c}))$ . By definition of  $F_{\mathbf{c}}$ ,  $F_{\mathbf{c}}([\boldsymbol{\pi}]) = F([\boldsymbol{\pi}]dg(\mathbf{c}))$ ,

let  $[\boldsymbol{\pi}]$  be the matrix in  $\mathcal{R}$  associated to  $\boldsymbol{\pi}$ :

$$[\boldsymbol{\pi}]_{i,j} = \frac{\pi_{i,j}}{\pi_{+j}} \text{ for each } i, j. \quad (3)$$

The intensity-invariant version  $[F]_{\mathbf{1}}$  of  $F$ , denoted simply as  $[F]$ , is defined by

$$[F](\boldsymbol{\pi}) = F([\boldsymbol{\pi}]) \text{ for each } \boldsymbol{\pi} \in \mathcal{P}. \quad (4)$$

Finally, note that any method that is defined on  $\mathcal{R}$ , such as in the budget example, can be extended (in a unique way) to an intensity-invariant method on  $\mathcal{P}$ . In what follows, any method defined on  $\mathcal{R}$  is extended that way to an intensity-invariant method on  $\mathcal{P}$ .

**Uniformity and exactness** The next two properties bear on some specific matrices, hereafter called row-balanced. A matrix is said to be *row-balanced* if each row receives the same total. Equivalently, rows obtain equal scores under the counting method. Row-balanced matrices constitute a kind of 'neutral' statements because there is no rationale for distinguishing between items if experts are not discriminated *a priori* since rows obtain identical totals. This is what is required by uniformity.

**Definition 3** *A method  $F$  is **uniform** on  $\mathcal{S}$  if  $F(\boldsymbol{\pi}) = \mathbf{e}_N$  for each row-balanced  $\boldsymbol{\pi}$  in  $\mathcal{S}$ .*

*A method is **exact** on  $\mathcal{S}$  if  $F(\boldsymbol{\pi}) = \mathbf{e}_N$  for  $\boldsymbol{\pi}$  in  $\mathcal{S}$  implies that  $\boldsymbol{\pi}$  is row-balanced.*

A method is uniform if it assigns equal scores to each row-balanced feasible statement matrix. Exactness asks the converse property that items obtain equal scores only if they receive identical totals.

Clearly the intensity-invariant version of a uniform (resp. exact) method on  $\mathcal{P}$  is uniform (resp. exact) on  $\mathcal{R}$  since  $[F](\boldsymbol{\pi}) = F([\boldsymbol{\pi}])$  for  $\boldsymbol{\pi}$  in  $\mathcal{R}$ . However  $[F]$  may not be uniform (or exact) on  $\mathcal{P}$ . Furthermore, there is no intensity-invariant method that is both uniform and exact on  $\mathcal{P}$ . To see this, start with a row-balanced matrix  $\boldsymbol{\pi}$  with distinct columns. Its ranking is  $\mathbf{e}_N$  by uniformity. The distinct columns can be multiplied by some factors so as to obtain a matrix that is not row-balanced. Intensity-invariance requires the ranking of this new matrix to be  $\mathbf{e}_N$ , in contradiction with exactness.

---

hence  $[F]_{\mathbf{c}}(\boldsymbol{\pi}) = F_{\mathbf{c}}([\boldsymbol{\pi}]) = [F_{\mathbf{c}}]_{\mathbf{1}}(\boldsymbol{\pi})$ .

**Homogeneity** The homogeneity property is very natural when statements are cardinal<sup>7</sup> as in the three first examples presented in Section 2. The relative valuations stated by an expert have a precise meaning, and if there is a single expert, we want the cardinal ranking to keep these relative valuations, hence to be proportional to the expert's statement. In particular, if the expert doubles the valuation on  $i$ , either in absolute or relative terms, the final score of  $i$  will be doubled relative to all other items. Homogeneity extends this property to the multi-expert setting. Starting with experts statements and multiplying each valuation on  $i$  by a factor,  $i$ 's relative position should be multiplied by the same factor. I spell out the property for the two situations in which absolute or relative statements matter.

Let us start with absolute statements. Let  $\boldsymbol{\pi}$  be in  $\mathcal{P}$  and multiply each valuation on item  $i$  by a positive scalar  $\rho_i$ . Homogeneity of  $F$  requires  $i$ 's score relative to other items to be multiplied by  $\rho_i$ . In matrix form, this writes as:  $F(dg(\boldsymbol{\rho})\boldsymbol{\pi})$  is the ranking proportional to  $dg(\boldsymbol{\rho})F(\boldsymbol{\pi})$  where  $\boldsymbol{\rho}$  is the vector with  $i$ 's component equal to  $\rho_i$  and all others equal to 1. The property is required for each row  $i$ , so iteration yields the following equivalent definition.

**Definition 4** *A method is homogeneous on absolute statements if for each  $\boldsymbol{\pi}$  in  $\mathcal{P}$  and positive  $n$ -vector  $\boldsymbol{\rho} = (\rho_i)$ ,  $F(dg(\boldsymbol{\rho})\boldsymbol{\pi})$  is the ranking proportional to  $dg(\boldsymbol{\rho})F(\boldsymbol{\pi})$ .*

Clearly, the counting method is homogeneous on absolute statements. To illustrate the property, let us consider the journals and representation examples.

Assume that between two periods, the citations received by a journal, say 1, from *each* journal increase by ten per cent, and all others are unchanged:  $\rho_1 = 1.1$ . Homogeneity on absolute statements requires the score of journal 1 to increase by ten percent relatively to others, hence to become  $\frac{1.1 \cdot r_1}{\sum_i r_i + 0.1 \cdot r_1} = \frac{1.1 \cdot r_1}{1 + 0.1 \cdot r_1}$ .

In the representation problem, let the number of votes in favor of party  $i$  be raised by 5 % in all districts between two elections, other numbers of votes unchanged (such an increase implies an increase in the total number of votes; this is possible since larger participation and demographic modifications make the number of cast votes

---

<sup>7</sup>Though, it has not yet been considered in the literature, as far as I know. The homogeneity axiom introduced in Palacios-Huerta and Volij (2004) is different since it bears on a given matrix that has two proportional rows.

variable). Homogeneity on absolute statements requires the voting weights of party  $i$  to be raised by 5 % relative to other parties.

Homogeneity on relative statements requires the same behavior when a factor modifies the relative valuations on an item: multiplying the shares on  $i$  relative to other items by a positive scalar  $\rho_i$  multiplies  $i$ 's score relative to other items by  $\rho_i$ . In matrix form,  $F([dg(\boldsymbol{\rho})\boldsymbol{\pi}])$  is the ranking proportional to  $dg(\boldsymbol{\rho})F(\boldsymbol{\pi})$  where  $\boldsymbol{\rho}$  is the vector with  $i$ 's component equal to  $\rho_i$  and all others equal to 1. The property is required for each row  $i$ , so iteration yields the following equivalent definition.

**Definition 5** *A method is homogeneous on relative statements if for each  $\boldsymbol{\pi}$  in  $\mathcal{R}$  and positive  $n$ -vector  $\boldsymbol{\rho} = (\rho_i)$ ,  $F([dg(\boldsymbol{\rho})\boldsymbol{\pi}])$  is the ranking proportional to  $dg(\boldsymbol{\rho})F(\boldsymbol{\pi})$ .*

Consider the budget example, in which experts state their preferred budget shares over issues (so statements are in  $\mathcal{R}$ ). Let two communities with the same number of experts. Their statements differ only by the fact that each expert in the second community assigns  $\alpha$  percent more to education relative to other issues than in the first. Homogeneity on relative statements requires the share devoted to education to be  $\rho$  percent larger relative to other issues in the second community than in the first; if the education share is 20% in the first community for example, it is  $(20 + 20\alpha)/(100 + 20\alpha)$  in the second. Let us illustrate with the following example. In each city, there are three experts and two issues -education and health. Statements in two different cities are represented in the following matrices.

$$\text{city 1 : } \begin{pmatrix} 2/3 & 1/3 & 1/2 \\ 1/3 & 2/3 & 1/2 \end{pmatrix} \quad \text{city 2 : } \begin{pmatrix} 4/5 & 1/2 & 2/3 \\ 1/5 & 1/2 & 1/3 \end{pmatrix}$$

$$\begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \rightarrow \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}$$

Each expert in city 2 wants the share on education to be the double relative to health than in city 1; the budget's share on education should be the double relative to health in city 2. If the method is uniform, the budget should be shared equally between the two issues in city 1; thus education should receive 2/3 of the budget in city 2 and health 1/3.

The distinction between homogeneity on absolute and relative statements does not matter for intensity-invariant methods, as stated in the next lemma (the proof is in the appendix).

**Lemma 1** *Let  $F$  be intensity-invariant.  $F$  is homogeneous on absolute statements if and only if it is homogeneous on relative statements.*

Thus, in the sequel, we simply refer to homogeneity for an intensity-invariant method. For a non intensity-invariant method, homogeneity is lost when intensity is factored out. Specifically, let method  $F$  be homogeneous on absolute statements but not intensity-invariant. The intensity-invariant version  $[F]$  may not be homogeneous (in whatever sense), as is illustrated with the counting method.

