Collective attention and ranking methods

Gabrielle Demange

PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

WORKING PAPER N° 2010 – 42

Collective attention and ranking methods

Gabrielle Demange

JEL Codes: D71, D83

Keywords: Ranking ; Scoring ; Invariant method ; Peers' method ; Attention ; Handicap ; Scaling matrix ; Dynamics through influence

Collective attention and ranking methods

Gabrielle DEMANGE[1]

May 7, 2012

Abstract

In a world with a tremendous amount of choices, ranking systems are becoming increasingly important in helping individuals to find information relevant to them. As such, rankings play a crucial role of influencing the attention that is devoted to the various alternatives. This role generates a feedback when the ranking is based on citations, as is the case for PageRank used by Google. The attention bias due to published rankings affects new stated opinions (citations), which will, in turn, affect the next ranking. The purpose of this paper is to investigate this feedback by studying some simple but reasonable dynamics. We show that the long run behavior of the process much depends on the preferences, in particular on their diversity, and on the used ranking method. Two main families of methods are investigated, one based on the notion of 'handicaps', the other one on the notion of peers' rankings.

Keywords ranking, scoring, invariant method, peers' method, attention, handicap, scaling matrix, dynamics through influence.

1 Introduction

The use of rankings is becoming pervasive in many areas, including the Web environment for ranking pages and academia for ranking researchers, journals, and universities. The public good aspect of information explains the use of rankings. The determination of a ranking involves a costly process of gathering and summarizing some relevant information on the alternatives in a particular topic. For example, the extremely popular PageRank used by Google is based on the huge hyperlink structure, in which a link from a page towards another one is interpreted as a (positive) reference. The publication of the ranking allows the individuals who find the information relevant to save on search costs. For that very reason, rankings have some influence on the attention that is devoted to the various alternatives. Presumably, the attention bias will affect the new statements (citations, or

[1]PSE-EHESS, address 48 bd Jourdan, 75014 Paris, France e-mail demange@pse.ens.fr.

links, or votes), which will, in turn, affect the subsequent ranking. The purpose of this paper is to investigate this feedback between rankings, attention intensities, and statements by studying some reasonable dynamics.

Let us first describe the ranking problems considered here. A ranking problem is described by a set of items to be ranked and a set of 'experts' who provide statements on the items on which the ranking will be based. A ranking is cardinal up to a multiplicative scalar, meaning that a ranking assigns relative scores to items. Let us describe some prominent problems. In a ranking of journals based on citations, journals are both the items to be ranked and the experts; a journal's statement is derived from the number of citations made by its articles towards articles published in the different journals. A similar structure applies for ranking Web pages on the basis of the links, taking a page's statements to be its links towards other pages. Here also the items to be ranked -the Web pages- coincide with the experts. This is not the case in our third example, a representation problem. The problem is to allocate voting weights in an assembly to parties as a result of the votes of various electoral bodies, regional for example. Here the items are the parties, the experts are the regions, and a region's statement is given by the number of votes gathered by each party in that region.

A ranking method assigns a ranking to the statements provided by the experts. It can thus be viewed as solving a multi-criteria problem seeking to aggregate preferences for instance. As a result, various methods are meaningful, none of which dominating the other. The counting method ranks items according to their received totals. The invariant method, on which Page Rank is based, determines which pages are influential on the basis that a page is influential if it is heavily cited by other influential pages. By its very definition the invariant method assigns weights to the experts in such a way that the weight as an expert is equal to the score as an item.[2] Other methods can be built by assigning not only scores to the items but also weights to the experts so that they are in an equilibrium relationship: the ranking is 'supported by weights' (Demange 2012-b).[3] The analysis considers such methods. The axiomatization approach of social choice theory provides a guide to evaluate and compare the various methods, typically on the basis of their behavior with respect to statements in a static framework.[4] The paper instead considers a dynamic framework, and takes a different perspective, rooted in the influence that published rankings have on attention.