The counting method is homogeneous on absolute statements but the intensity-invariant version is not. In the following example,  $\pi'$  is obtained by multiplying the first row of  $\pi$  by 2:

$$\pi = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \pi' = \begin{pmatrix} 4 & 2 \\ 1 & 2 \end{pmatrix}, [\pi'] = \begin{pmatrix} 4/5 & 1/2 \\ 1/5 & 1/2 \end{pmatrix}.$$

The intensity-invariant version of the counting method assigns equal scores,  $(1/2, 1/2)$ , to  $\pi$  but  $(13/20, 7/20)$  to  $\pi'$  (and to  $[\pi']$ ) instead of  $(2/3, 1/3)$  as required by homogeneity. The reason is that expert 1 likes relatively more item 1 than expert 2 does, hence its total increases more than that of expert 2 when item 1's statements are doubled. Thus, when normalizing  $\pi'$ , the adjustment on expert 1's statements is larger than on 2's. This explains why the total of item 1's shares is less than doubled relative to that of item 2.

### 2.3 Examples: the Invariant and HITS methods

This section introduces well-known methods that differ from the counting one. Whereas the counting method treats experts equally, these methods treat them differently according to their statements.

The next two methods, sometimes referred to as 'eigenvalue' methods, are the Liebowitz-Palmer method<sup>8</sup> and its intensity-invariant version, called the invariant method. The sets of items and experts coincide ( $N = M$ ). The methods are 'peers'

---

<sup>8</sup>This terminology refers to the work of Liebowitz and Palmer (1984), who use an approximation



methods based on the premise that the statements made by a peer (as an expert) should be weighed by his score. This induces a loopback effect: the scores are proportional to the weighted sums of their valuations where the weights are the scores. Specifically the method looks for  $\mathbf{r}$  in  $\Delta_N$  that satisfies

$$\text{for some positive } \lambda, r_i = \lambda \sum_{j \in N} \pi_{i,j} r_j \text{ for each } i. \quad (5)$$

According to (5), the positive vector  $\mathbf{r}$  is an eigenvector of matrix  $\boldsymbol{\pi}$ . By Perron-Frobenius theorem on matrices with positive elements, such an eigenvector exists, and is unique up to a scalar: (5) well defines a method.

The invariant method is the intensity-invariant version<sup>9</sup> of the Liebowitz-Palmer method. Since a normalized matrix has its largest eigenvalue equal to one,<sup>10</sup> the invariant ranking of  $\boldsymbol{\pi}$  is the unique  $\mathbf{r}$  in the simplex that satisfies

$$r_i = \sum_{j \in N} [\pi]_{i,j} r_j \text{ for each } i. \quad (6)$$

The method is intensity-invariant, uniform, exact, but not homogeneous.

In the Web setting (or more generally a network setting) the methods proposed by Du *et al.* (2012) can be seen as a generalization of the invariant method. They associate an exchange economy to the network and assign the resulting equilibrium prices of the nodes as their scores. For Cobb-Douglas preferences, the invariant ranking is obtained. Though different ranking methods are obtained for different families of preference, all methods reflect the idea that the higher the score of a node, the more valuable its recommendation. In my view, this property qualifies a method as a 'pure' peer's method (see Demange 2012-b for a definition and an analysis of this property in a dynamical framework). The property is not satisfied by the next method.

---

of the method for ranking economic journals. The methods or some variants have been (re)defined and used in various contexts: in sociology by Katz (1953) and Bonacich (1987), in academics for ranking journals by Pinski and Narin (1976).

<sup>9</sup>It serves as a basis to PageRank to rank Web pages using the link structure between pages by Brin and Page (1998). Because the matrix has many zeros, a perturbation is used to make it irreducible.

<sup>10</sup>Recall that a matrix and its transpose have identical eigenvalues. The set of equations  $\sum_i \pi_{i,j} = 1$  for each  $j$  implies that  $\mathbf{1}_N$ , is an eigenvector of the transpose of  $\boldsymbol{\pi}$  with eigenvalue 1. Since  $\mathbf{1}_N$  positive, 1 is the dominant eigenvalue.

The Hyperlink-Induced Topic Search (HITS) method, introduced by Kleinberg (1999), assigns scores to a set of Web pages on the basis of their link structure, as the invariant method used in PageRank. Thus, the two sets of items and experts coincide. The method however distinguishes two weights for each Web page, one associated with the relevance or authority of a page, the other with the adequacy of a page to point toward the relevant pages. The first set of weights defines the ranking, which should help users to find the relevant pages. The second set of weights identifies the pages -called 'hubs'- that are important because they point to relevant pages (but might not be useful to Internet users). Specifically the method assigns the ranking  $\mathbf{r}$  and the experts weights  $\mathbf{q}$  in  $\Delta_N$  that satisfy

$$r_i = \sum_j \pi_{i,j} q_j \text{ for each } i \text{ and } q_j = \lambda \sum_i \pi_{i,j} r_i \text{ for each } j \quad (7)$$

for some positive  $\lambda$ . As argued by Kleinberg (1999), hubs and authorities exhibit a *mutually reinforcing relationship*: a good authority is a page that is pointed to by many good hubs, a good hub is one that points to many good authorities. The method is well defined: (7) writes in matrix form as  $\mathbf{r} = \boldsymbol{\pi} \mathbf{q}$  and  $\mathbf{q} = \lambda \tilde{\boldsymbol{\pi}} \mathbf{r}$  where  $\tilde{\boldsymbol{\pi}}$  is the transpose of  $\boldsymbol{\pi}$ . Thus  $\mathbf{r} = \lambda \boldsymbol{\pi} \tilde{\boldsymbol{\pi}} \mathbf{r}$  and  $\mathbf{q} = \lambda \tilde{\boldsymbol{\pi}} \boldsymbol{\pi} \mathbf{q}$ . The 'authority' weights  $\mathbf{r}$  and the 'hub' weights  $\mathbf{q}$  are respectively the unique normalized principal eigenvectors of the positive matrices  $\boldsymbol{\pi} \tilde{\boldsymbol{\pi}}$  and  $\tilde{\boldsymbol{\pi}} \boldsymbol{\pi}$ .<sup>11</sup> The method is uniform, exact, but not homogeneous on absolute statements.

**Supporting weights** Observe that the counting, invariant, and HITS methods can all be viewed as assigning not only scores to items but also 'weights' to experts. As can be seen from (1), (6) and (7), each of the three methods assigns to each  $\boldsymbol{\pi}$  a ranking  $\mathbf{r} = F(\boldsymbol{\pi})$  and weights  $\mathbf{q} = Q(\boldsymbol{\pi})$  to experts so that each score writes as the sum of its valuations weighted by the experts weights:

$$r_i = \sum_{j \in M} \pi_{i,j} q_j \text{ for each } i \text{ in } N \text{ where } \mathbf{r} = F(\boldsymbol{\pi}) \text{ and } \mathbf{q} = Q(\boldsymbol{\pi}) \quad (8)$$

where furthermore the ranking  $F(\boldsymbol{\pi})$  and the weights  $Q(\boldsymbol{\pi})$  are related in an equilibrium relationship. Let us spell out this in each case.

---

<sup>11</sup> $1/\lambda$  is the dominant eigenvalue of  $\boldsymbol{\pi} \tilde{\boldsymbol{\pi}}$  and is not equal to 1 in general even for a normalized  $\boldsymbol{\pi}$ .

The counting method assigns equal weights to the experts, whatever the statements :  $Q(\boldsymbol{\pi}) = 1/m$ . The equilibrium relationship is trivial, based on the premise that no distinction should be made between experts whatsoever. The invariant method, by its very definition, assigns weights to experts so as to equalize scores to weights:  $F(\boldsymbol{\pi}) = Q(\boldsymbol{\pi})$ . The HITS method assigns weights so as to form an equilibrium relationship between 'authority' scores and 'hubs' scores:  $Q(\boldsymbol{\pi})$  is the normalized vector proportional to  $\tilde{\boldsymbol{\pi}}F(\boldsymbol{\pi})$ .

Thus, apart from the counting method, experts are treated differently as the weights assigned to their statements typically differ. This differential treatment depends on the whole profile of statements and results from an equilibrium relationship. One can extend this insight and define a variety of methods by varying the equilibrium relationship. More precisely, given a relationship between rankings and weights, let us consider a method that simultaneously assigns scores to the items and weights to the experts so that (1) the ranking writes as a weighted combination of the experts' statements and (2) the scores and the weights are in equilibrium. This is what is performed in the next section.

### 3 The handicap-based method

This section first introduces a method based on an equilibrium relationship between items' scores and experts' weights supported by the notion of handicaps (Proposition 1). Proposition 2 states two characterizations of the method using the axioms introduced in the previous section. Finally, Proposition 3 provides characterizations of the counting method that shows the similarities between the handicap-based and counting methods. Statements are assumed here to be all positive. Null entries are considered in Section 4.