Attention intensities describe a possible bias, independent of the preferences, in the care with which alternatives are assessed. The dynamics are built on two ingredients. The first one is a simple model linking statements to preferences and attention intensities. The statements provided

[2]That this uniquely defines the method is not trivial, and relies on the Perron Frobenious theorem, as explained in section 2.2.

[3]The counting method also assigns weights to the experts, identical whatever their statements.

[4]Recent contributions in the case of peers settings are Palacios-Huerta and Volij (2004), Slutzki and Volij (2006), Altman and Tennenholtz (2005), Clippel, Moulin, and Tideman (2008).

by an expert are both determined by his attention but also by 'intrinsic' preferences. The second ingredient is that rankings play the crucial role of modifying attention intensities. In a context in which the number of alternatives to examine is huge, experts cannot carefully assess each one and tend to pay more attention to those whose score is higher. For example, while working on a paper, a researcher tends to read more the journals whose ranks are higher. This is described by an 'influence function', which specifies how the current ranking modifies experts' attention intensities. In a recurrent framework, the two ingredients just described induce a joint dynamics for rankings and statements: the current ranking modifies attention intensities, hence the next statements on which next ranking is based. Our aim is to investigate how the dynamics is affected by the preferences and the ranking method.

A first intuition is that, as past statements have an impact on future statements through rankings computation, we might expect 'the rich to get richer'. For example, journals with a lot of past citations are more likely to be cited again, which may result in an improvement in their scores. However, the experts' statements depend not only on attention intensities but also on preferences. The strength of the snowball effect is mitigated by the diversity of preferences. When preferences are similar, the score of the item preferred by all experts indeed becomes arbitrarily large relative to others. However, when preferences are diverse,several items may keep a positive score and furthemore the long run outcome can be drastically affected by the used method. Contrasted results are obtained for two different classes of methods. The first class includes the counting method and the handicap-based method (introduced in Demange (2012-b)). These methods guarantee that, given preferences for the experts, there is a unique possible stable limit point (see more precisely Proposition 1).

The second class is the class of peers' methods, as defined in the paper. Peers' methods require the sets of experts and items to coincide but not only. The rationale behind a peers' method is that the ability of an individual to perform (measured by his score) is correlated with his ability to judge others' performance. In particular, for a method supported by weights, a minimal requirement is that an item which obtains a small score is assigned a small expert's weight. This defines a peers' method. The invariant method is a peers' method since scores and experts' weights are equalized. In Demange (2012-a), I show that the dynamics for the invariant method may admit multiple limit points. According to our results, such multiplicity is bound to happen: Whatever peers' method, the dynamics may admit multiple limit points for some preferences, each one corresponding to a different support (the support is the subset of items that keep a positive score). Furthermore, the supports of the limit points are independent of the peers' method. Such result illustrates the self-enforcing aspect of a peers' method: a subgroup of experts that cite themselves more than their outsiders may obtain high scores supported by high experts' weights, independently of the preferences of the outsiders. Self-enforceability here is not obtained through plain manipulation but through the coordination device induced by the influence of the ranking on attention.

An influence function is a crude but simple way of modeling some form of bounded rationality or myopic behavior due to search costs or persuasion bias for example. An alternative framework is to base the influence through a Bayesian reasoning, which requires to make explicit the type of information individuals are searching for. Not only a bayesian framework is much more complex in the problems we are interested in because information is typically multi-dimensional, but also there is evidence that people might not be Bayesian. In particular, the marketing literature puts forward a persuasion bias, according to which the repetition of the same information has an effect (see deMarzo, Vayanos, and Zwiebel (2003) for a discussion and references on this aspect). An influence function may be interpreted as modeling a persuasion bias, in which a higher score is akin to a high repetition of the same piece of information.

This paper is about the convergence of behaviors and statements. This is also the concern of the large literature that analyzes the influence of opinions channelled by 'neighbors' in a network. Most studies consider a situation is one in which individuals receive private signals about a state of the world and repeatedly communicate with their neighbors. The influence of communicated opinions may be driven by bayesian revision (see e.g. Goyal (2005) and the references therein) or specified by some up-dating rule as initiated by DeGroot (2008) (see Golub and Jackson (2008) for a recent development). A main question is whether (non-strategic) communication will lead opinions to converge to a common belief and, if they converge, how the limit belief relates to the initial opinions and the network structure. In our setting instead, information, which is embodied in the ranking, is made public, available to all. This is precisely the source of the coordination of attention.

Researchers in computer science have also concerns about the influence of the rankings provided by search engines. The main criticism is that rankings are biased towards already popular Web pages, thus preventing the rise in popularity of recently created high quality pages. Proposals to correct the bias, such as introducing some randomness in the rankings (Pandey et al. (2005)), or to account for the date of creation of a page in the computation of the ranking (Cho, Roy, and Adams (2005)) have been advanced.

The paper is organized as follows. Next section presents ranking methods, gives examples, and defines some properties. Section 3 is devoted to the dynamics under a linear influence function for the generalized handicap-based methods and the peers' methods. Section 4 presents some extensions and concluding remarks. Some technical proofs are given in Section 5.

2 Ranking methods

This section describes the framework and presents some known methods.

2.1 The framework

A number of items have to be ranked or scored on the basis of the evaluation of a group of experts.

Items can be individuals, journals, articles, political parties. Experts can be judges, voters, or the items themselves in the case of judgment by peers. Let $N = \{1, \cdots, n\}$ be the set of n items to be scored. Let $M = \{1, \cdots, m\}$ be the set of m experts. 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 some evaluations, gathered in a statement matrix, on which the ranking of the items will be based. The statement of expert j is described by a n-vector, $\boldsymbol{\pi}_{.j} = (\pi_{i,j})$, in which $\pi_{i,j}$ is the evaluation of j over item i. The $n \times m$ matrix $\boldsymbol{\pi} = (\pi_{i,j})$ in which j's column is $\boldsymbol{\pi}_{.j}$ is called the *statement matrix*. Statement matrices are first restricted to be positive, that is $\pi_{i,j}$ to be all positive.

The settings described in the introduction are cast into this framework. For ranking journals based on citations, N and M are both given by the set of journals to be compared, statements are the number of citations by articles in journal j towards 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: this gives matrix $\boldsymbol{\pi}$ in which the value $\pi_{i,j} = \frac{C_{i,j}}{n_j}$ is the average number of references of an article from j to i. For ranking Web pages based on the link structure, the two sets of items and experts N and M coincide with the set of 'relevant' pages. Hence the statement matrix $\boldsymbol{\pi}$ is the adjacent matrix of the Web network: it has $\pi_{i,j}$ equal to 1 if page j points to i and 0 otherwise. The matrix has many zeros because many pages are not pointing to each other. (Often, a perturbation technique makes the matrix positive.)

Given statements, one seeks for a ranking that evaluates the relative strength of the items. Specifically, a ranking assigns to each i a non-negative number r_i, called the *score* of i. Since only the relative values matter, the sum of the scores can be normalized to 1. A method assigns a ranking to each possible statement matrix. This yields the following definition.

Definition 1 *A ranking of N is described by a vector $\mathbf{r}$ in the simplex $\Delta_N = \{\mathbf{r} = (r_i) \in \Re^n, r_i \geq 0, \sum_i r_i = 1\}$. Given a set of experts M, a ranking method F assigns to each positive $n \times m$ matrix $\boldsymbol{\pi}$ a ranking $\mathbf{r} = F(\boldsymbol{\pi})$ in Δ_N.*

In the sequel, $\boldsymbol{y} \propto \boldsymbol{x}$ indicates that the two vectors $\boldsymbol{y}$ and $\boldsymbol{x}$ are proportional. Observe that $\mathbf{r} \propto \boldsymbol{x}$ uniquely defines the ranking $\mathbf{r}$ proportional to a positive n-vector $\boldsymbol{x}$.