#### 3.1 Definition and properties

The purpose of handicaps is to equalize the strengths between items. Handicaps and scores may be seen as inversely related: saying that the handicap of  $i$  is twice that of  $\ell$  means that the score of  $i$  is half that of  $\ell$ . So handicaps are assigned to a ranking  $\mathbf{r}$  by the relationship  $h_i = 1/r_i$ . The handicap-based method is based on

an equilibrium relationship between handicaps and experts' weights : it looks for handicaps that equalize items' weighted counts and for experts' weights that equalize the distributed handicaps across experts. Proposition 1 states that this leads to a well defined method.

**Proposition 1** *Given a positive matrix  $\pi$ , there is a unique  $\mathbf{r} = (r_i)$  in  $\Delta_N$  such that*

$$\sum_j (\pi_{i,j} q_j) \frac{1}{r_i} = 1 \text{ for each } i \text{ where } \sum_i (\frac{\pi_{i,j}}{r_i}) q_j = \frac{n}{m} \text{ for each } j. \quad (9)$$

*The **handicap-based method**  $H$  assigns to each matrix this unique ranking  $\mathbf{r}$  and the supporting weights  $\mathbf{q}$ .  $H$  is intensity-invariant; furthermore it is uniform, exact, and homogeneous.*

Defining handicaps as  $h_i = 1/r_i$ , (9) handicaps and weights exhibit a mutual relationship:

$$\sum_j (\pi_{i,j} q_j) h_i = 1 \text{ for each } i \text{ where } \sum_i (\pi_{i,j} h_i) q_j = \frac{n}{m} \text{ for each } j. \quad (10)$$

According to the first set of equations, handicaps equalize items' weighted counts and according to the second set, experts' weights equalize their distributed handicap 'points' defined as  $\sum_i \pi_{i,j} h_i$ .

The differences in the items' scores reflect the differences in their valuations, i.e., the rows of the matrix, whereas the differences in the experts' weights reflect the differences in their statements, i.e., the columns. The experts' weights reflect the ability to make 'relevant' statements where relevance is measured with respect to the final ranking (as in the HITS method but using a different measure). In general, a large weight indicates that the expert tends to assign relatively large valuations to items with low handicaps (hence large scores) and relatively low valuations to items with high handicaps. The term 'relatively' is important because in absolute terms items with low handicaps tend to receive large valuations. To make this point clear, an expert whose statements coincide with the final ranking has a weight equal to the average,  $1/m$ . Another expert whose statements are 'biased' toward items with large scores will receive a larger weight. In particular, the maximal expert's weight is obtained for a statement almost concentrated on the item with the largest score (and

is equal to that score). Similarly, the minimal weight is obtained by an expert whose statement are concentrated on that item.

The proof is in the appendix. The existence of a ranking solution to (9) relies on the following observation. The method can be seen as searching for a ranking and experts' weights that transform the statements into a matrix that is both row and column balanced. Specifically, (9) requires the matrix  $\mathbf{p}$  of general element  $p_{i,j} = \frac{1}{r_i} \pi_{i,j} q_j$  obtained from  $\boldsymbol{\pi}$  by multiplication of its rows by the items' handicaps and of its columns by the experts' weights to have the same total in each row as well as in each column

$$\sum_j p_{i,j} = 1 \text{ for each } i \text{ and } \sum_i p_{i,j} = n/m \text{ for each } j. \quad (11)$$

The problem of adjusting a given matrix  $\boldsymbol{\pi}$  by multiplication of its rows and its columns by some numbers so as to meet some totals constraints is a standard matrix scaling problem, known as the RAS model in economics or the log-fitting model in statistics. Relying on these results the matrix  $\mathbf{p}$  is unique. It remains to show that the multipliers are uniquely defined when  $\mathbf{r}$  is in the simplex.<sup>12</sup>

The handicap-based ranking and the associated experts' weights can be computed through an iterative process by using the iterative scaling algorithm RAS (Bacharach 1965). Let  $\boldsymbol{\pi}$  be a positive normalized matrix. The procedure starts by assigning equal weights to experts,  $q_j^0 = 1/m$  each  $j$ , and the handicap points,  $h_i^0$  to  $i$ , inverse to the weighted items totals:  $h_i^0 (\sum_i \pi_{i,j}) q_j^0 = 1$  for each  $i$ , that is  $r_i^0 = (1/h_i^0)$  is the counting score. Define the handicap points distributed by  $j$  as  $\sum_i h_i^0 \pi_{i,j}$ . If they are identical across  $j$ , the process stops, and the handicap-based ranking is the counting ranking. Otherwise, define experts' weights  $\mathbf{q}^1$  so as to equalize the distributed points across experts to  $\frac{n}{m}$ :

$$\left( \sum_i h_i^0 \pi_{i,j} \right) q_j^1 = \frac{n}{m} \text{ for each } j.$$

(The vector  $\mathbf{q}^1$  is not necessarily in the simplex.) Compute handicaps  $\mathbf{h}^1$  so as to equalize the items totals weighted by  $\mathbf{q}^1$ ; the procedure starts over again, alternating

---

<sup>12</sup>The approaches differ in that the object of interest of the RAS method is the final adjusted matrix  $\mathbf{p}$  whereas we are interested in the (relative) values of the adjustment on rows and columns so as to define the handicaps and experts' weights. See for example Balinski and Demange (1989-b).

row-scaling and column-scaling: for each  $\tau = 1, \dots$

$$h_i^\tau \left( \sum_j \pi_{i,j} q_j^\tau \right) = 1 \text{ for each } i \quad (12)$$

$$\left( \sum_i h_i^\tau \pi_{i,j} \right) q_j^{\tau+1} = \frac{n}{m} \text{ for each } j. \quad (13)$$

The process can be shown to converge<sup>13</sup> to some positive vectors  $(\mathbf{h}, \mathbf{q})$  that satisfy (10). Take the vector  $\mathbf{r}$  associated to handicap  $\mathbf{h}$ :  $r_i = \lambda/h_i$  where  $\lambda$  is chosen to have the sum  $\sum r_i$  equal to 1 and weights equal to  $\lambda \mathbf{q}$ . Since the vectors  $(\mathbf{r}, \lambda \mathbf{q})$  satisfy (9) and  $\mathbf{r}$  belongs to the simplex,  $\mathbf{r}$  is the handicap-based ranking.

### 3.2 Characterization of the handicap-based and the counting methods

The next proposition provides a characterization of the handicap-based method.

#### Proposition 2

- (a) *The handicap-based method is the only ranking method that is uniform on  $\mathcal{R}$ , intensity-invariant, and homogeneous.*
- (b) *The handicap-based method is the only ranking method that is exact on  $\mathcal{R}$ , intensity-invariant, and homogeneous.*

The next proposition provides a characterization of the counting method.

#### Proposition 3

- (a) *The counting method is the only method that is homogeneous on absolute statements and uniform on  $\mathcal{P}$ .*
- (b) *The counting method is the only method that is homogeneous on absolute statements and exact on  $\mathcal{P}$ .*

Comparing the characterizations of the handicap-based and counting methods shows that their main difference stems from intensity-invariance. Put differently, the handicap-based method is adequate for relative statements and the counting method for absolute ones.

---

<sup>13</sup>The convergence is not straightforward, especially when some elements are null, as can be seen in the vast literature on the subject (see the recent survey of Pukelsheim (1976) for example). With all elements positive, convergence is ensured at a geometric rate that depends on the final matrix  $p$  (Soules 1999).

The ordinal ranking associated to the counting method has been axiomatized in the 'tournament' setting. Rubinstein (1976) considers a simple tournament in which  $i$  wins over  $j$  or  $j$  wins over  $i$  ( $\pi_{i,j}$  is 0 or 1 and  $\pi_{i,j} + \pi_{j,i} = 1$ ). Brink and Gilles (2003) consider the 0-1 setting and Brink and Gilles (2009) consider weighted directed graphs, which correspond to non-negative matrices with null elements on the diagonal. The two first papers rely on an axiom of independence of alternatives (IIA), meaning that the ordering of two items, say  $i$  and  $k$ , only depend on the items beaten by  $i$  or  $k$  and on those that beat one of them. In other words, the scores are determined by the neighbors in the graph and all the nodes that are not directly linked to  $i$  and  $k$  have no impact on their scores. Our characterization of the counting method clearly differs since it does not rely on IIA. Brink and Gilles (2009) uses an axiom dealing with the sum of matrices, hence is also different.

Let us add a final remark on IIA. IIA is clearly violated by all methods introduced here except the counting method; indeed their aim is precisely to account for the whole statements profile to derive expert's weights. More generally, IIA is not satisfied by any method for which the weights vary with the statements because then the score of an item depends on the statements over all items via the experts weights. This justifies to give up IIA.

### 3.3 Comparison between methods

The counting, HITS, handicap-based, and invariant methods all assign  $\mathbf{r}$  to unanimous statements (unanimous means  $\pi_{i,j} = r_i$  for a common  $\mathbf{r}$ ). The weights are all equal, equal to  $1/m$ , for all methods but the invariant ranking, which has weights equal to  $\mathbf{r}$ . This implies that when preferences do not differ much, the counting, HITS and handicap-based methods behave similarly but not the invariant method.

The handicap-based and HITS methods differ in the way experts' weights depend on a ranking  $\mathbf{r}$ . The experts' weights are the harmonic means of  $\mathbf{r}$  for the handicap-based method and the average means for the HITS (up to a multiplicative factor), each weighted with the expert's statement. This 'direct' effect induces an indirect effect because of the impact of the scores and the overall equilibrium effect is not easy to assess. However the less variable the ratio  $\pi_{i,j}/r_i$  with respect to  $i$ , the closer the average and harmonic means are. One could thus expect the two methods to produce

close rankings when these ratios computed at the handicap-based ranking are not too variable.

This extends the previous remark on unanimous statements since the ratios are constant in that case.

**An example** The invariant method is intensity-invariant and exact but not homogeneous. The following example shows that its behavior with respect to changes in statements may present serious drawbacks in some contexts. Let  $\boldsymbol{\pi}$  and  $\boldsymbol{\pi}'$  be

$$\boldsymbol{\pi} = \begin{pmatrix} 0 & 1/3 & 1 \\ 1 & 0 & 1/3 \\ 1/3 & 1 & 0 \end{pmatrix} \quad \boldsymbol{\pi}' = \begin{pmatrix} 0 & 10/19 & 10/13 \\ 9/10 & 0 & 3/13 \\ 1/10 & 9/19 & 0 \end{pmatrix},$$

(I choose null diagonal elements to show that the behavior is not due to self-citations). Comparing the statements in the two matrices, those for 1 relative to 2 are 10/9 times larger in  $\boldsymbol{\pi}'$  than in  $\boldsymbol{\pi}$ , and those for 2 relative to 3 are 3 times larger. That is,  $\boldsymbol{\pi}' = [dg(\boldsymbol{\rho})\boldsymbol{\pi}]$  for  $\boldsymbol{\rho} = (10/9, 1, 1/3)$ . Since  $\boldsymbol{\pi}$  is balanced, both the invariant and handicap-based methods assign  $(1/3, 1/3, 1/3)$  to  $\boldsymbol{\pi}$ . For  $\boldsymbol{\pi}'$ , the handicap-based ranking is proportional to  $(10/9, 1, 1/3)$  by homogeneity and the invariant ranking is (approximately)  $(0.38, 0.395, 0.225)$  to  $\boldsymbol{\pi}'$ . Thus, although the statements on 1 relative to 2 increase from  $\boldsymbol{\pi}$  to  $\boldsymbol{\pi}'$ , each one multiplied by  $\rho_1 = 10/9$ , the invariant score of 2 becomes larger than 1's in  $\boldsymbol{\pi}'$ .

This non-monotone behavior is explained by the dual role of the scores in the invariant method and the induced loopback effect. Item 1 receives a high valuation from expert 3 at  $\boldsymbol{\pi}$ . As item 3 gets less support (the scaling down by  $\rho_3 = 1/3$ ), its score is decreased, hence its weight as well. As a result, the high valuation from 3 counts less and this produces a negative effect on item 1's score that counterbalances the direct positive effect due to  $\rho_1$ . As for item 2, it receives a high valuation from 1 at  $\boldsymbol{\pi}$ . So it benefits indirectly from both the increase in 1's valuations and the decrease in 3's, since they result in an increase in 1's score and a decrease in 3's score. This explains why the score of 2 not only increases but ends up larger than 1's score.

Finally, note that 1's invariant score may end up larger than 2's if we choose different values for  $\rho_1$  and  $\rho_3$  (still respectively larger and smaller than 1). As should be clear from the above argument, there are effects possibly in opposite directions due



to the double role of the scores in the invariant method. The final order of the scores depends on the relative intensity of these various effects, which, in turn, depends on the matrix and the values of the  $\rho_i$ .

### 3.4 Illustration: Journals Rankings

This section illustrates the differences between the invariant, the handicap-based and the HITS methods for the rankings of 37 journals, using the same data as in Palacios-Huerta and Volij (04). Scores and weights are given per article. Let  $\pi_{i,j} = \frac{C_{i,j}}{C_{+,j}}$  be the total share of citations sent by (all articles of)  $j$  received by (all articles of)  $i$  and  $n_i$  denote the number of articles in journal  $i$ . The invariant ranking  $\mathbf{r}$  per article satisfies

$$r_i n_i = \sum_j \pi_{i,j} r_j n_j \text{ for each } i$$

and the handicap-based scores and weights  $\mathbf{r}$  and  $\mathbf{q}$  satisfy

$$n_i r_i = \sum_j \pi_{i,j} n_j q_j \text{ for each } i \text{ and } \frac{1}{n_j q_j} = \sum_i \frac{\pi_{i,j} n_i}{r_i} \text{ for each } j,$$

i.e.  $(n_i r_i)$  and  $(n_j q_j)$  are the scores and weights of journals.

Table 1 reports the handicap-based, HITS, and invariant rankings in the three first columns, the ratio of the handicap-based score over the invariant one in the fourth, the weights for the handicap-based and HITS methods in the fifth and sixth column and finally the correlation between the overall citations with a journal's citations (i.e. the correlation between  $\sum_j \mathbf{C}_{.,j}$  and  $\mathbf{C}_{.,j}$ ). Rankings are normalized with a constant sum for  $\sum_i r_i n_i$  instead of a constant sum for  $\sum_i r_i$ . Since the weights satisfy  $\sum_i r_i n_i = \sum_i q_i n_i$ , such a normalization gives the same total to weights whatever the used ranking. The results are presented so that  $\sum_i r_i n_i = 100a$  where  $a = (\sum n_i)/37$  is the average number of articles per journal, which yields an average score  $(\sum_i r_i n_i)/(\sum n_i)$  equal to 2.7.

The handicap-based and HITS methods produce similar rankings. Their weights are in line and convey meaningful information and they differ significantly from the invariant weights (which are the invariant scores). Interestingly *Journal of Economic Literature* (JEL) has by far the largest weight overall journals, which supports the meaningfulness of the method, followed by *Review of economic studies* (RES), *Quarterly Journal of Economics* (QJE), and *Rand Journal of Economics* (Rand). As an

	Handicap- Based	HITS	Inv	<u>Hand</u> Inv	$q_{hand}$	$q_{HITS}$	Corr
Quarterly J. Econ	10.018	9.7381	11.4412	0.8756	5.3430	4.6379	0.7272
Econometrica	9.6496	9.7683	11.7444	0.8216	3.4210	3.4836	0.6771
J Econ Literature	9.6480	9.4887	9.2594	1.0420	8.1585	8.4002	0.6116
J. Political Economy	7.2226	7.0028	7.5573	0.9557	4.1462	3.7403	0.8151
A Econ Rev.	7.0113	6.9074	7.5236	0.9319	2.4798	2.5582	0.7873
Rev. of Econ. Studies	5.9858	5.8574	7.4224	0.8064	5.5907	5.4340	0.7224
J. of Monetary Econ.	4.2653	4.1510	5.2428	0.8136	4.3299	3.9331	0.5466
J. of Econ. Perspectives	3.6309	3.4951	3.6507	0.9946	3.6657	3.3049	0.5460
J. of Econ. Theory	3.2109	3.3893	4.0030	0.8021	1.8915	1.9902	0.4856
Rand J. of Econ.	3.1024	2.9416	2.2646	1.3700	4.8855	4.3174	0.4445
Games and Econ. Behavior	3.0573	3.3397	3.7402	0.8174	2.5863	2.7514	0.4168
J. of Financial Econ.	2.9606	2.4102	1.7404	1.7010	3.7100	2.8174	0.1117
J. of Econometrics	2.9066	2.8103	2.4406	1.1909	1.8608	1.7978	0.3742
J. of Labor Econ.	2.8038	2.8306	2.1191	1.3231	3.8302	4.3557	0.4292
J. of Human Resources	2.5879	2.6167	2.0470	1.2643	3.4522	3.3428	0.2456
J. of Risk and Uncertainty	2.4000	2.5413	1.8740	1.2807	3.5808	3.8032	-0.0186
J. of Env. Econ. and Manag.	2.3330	2.2943	1.4148	1.6490	3.1345	2.9811	0.1441
Econometric Theory	2.3312	2.2006	1.8501	1.2601	3.2729	2.7537	0.2903
J. of Bus. and Ec. Stat.	2.1956	2.2257	1.6560	1.3258	3.0493	3.1691	0.4229
Rev. of Econ. and Stat.	2.0609	2.0573	1.8967	1.0865	2.3898	2.3346	0.6789
Social Choice and Welfare	2.0237	2.3635	1.4593	1.3868	3.2979	3.8801	0.1242
J. of Public Econ.	2.0046	2.0367	1.8640	1.0754	2.2278	2.1119	0.6344
J. of Int. Econ.	1.8998	1.9417	1.3240	1.4349	3.0395	3.2132	0.3459
Int. Econ. Rev.	1.8632	1.8813	1.8004	1.0349	3.4277	3.3640	0.8604
J. of Applied Econometrics	1.7499	1.8086	1.4695	1.1909	3.6414	4.0607	0.3097
Econ. J.	1.7492	1.7849	1.3943	1.2546	1.8269	1.7183	0.5090
AER Proceedings	1.6863	1.6471	1.6315	1.0336	2.1315	1.9946	0.5815
European Econ. Rev.	1.6466	1.6414	1.5101	1.0904	1.8436	1.6745	0.7475
Int. J. of Game Theory	1.4510	1.7037	1.4368	1.0099	4.1551	5.3043	0.2916
Econ. Theory	1.4370	1.4534	2.1151	0.6794	2.1391	2.3458	0.4543
J. of Ec. Dyn. and Control	1.1781	1.2508	1.2210	0.9649	2.0642	2.2980	0.4078
J. of Mathematical Econ.	1.0797	1.2176	1.1731	0.9204	1.9834	2.3951	0.0467
Econ. Inquiry	0.8495	0.8886	0.7209	1.1784	3.0400	3.1831	0.6491
J. of Ec. Behavior and Org.	0.8029	0.8895	0.5882	1.3651	1.7462	2.1252	0.4844
Scandinavian J. of Econ.	0.7770	0.8062	0.4717	1.6470	3.8152	4.3015	0.6248
Oxford Bull. of Ecs. and Stat.	0.6922	0.7536	0.3138	2.2058	2.8363	3.3401	0.2086
Econ. Letters	0.4573	0.4739	0.3554	1.2866	0.6645	0.6700	0.4839