The counting method is the simplest method; it assigns scores to items in proportion to their received totals. Let π_{i+} denote the total of i's row: $\pi_{i+} = \sum_{j\in M} \pi_{i,j}$.

$$\mathbf{r} \propto \boldsymbol{\pi}_{.+} \text{ or } r_i = (\frac{1}{\sum_{\ell,j} \pi_{\ell,j}}) \sum_{j\in M} \pi_{i,j} \text{ for each } i \tag{1}$$

It turns out that most methods can be viewed as assigning not only scores to items but also weights to experts, as illustrated in Section 2.2. The counting method for instance assigns equal weights to the experts, whatever their statements. Formally[5]

Definition 2 *A method F is supported by weights if it assigns weights to experts, $Q^F = (Q^F_j)$, where each Q^F_j is a positive function defined over the set of positive matrices, so that for each $\boldsymbol{\pi}$*

$$F_i(\boldsymbol{\pi}) = \sum_{j\in M} \pi_{i,j} Q^F_j(\boldsymbol{\pi}) \text{ for each } i. \tag{2}$$

According to (2) the ranking $F(\boldsymbol{\pi})$ is a combination of the experts' statements. In the methods described in next section, the ranking and the supporting weights form an equilibrium relationship. In line with this interpretation, the weights are well defined, even when there are multiple ways to write the ranking as a combination of the columns. For the counting method for example, the equilibrium relationship is trivial, based on the premise that no distinction should be made between experts whatsoever. For m larger than n, experts' statements are linearly dependent, and the counting ranking can be written as a combination of the statements in many different ways. The weights, however, are well defined, all equal to $1/\sum_{\ell,j} \pi_{\ell,j}$.

The property of supporting weights is useful to define new methods by adjusting the weights as defined in section 2.2 or to define properly what a peers' method is (see section 3.3).

Normalized matrices We will deal with matrices whose column's sums are fixed. In the sequel, each (fixed) column's sum takes value 1, $\pi_{+j} = 1$. This has no impact on the analysis: the column's sum of each expert can take a value distinct from one and distinct across experts. There are two independent justifications for considering fixed column's sums. The first one is linked with the dynamics on statement matrices, as explained in Section 3. The second one is that for some methods, the ranking is not affected by the 'intensity' of an expert's statement, namely by the total of its evaluations. Formally a method F is **intensity-invariant** if $F(\boldsymbol{\pi}') = F(\boldsymbol{\pi})$ for $\boldsymbol{\pi}'$ the matrix obtained from $\boldsymbol{\pi}$ by multiplying each column j by any positive μ_j. As a result, totals can be scaled to 1 for example by dividing column j by the total π_{+j}, $\pi_{+j} = \sum_{i\in N} \pi_{i,j}$. Thus, intensity invariance is equivalent to

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

[5]This notion is introduced in Demange (2012-b) and used to define the handicap-based method.

Intensity-invariance is a desirable property in various contexts (see Demange (2012-b)). Though we will not require intensity-invariance, it is important to check that our analysis allows for intensity-invariance, since the property is often required. Intensity-invariant methods are easily constructed by first 'factoring out' intensity in the expert's statement. Specifically the *intensity-invariant* version $[F]$ of method F is defined by setting $[F](\boldsymbol{\pi}) = F([\boldsymbol{\pi}])$ for each $\boldsymbol{\pi}$. In the case of journals for example, factoring out reference intensity avoids to introduce bias due to the fact that the average number of cites per article differs across journals.[6] In the Web environment, one deflates a link from a site by the total number of links from that site.