Table 1: Rankings and weights per article

illustration of the fact that weights pertain to a different property than the scores, the handicap-based score is roughly equal to its invariant score for JEL, is lower for RES and larger for Rand. The rather low weight of AER, little less than the 'average' in both the handicap-based and HITS methods suggests that AER tends to refer no more than the average to the top journals (see the discussion in section 3.1). The last column illustrates that correlation with the overall citations is not a good indicator of the weights, as could also be expected from our previous discussion.

The top 6 journals remain the same without any ambiguity. The order varies, for example, the order between the two first journals, QJE and *Econometrica* is reversed. However, since their scores are very close between each other in each ranking, the simple ordering amplifies very small differences in scores (which shows the benefit of considering a cardinal ranking). Apart from JEL, the scores of the 5 other journals in the top 6 are lower than in their invariant scores, with the largest decrease for RES. The main intuition for this decrease is that these journals receive proportionately more citations from top journals than the average. Since the weights of top journals are typically lower than their invariant scores, this explains the decrease in their score. JEL instead has a more dispersed scope of citations (again, recall that this is not related to the fact that JEL has a large weight). This type of argument also explains why the scores of most theory journals, except Journal of Mathematical Economics and Social of Choice and Welfare (SCW), decrease because they also receive citations proportionately more from top journals. The journals with the largest handicap-based score relative to their invariant score are *Journal of Financial Economics*, Rand, *Oxford Bulletin of Economics and Statistics*, and SCW.

In my view, one should not pay too much attention to the differences in these rankings; however the weights computed by the handicap-based and the HITS methods convey meaningful information.

## 4 Extending the handicap-method

So far the handicap-based method has been defined for positive statements. In some settings, the positivity assumption is too strong. In the budget example (Section 2), an expert may not want to allocate funds to a particular issue. In a voting context, in which experts are asked to grade the items, they may be allowed to assign a zero

to an item, meaning that they find this item not acceptable, or 'infinitely below' an item with a positive grade. Zero elements in the statement matrix raise difficulties for defining the handicap-based ranking, as is also the case for the invariant method. This section investigates for which matrices these difficulties can be solved. We look for statement matrices for which a handicap-based ranking is well defined, meaning that there is a ranking that satisfies (9) and it is unique. Uniqueness is important because it relates to the continuity of the method: without uniqueness, perturbing a matrix by replacing non-zeros by small but positive elements give different answers depending on the perturbation.

Given a statement matrix  $\boldsymbol{\pi}$ ,  $\pi_{i,j} \geq 0$ , we assume that  $\boldsymbol{\pi}$  has no null row nor null column (these could be deleted). Let us introduce some notation. Let  $I(j) = \{i | \pi_{i,j} > 0\}$  be the set of items cited by  $j$ . For  $J$  subset of  $M$ ,  $I(J) = \cup_{j \in J} I(j)$  is the set of items cited by at least one expert in  $J$ . Consider the bipartite graph  $G$  with sets of nodes  $N$  and  $M$  where  $(i, j)$  is an edge for the  $i$  in  $N$  and  $j$  in  $M$  with positive  $\pi_{i,j}$ . In the peers' setting in which the items are also the experts, the two sets  $N$  and  $M$  are replica of each other so that the bipartite graph distinguishes the two roles as expert or item for each element.  $G$  is *items-connected* if each pair of items is linked by a path: for each pair of items  $i, \ell$ , there is a sequence alternating items and experts, starting at  $i$  and ending at  $\ell$ :  $i = i_0, j_0, i_1, \dots, i_t, j_t, \dots, j_k, i_{k+1} = \ell$  such that each expert  $j_t$  in the sequence cites the two adjacent items  $i_t$  and  $i_{t+1}$ . When  $G$  is *items-connected*,  $G$  is *connected* as well: since each expert assigns a positive statement on one item at least, there is a path between any two elements, item or expert.<sup>14</sup>

Consider a statement matrix for which a ranking as defined by (9) exists. The matrix  $\boldsymbol{p}$  defined by  $p_{i,j} = \frac{1}{r_i} \pi_{i,j} q_j$  is balanced and has exactly the same null cells as  $\boldsymbol{\pi}$ . A necessary condition for the existence of a handicap-based ranking follows: there must exist a balanced matrix that has exactly the same null entries as  $\boldsymbol{\pi}$ . This existence problem can be put as a flow problem solved by supply-demand conditions, as described in the next proposition.

**Proposition 4** *Given a non-negative matrix  $\boldsymbol{\pi}$ , there is a unique ranking  $\mathbf{r} = (r_i)$  in*

---

<sup>14</sup>To obtain a path between two experts for example take two items that are cited respectively by each one, a path between the two experts is obtained by taking the path linking the two items (if they differ), and adding the link between each item and the expert who cites it.

$\Delta_N$  such that (9) holds:

$$\sum_j (\pi_{i,j} q_j) \frac{1}{r_i} = 1 \text{ for each } i \text{ where } \sum_i \left(\frac{\pi_{i,j}}{r_i}\right) q_j = \frac{n}{m} \text{ for each } j$$

if and only if the bipartite graph  $G$  associated to  $\pi$  is items-connected and

$$\frac{1}{m}|J| < \frac{1}{n}|I(J)| \text{ for any strict subset } J \text{ of } M. \quad (14)$$

The items-connectedness of graph  $G$  ensures the uniqueness of the ranking. This is a natural condition. If  $G$  is not items-connected, then there are two disjoint sets of experts who cite two disjoint sets of items: these experts do not share any common interest. In such a situation, it makes sense that no unique ranking can reflect the statements of all experts.

Conditions (14) when  $n = m$  require that the number of items cited by a given subgroup of experts should exceed the number of these experts: cites should be sufficiently 'disseminated'. When  $n = m$  the property is required on the proportion of items in  $N$  and the proportion of experts in  $M$ . Conditions (14) also require the dual property that the proportion of experts who cite a given subset of items should exceed the proportion of these items in the whole set  $N$ : taking the complements to the subsets in  $N$  and  $M$  (14) write<sup>15</sup>

$$|I| < \frac{n}{m}|J(I)| \text{ for any strict subset } I \text{ of } N. \quad (15)$$

To get an intuition for these conditions, observe that a balanced matrix which has null elements for each pair where the element in  $\pi$  is null corresponds to a flow in the bipartite graph  $G$  such that each  $j$  sends  $1/m$  unit and each node  $i$  in  $N$  receives  $1/n$  units. The existence of such a flow is solved by well-known supply-demand conditions, which are the same inequalities as in (14) but only required weakly. Requiring the inequalities to be strict ensure that the flow can be made strictly positive on each link (if not, the flow can be perturbed by having a small enough amount going through a link without flow and finding a cycle without creating zero).

The invariant or the HITS methods also cannot be extended in a unique way to all non-negative matrices. The matrix  $\pi$  or  $\pi\tilde{\pi}$  should admit a unique ranking

---

<sup>15</sup>Given a strict subset  $I$ , let  $J = M - J(I)$ . If  $J$  is empty, (15) is met. If not, apply (14) to  $J$ :  $\frac{1}{m}(m - |J(I)|) < \frac{1}{n}|I(J(I))|$ . By definition, the experts cited by  $M - J(I)$  do not cite any items in  $I$ . So  $I(J(I)) \subset N - I$ , which implies  $|I(J(I))| \leq n - |I|$ , and finally  $|I| < \frac{n}{m}|J(I)|$ .

associated to the largest eigenvalue, as is ensured by its irreducibility<sup>16</sup>, thanks to Perron-Frobenius theorem. Comparing these conditions with those in Proposition 4 shows again that the invariant method behaves quite differently than the HITS and handicap-based methods.

The items-connectedness of  $G$  is equivalent to the irreducibility of  $\boldsymbol{\pi}\tilde{\boldsymbol{\pi}}$ . The element  $(i, \ell)$  of  $\mathbf{a} = \boldsymbol{\pi}\tilde{\boldsymbol{\pi}}$  writes  $\sum_j \pi_{i,j}\pi_{\ell,j}$ , hence is positive if and only if there is an expert who cites both  $i$  and  $\ell$ . Thus  $a_{i,\ell}^{(t)}$  is positive if there is a path with  $t$  experts linking  $i$  to  $\ell$  in the bipartite graph  $G$ . Hence the items-connectedness ensures that the HITS ranking is uniquely defined. Using a similar argument, the experts' (or 'hubs') weights are also well defined: any two experts are linked so that the matrix  $\tilde{\boldsymbol{\pi}}\boldsymbol{\pi}$  is irreducible.

For the invariant method, the ranking is uniquely defined when the matrix  $\boldsymbol{\pi}$  is irreducible. This condition differs from the irreducibility of  $\boldsymbol{\pi}\tilde{\boldsymbol{\pi}}$ , as shown by the following example. There are three items/experts; 1 cites 2, 2 cites 3, and 3 cites 1. The matrix<sup>17</sup>  $\boldsymbol{\pi}$  is irreducible but  $\boldsymbol{\pi}\tilde{\boldsymbol{\pi}}$  is null because each expert cites a different item.

To end this section, let us make a couple of remarks on the interpretation of zeros and the 0-1 setting. We have interpreted a zero as meaning 'non-acceptable'. This is reflected by the homogeneity axiom: multiplying statements on an item does not change its zero. Hence perturbing a zero element of an expert's statement into a small but positive element may change it substantially. However, when the conditions for existence of the handicap-based ranking are satisfied, such perturbations all lead to the same ranking. This explains why the existence conditions are rather strong. In other contexts such as in the Web, a zero (the absence of a link) may not mean 'non-acceptable', but simply that an item is not 'known'. This discussion suggests to be more careful about what a zero means. Depending on the meaning, different axioms might be relevant.

As to the 0-1 setting, the method can be used, but the characterization does not

---

<sup>16</sup>Recall that a non-negative square matrix  $\mathbf{a}$  is irreducible if for each pair  $(i, \ell)$  there is an integer  $t$  such that the  $(i, \ell)$  element of the  $t$ -product matrix  $\mathbf{a}^{(t)}$  is positive.

<sup>17</sup> $\boldsymbol{\pi} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ .

apply because the homogeneity axiom does not make sense: multiplying a row by some constant makes it non-feasible.<sup>18</sup> Furthermore, the existence conditions are strong in that context (since they require many values equal to 1). Nevertheless, it would be interesting to investigate how the method behave on Internet data, using the standard perturbation.

## 5 Concluding remarks

This paper has introduced a new ranking method and characterized it through properties. The handicap-based method is adequate in environments where either the ‘intensity’ of statements is not controlled and intensity-invariance is required, or statements are relative evaluations, as when individuals express their preferences as to how a budget should be allocated between various issues. The handicap-based method is, in a sense made precise in the paper, the counterpart of the counting method in these environments.

Several developments are worth investigating.

First, viewing a ranking method as assigning simultaneously scores to items and weights to experts in an equilibrium relationship provides a useful tool for deriving new interesting methods. How the weights vary with the whole experts’ statements profile reflects how the differences across these statements are treated. In addition, in a context where items and experts coincide, two indices are assigned to each item -a score and a weight- thereby providing more information than a single ranking.

Second, in the situation such as the Web where there are many zeros, most methods need to perturb the data in order to be applied. This raises the question of the robustness of the outcome to such perturbations. A possible route other than perturbation would be to make the set of experts endogenous, not necessarily identical to the set of items to rank.

Finally, the analysis of rankings in a dynamical setting and the extent of their influence are important topics that need to be explored.<sup>19</sup>

---

<sup>18</sup>The same remark applies to the axiomatizations proposed by Palacios-Huerta and Volij (2004) and Slutzki and Volij (2006).

<sup>19</sup>For some analysis on the subject, see Demange 2012-a and 2012-b).

## 6 Proofs

**Proof of Lemma 1** Let  $F$  be intensity-invariant:

$$F([\boldsymbol{\pi}]) = F(\boldsymbol{\pi}) \text{ and } F([dg(\boldsymbol{\rho})\boldsymbol{\pi}]) = F(dg(\boldsymbol{\rho})\boldsymbol{\pi}). \quad (16)$$

Let  $F$  be homogeneous on absolute statements:  $F(dg(\boldsymbol{\rho})\boldsymbol{\pi})$  is proportional to  $dg(\boldsymbol{\rho})F(\boldsymbol{\pi})$  for any  $\boldsymbol{\pi}$  in  $\mathcal{P}$ , in particular for  $\boldsymbol{\pi}$  in  $\mathcal{R}$ . Thus  $F([dg(\boldsymbol{\rho})\boldsymbol{\pi}])$ , which is equal to  $F(dg(\boldsymbol{\rho})\boldsymbol{\pi})$ , is proportional to  $dg(\boldsymbol{\rho})F(\boldsymbol{\pi})$ :  $F$  is homogeneous on relative statements.

To show the converse, let  $F$  be homogeneous on relative statements and  $\boldsymbol{\pi}$  be in  $\mathcal{P}$ . Observe that  $[dg(\boldsymbol{\rho})\boldsymbol{\pi}] = [dg(\boldsymbol{\rho})[\boldsymbol{\pi}]]$ . Hence by (16)  $F(dg(\boldsymbol{\rho})\boldsymbol{\pi}) = F([dg(\boldsymbol{\rho})[\boldsymbol{\pi}]])$ . Since  $[\boldsymbol{\pi}]$  is in  $\mathcal{R}$ , homogeneity on  $\mathcal{R}$  implies that  $F([dg(\boldsymbol{\rho})[\boldsymbol{\pi}]])$  is proportional to  $dg(\boldsymbol{\rho})F([\boldsymbol{\pi}])$ . Using  $F([\boldsymbol{\pi}]) = F(\boldsymbol{\pi})$ , we finally obtain that  $F(dg(\boldsymbol{\rho})\boldsymbol{\pi})$  is proportional to  $dg(\boldsymbol{\rho})F(\boldsymbol{\pi})$ . This proves that  $F$  is homogeneous on absolute statements. ■

**Proof of Proposition 1.** The proof of existence and uniqueness of  $H$  involves two steps.

The first step shows the existence of  $\mathbf{r}$  in the simplex that satisfies (9). As said in the text, this is equivalent to the fact that matrix  $\mathbf{p}$  with general element  $p_{i,j} = \frac{1}{r_i}\pi_{i,j}q_j$  is  $(1, n/m)$ -balanced. The proof relies on the known result about matrix scaling: there is a unique  $(1, n/m)$ -balanced matrix  $\mathbf{p}$  that is obtained from  $\boldsymbol{\pi}$  by multiplication of its rows and its columns by some numbers. A simple proof relies on a convex program. We recall the argument here for positive matrices. Consider the program (ln denotes the natural logarithm):

$$\begin{aligned} \mathcal{P} : \underset{\mathbf{p}}{\text{minimize}} \quad & \sum_{i,j} p_{i,j} [\ln(\pi_{i,j}/p_{i,j}) - 1] \text{ over the } \mathbf{p} = (p_{i,j}) > 0 \\ \text{subject to} \quad & \sum_j p_{i,j} = 1 \text{ for each } i \text{ and } \sum_i p_{i,j} = n/m \text{ for each } j. \end{aligned} \quad (11)$$

The program is convex with a strictly convex objective function and a feasible set with a non empty relative interior. Hence the solution  $\mathbf{p}$  is unique, characterized by the first order conditions on the Lagrangian: There are multipliers  $\alpha_i$  and  $\beta_j$  associated respectively to the constraints (11) on the totals of row  $i$  and column  $j$  such that

$$\ln(\pi_{i,j}) - \ln(p_{i,j}) = \alpha_i + \beta_j. \quad (17)$$

Set  $r_i = \exp(\alpha_i)$  and  $q_j = \exp(-\beta_j)$ , (17) writes  $p_{i,j} = \frac{1}{r_i}\pi_{i,j}q_j$ . Plugging these expressions into the constraints (11) give (9). It remains to show that  $\mathbf{r}$  can be



chosen to be in the simplex. The multipliers  $\boldsymbol{\alpha}$  and  $\boldsymbol{\beta}$  are defined up to an additive constant, i.e.,  $(\alpha_i + c)$ ,  $(\beta_j - c)$  satisfy (17) for any  $c$  if  $\boldsymbol{\alpha}, \boldsymbol{\beta}$  does (this is due to the fact the linear system (11) is formed with linearly dependent equations). Thus there is a vector  $\mathbf{r}$ ,  $r_i = \exp(\alpha_i + c)$ , that belongs to the simplex for an appropriate value of  $c$ .

The second step proves that  $\mathbf{r}$  is unique. Given  $\mathbf{r}$  that satisfies (9), the matrix  $\mathbf{p}$  defined by  $p_{i,j} = \frac{1}{r_i} \pi_{i,j} q_j$  satisfies the constraints (11). Furthermore, taking the log of  $p_{i,j} = \frac{1}{r_i} \pi_{i,j} q_j$ , the first order conditions (17) are met for  $\alpha_i = \ln(r_i)$ ,  $\beta_j = -\ln(q_j)$ . Hence  $\mathbf{p}$  is the unique solution to  $\mathcal{P}$ . So if there are two rankings that satisfy (9),  $\mathbf{r}$  and  $\mathbf{r}'$ , the corresponding values  $(\alpha_i)$   $(\beta_j)$  and  $(\alpha'_i)$   $(\beta'_j)$  satisfy (17) for the same matrix  $\mathbf{p}$ . Taking the difference yields

$$(\alpha_i - \alpha'_i) + (\beta_j - \beta'_j) = 0 \text{ for each } i, j. \quad (18)$$

Hence  $\alpha'_i = \alpha_i + c$  for some scalar  $c$ . There can be only one value for  $c$  so that  $\mathbf{r}$  defined by  $r_i = \exp(\alpha_i)$  belongs to  $\Delta_N$ . This proves that  $H$  is a well defined method. ■

Let us now prove the properties of  $H$

$H$  is intensity-invariant. Let  $\boldsymbol{\pi}'$  be obtained from  $\boldsymbol{\pi}$  by multiplying column  $j$  of  $\boldsymbol{\pi}$  by  $\mu_j$ . Letting  $\mathbf{q}'$  be the vector obtained from  $\mathbf{q}$  by dividing  $q_j$  by  $\mu_j$ , the vectors  $\mathbf{r}$  and  $\mathbf{q}'$  satisfy (9) for  $\boldsymbol{\pi}'$ . By the uniqueness result proved previously,  $H(\boldsymbol{\pi}')$  is equal to  $\mathbf{r}$ .

$H$  is uniform on  $\mathcal{R}$ . Let matrix  $\boldsymbol{\pi}$  in  $\mathcal{R}$  be row-balanced. It satisfies  $\sum_j \pi_{i,j} = m/n$  for each  $i$  and  $\sum_i \pi_{i,j} = 1$  for each  $j$ . Hence, the conditions

$$\sum_j (\pi_{i,j} q_j) \frac{1}{r_i} = 1 \text{ for each } i \text{ where } \sum_i \left( \frac{\pi_{i,j}}{r_i} \right) q_j = \frac{n}{m} \text{ for each } j \quad (9)$$

are satisfied by taking equal scores and equal experts' weights, that is, each  $r_i$  equals to  $1/n$  and each  $q_j$  equals to  $1/m$ . (Equivalently  $\mathbf{p} = \boldsymbol{\pi} \frac{n}{m}$  satisfies (11).) Hence, by the uniqueness result,  $H(\boldsymbol{\pi}) = \mathbf{e}_N$ .

$H$  is exact on  $\mathcal{R}$ . Let  $H(\boldsymbol{\pi}) = \mathbf{e}_N$  for a matrix  $\boldsymbol{\pi}$  in  $\mathcal{R}$ . We need to show that  $\boldsymbol{\pi}$  is row-balanced. Applying  $\mathbf{r} = \mathbf{e}_N$  to the second equation in (9) yields that the weight vector  $\mathbf{q}$  satisfies  $(\sum_i \pi_{i,j}) q_j = 1/m$  for each  $j$ . This implies  $q_j = 1/m$  since  $\boldsymbol{\pi}$  is in  $\mathcal{R}$ . Plugging  $r_i = 1/n$  and  $q_j = 1/m$  for each  $i, j$  into the first set of equations of (9) we obtain that each row's total is equal to  $m/n$ : the matrix  $\boldsymbol{\pi}$  is row-balanced.

$H$  is homogeneous. Let  $\boldsymbol{\pi}' = dg(\boldsymbol{\rho})\boldsymbol{\pi}$  for a positive vector  $\boldsymbol{\rho}$ . By the definition of  $H$ , we have  $\pi_{i,j} = \frac{r_i p_{i,j}}{q_j}$  where  $\mathbf{p}$  is a  $(1, n/m)$ -balanced matrix. Hence  $\pi'_{i,j} = \rho_i \pi_{i,j} = \frac{\rho_i r_i p_{i,j}}{q_j}$  so that  $\mathbf{p}$  is obtained from  $\boldsymbol{\pi}'$  by dividing each row  $i$  by  $\rho_i r_i$  and multiplying each column  $j$  by  $q_j$ . By the uniqueness result stated in the first part of Proposition 1, this implies that the handicap-based ranking  $\mathbf{r}'$  associated to  $\boldsymbol{\pi}'$  is the vector in  $\Delta_N$  proportional to  $(\rho_i r_i)$ . Thus,  $H(\boldsymbol{\pi}') = (\frac{\rho_i r_i}{\sum_i \rho_i r_i})$ , namely multiplying the rows of  $\boldsymbol{\pi}$  by some vector multiplies the scores in the same proportions: this proves homogeneity. ■

### Proof of Proposition 2.

Proof of (a). Let method  $F$  be uniform on  $\mathcal{R}$ , intensity-invariant and homogeneous.

$F$  and  $H$  coincide if they coincide on the set  $\mathcal{R}$  of relative statement matrices since both methods are intensity-invariant. Given  $\boldsymbol{\pi}$  consider  $\mathbf{r} = H(\boldsymbol{\pi})$ ,  $\mathbf{q}$ , and the balanced matrix  $\mathbf{p}$  associated by the handicap-based method,  $p_{i,j} = \frac{q_j \pi_{i,j}}{r_i}$ . By uniformity on  $\mathcal{R}$ ,  $F(\mathbf{p}) = \mathbf{e}_N$ . Let  $\mathbf{p}' = dg(\mathbf{r})\mathbf{p}$ . Homogeneity implies  $F(\mathbf{p}') = \mathbf{r}$ . The normalized matrix of  $\mathbf{p}'$  is  $\boldsymbol{\pi}$ : since matrix  $\boldsymbol{\pi}$  is in  $\mathcal{R}$ ,  $\sum_i \pi_{i,j} = 1$  for each  $j$ , which writes  $\sum_i p_{i,j} r_i = q_j$  thus  $[\mathbf{p}'] = \boldsymbol{\pi}$ . By intensity-invariance  $F([\mathbf{p}']) = F(\mathbf{p}') = \mathbf{r}$ , which finally gives  $F(\boldsymbol{\pi}) = H(\boldsymbol{\pi})$ , the desired result.

Proof of (b). Let method  $F$  be exact on  $\mathcal{R}$ , intensity-invariant and homogeneous.

Given  $\mathbf{r} = F(\boldsymbol{\pi})$ , divide each row  $i$  by  $r_i$  so as to obtain matrix  $\boldsymbol{\pi}' = dg(1/r_1, \dots, 1/r_n)\boldsymbol{\pi}$ . Homogeneity implies that the scores are equalized:  $F(\boldsymbol{\pi}') = \mathbf{e}_N$  thanks to Lemma 1. By intensity-invariance of  $F$  we have  $F([\boldsymbol{\pi}']) = F(\boldsymbol{\pi}') = \mathbf{e}_N$ . Now, exactness on  $\mathcal{R}$  implies that  $[\boldsymbol{\pi}']$  is row-balanced, hence each row sums to  $m/n$ . Since  $[\boldsymbol{\pi}']_{i,j} = \frac{\pi_{i,j}}{r_i \sum_\ell \pi_{\ell,j}/r_\ell}$ , this writes

$$\sum_j \frac{\pi_{i,j}}{r_i \sum_\ell \pi_{\ell,j}/r_\ell} = \frac{m}{n} \text{ for each } i,$$

$$\text{or } r_i = \sum_j \pi_{i,j} q_j \text{ for each } i, \text{ where } \frac{1}{q_j} = \frac{m}{n} \sum_\ell \frac{\pi_{\ell,j}}{r_\ell} \text{ for each } j.$$

Thus  $\mathbf{r}$  and  $\mathbf{q}$  satisfy (9). Since  $\mathbf{r}$  is in  $\Delta_N$ ,  $\mathbf{r}$  is equal to  $H(\boldsymbol{\pi})$ , the desired property. ■

### Proof of Proposition 3.

Proof of (a). Clearly the counting method is homogeneous on absolute statements and uniform on  $\mathcal{P}$ . To show the reverse, let method  $F$  satisfies these properties. Given a matrix  $\boldsymbol{\pi}$ , divide each row  $i$  by its total  $\pi_{i+} = \sum_{j \in M} \pi_{i,j}$  and denote by  $\boldsymbol{\pi}'$  the obtained matrix:  $\boldsymbol{\pi}' = dg(1/\pi_{1+}, \dots, 1/\pi_{n+})\boldsymbol{\pi}$ . By the a-homogeneity of  $F$ , the

ranking assigned by  $F$  to  $\boldsymbol{\pi}'$  is obtained by dividing each component  $i$  of  $F(\boldsymbol{\pi})$  by  $\pi_{i+}$  and normalizing: for some positive  $\lambda$ ,  $F_i(\boldsymbol{\pi}') = \lambda F_i(\boldsymbol{\pi})/\pi_{i+}$  for each  $i$ . Since  $\boldsymbol{\pi}'$  is row-balanced  $F(\boldsymbol{\pi}') = \mathbf{e}_N$ . This yields that  $F_i(\boldsymbol{\pi})/\pi_{i+}$  is constant across  $i$ :  $F(\boldsymbol{\pi})$  is the counting ranking of  $\boldsymbol{\pi}$ , the desired result.

Proof of (b). Clearly the counting method is homogeneous on absolute statements and exact on  $\mathcal{P}$ . To show the reverse, let method  $F$  satisfy these properties. Given  $\mathbf{r} = F(\boldsymbol{\pi})$ , divide each row  $i$  by  $r_i$  so as to obtain matrix  $\boldsymbol{\pi}' = dg(1/r_1, \dots, 1/r_n)\boldsymbol{\pi}$ .  $F$  is a-homogeneity implies that the scores are equalized:  $F(\boldsymbol{\pi}') = \mathbf{e}_N$  thanks to Lemma 1. Exactness implies that  $\boldsymbol{\pi}'$  is row-balanced: for some positive  $\lambda$ ,  $\sum_j \pi'_{i,j} = \lambda$  for each  $i$ . Hence  $\sum_j \pi_{i,j} = \lambda r_i$  for each  $i$ :  $\mathbf{r}$  is the counting ranking of  $\boldsymbol{\pi}$ , the desired result. ■

**Proof of Proposition 4** Under the stated conditions, the proof of Proposition 1 extends to a matrix  $\boldsymbol{\pi}$  with some null elements as follows. The objective in  $\mathcal{P}$  takes the sum over the  $(i, j)$  for which  $\pi_{i,j}$  is positive, namely the  $(i, j)$  in  $G$ :

$$\begin{aligned} \mathcal{P} : \underset{\mathbf{p}}{\text{minimize}} \quad & \sum_{i,j \in G} p_{i,j} [\ln(\pi_{i,j}/p_{i,j}) - 1] \text{ over the } \mathbf{p} = (p_{i,j}) \geq 0 \\ \text{subject to} \quad & \sum_j p_{i,j} = 1 \text{ for each } i \text{ and } \sum_i p_{i,j} = n/m \text{ for each } j. \end{aligned} \quad (11)$$

Under conditions (14), the feasible set defined by (11) has a nonempty interior. Hence the Kuhn and Tucker theorem applies. A solution  $\mathbf{p}$  is associated to multipliers  $\alpha_i$  and  $\beta_j$  such that

$$\ln(\pi_{i,j}) - \ln(p_{i,j}) = \alpha_i + \beta_j, \text{ for each } (i, j) \in G \quad (19)$$

Following the same arguments as in Proposition 1, setting  $r_i = \exp(\alpha_i)$  for each  $i$  up to a multiplicative constant ensures the existence of  $\mathbf{r}$  that satisfies (9).

Let us consider uniqueness. Let  $\mathbf{r}$  a solution to the conditions (9) on a handicap-based ranking.  $\mathbf{r}$  is associated to a balanced matrix  $\mathbf{p}$  and the values defined by  $\alpha_i = \ln(r_i), \beta_j = -\ln(q_j)$  satisfy (19). Hence  $\mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}$  satisfy the Kuhn and Tucker conditions associated to  $\mathcal{P}$ . Since the program  $\mathcal{P}$  is strictly convex a solution  $\mathbf{p}$  is unique. Thus, if there are two rankings solutions to (9),  $\mathbf{r}$  and  $\mathbf{r}'$ , taking the difference in (19) for their corresponding values yields

$$(\alpha_i - \alpha'_i) + (\beta_j - \beta'_j) = 0 \text{ for each } (i, j) \in G. \quad (20)$$

This implies that along a path linking two items the values  $(\alpha_i - \alpha'_i)$  are all equal. Using the same argument as for Proposition 1, the uniqueness of  $\mathbf{r}$  follows if  $G$  is connected:  $\alpha'_i - \alpha_i$  is constant across all  $i$ , and there can be only one  $\mathbf{r}$  defined by  $r_i = \exp(\alpha_i)$  that belongs to  $\Delta_N$ . On the other hand, if  $G$  is not connected, there are no links between the values on each component and uniqueness fails. ■

## References

- Altman A. and M. Tennenholtz (2005) “Ranking systems: the PageRank axioms” EC '05 Proceedings of the 6th ACM conference on Electronic commerce.
- Altman A. and M. Tennenholtz (2008) “Axiomatic foundations of ranking systems”, *Journal of Artificial Intelligence Research*, 31(1), 473-495.
- Amir R. (2002) “Impact-adjusted Citations as a measure of Journal quality”, CORE discussion paper 74.
- Bacharach M. (1965) “Estimating Nonnegative Matrices from Marginal Data”, *International Economic Review*, 6(3), 294-310.
- Balinski M.L. and G. Demange (1989) “Algorithm for Proportional Matrices in Reals and Integers”, *Mathematical Programming*, 45(1), 193-210.
- Bonacich P. (1987) “Power and centrality: a family of measures”, *American Journal of Sociology*, 92(5), 1170-1182.
- Brin S., and L. Page (1998) “The anatomy of large-scale hypertextual web search engine”, *Comp Networks and ISDN systems*, 30(1-7), 107-117.
- van den Brink R. and R. P. Gilles (2003) “Ranking by outdegree for directed graphs”, *Discrete Mathematics* 271, 261–270.
- van den Brink R. and R. P. Gilles (2009) “The outflow ranking method for weighted directed graphs” *European Journal of Operational Research* 193, 484–491.
- de Clippel G., H. Moulin, and N. Tideman (2008) “Impartial division of a dollar”, *Journal of Economic Theory*, 139(1), 176-191.
- Demange, G. (2012) “On the influence of a ranking system”, *Social Choice and Welfare*, Vol 39, Issue 2, 431-455.

- Demange, G. (2012) “Collective attention and ranking methods”, forthcoming in *Journal of Dynamics and Games*.
- Du, Y., Lehrer E. and Pauzner (2012) “Competitive economy as a ranking device over networks”, mimeo.
- Katz L. (1953) “A new status index derived from sociometric analysis”, *Psychometrika*, 18(1), 39-43.
- Kleinberg N. (1999) “Authoritative sources in a hyperlinked environment”, *Journal of the ACM*, 46(5), 604-632.
- Liebowitz, S. J., and J. C. Palmer (1984) “Assessing the Relative Impacts of Economics Journals”, *Journal of Economic Literature*, 22(1), 77-88.
- Palacios-Huerta, I., and O. Volij (2004) “The Measurement of Intellectual Influence”, *Econometrica*, 72(3), 963-977.
- Pandey S., S. Roy, C. Olston, J. Cho, and S. Chakrabarti (2005) “Shuffling a Stacked Deck: The Case for Partially Randomized Ranking of Search Engine Results”, *Proceedings of the 31st international conference on Very large data bases*, 781-792.
- Pinski, G., and F. Narin (1976) “Citation Influence for Journal Aggregates of Scientific Publications: Theory, with Application to the Literature of Physics”, *Information Processing and Management*, 12(5), 297-312.
- Pukelsheim F. (2012) “Biproportional Matrix Scaling and the Iterative Proportional Fitting procedure” Augsburg preprint.
- Rubinstein, A. (1980) “Ranking the Participants in a Tournament” *SIAM Journal on Applied Mathematics*, Vol. 38, No. 1, 108-111.
- Slutzki G. and O. Volij (2006) “Scoring of Web pages and tournaments-axiomatizations”, *Social choice and welfare*, 26(1), 75-92.
- Soules G. (1991) The Rate of Convergence of Sinkhorn Balancing *Linear algebra and its applications* 150, 3-40.
- de Sousa Santos, B. (1998) “Participatory budgeting in Porto Alegre: toward a redistributive democracy” *Politics and Society*, 26, 461-510.

Woeginger G.J. (2008) “An axiomatic characterization of the Hirsch-index”, *Mathematical Social Sciences*, 56(2), 224-232.