A matrix whose columns sums are equal to 1 is called *normalized.* The set of positive normalized matrices is denoted by $\mathcal{M}$. Observe that the experts weights satisfy $\sum_j \pi_{+j} Q_j^F(\boldsymbol{\pi}) = 1$.[7] So for a normalized matrix, they sum to 1, i.e. belong to the simplex Δ_M.

2.2 The invariant, Hits, and handicap-based methods

In opposite to the counting method, the following methods differentiate the experts depending on the statement matrices.

The Liebowitz-Palmer (LP)[8] and its intensity-invariant version, called the invariant method, are defined in a peer's context when items and experts coincide ($N = M$). The premise of the methods is that the statements made by a peer as an expert should be weighted by the received score as an item. This induces a loopback effect: a score of an item is defined as proportional to the sum of the received shares weighted by the experts' scores. Specifically the LP method looks for $\mathbf{r}$ in Δ_N that satisfies

$$\text{for some positive } \lambda,\ r_i = \lambda \sum_{j \in N} \pi_{i,j} r_j \text{ for each } i. \tag{4}$$

Equations (4) say that 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. The method is thus well defined.

[6] Recall that $\pi_{i,j}$ is the average number of references of an article from j to i. The j's statement vector in $[\pi]$ describes in which proportion the cites made by an article in j are received by i on average. See Palacios-Huerta and Volij (2004) for an analysis of the impact of cite intensities on the ranking of economic journals.

[7] To see this, let $\mathbf{r} = F(\boldsymbol{\pi})$ be in Δ_N and $\boldsymbol{q} = Q^F(\boldsymbol{\pi})$ satisfy the relationships $r_i = \sum_{j \in M} \pi_{i,j} q_j$ for each i in N. Summing over i and exchanging sums yields:

$$\sum_i r_i = \sum_i \sum_j \pi_{i,j} q_j = \sum_j (\sum_i \pi_{i,j}) q_j = \sum_j \pi_{+j} q_j.$$

[8] This terminology refers to the work of Liebowitz and Palmer (1984), who use an approximation of the method for ranking economic journals. The methods and some variants have been (re)defined and used in different contexts: in sociology by Katz (1953) and Bonacich (1987), in academics for ranking journals by Pinski and Narin (1976), and on the Web for ranking pages using the link structure between pages by Brin and Page (1998).

The **invariant** method is the intensity-invariant version of the LP method, obtained by factoring out intensity. Since a normalized matrix has its largest eigenvalue equal to one[9] the invariant ranking of $\boldsymbol{\pi}$ is the unique $\mathbf{r}$ in the simplex that satisfies

$$r_i = \sum_{j \in N} [\pi]_{i,j} r_j \text{ for each } i. \tag{5}$$

The score of an item is equal to the sum of the received shares weighted by the experts' scores. By its very definition, the invariant method assigns weights to experts in such a way that the scores are equalized to the weights.

The **Hits** method, introduced by Kleinberg (1999), also ranks Web pages on the basis of their link structure. Given a relevant set of pages, N, the Hits method defines a ranking of these pages, based on the links within N. Thus, as for the invariant method, the two sets of items and experts coincide. The method distinguishes two weights for each 'page', one associated with the relevance or authority of a page, the other with the adequacy of a page to point towards 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 be not useful to Internet users). Specifically the method assigns the ranking $\mathbf{r}$ and the experts weights $\boldsymbol{q}$ that satisfy for some positive λ

$$r_i = \sum_j \pi_{i,j} q_j \text{ for each } i \text{ and } q_j = \lambda \sum_i \pi_{i,j} r_i \text{ for each } j. \tag{6}$$

In matrix form, (6) writes as $\mathbf{r} = \boldsymbol{\pi}\boldsymbol{q}$ and $\boldsymbol{q} = \lambda\widetilde{\boldsymbol{\pi}}\mathbf{r}$ where $\widetilde{\boldsymbol{\pi}}$ is the transpose of $\boldsymbol{\pi}$. Thus the 'authority' weights $\mathbf{r}$ and the 'hub' weights $\boldsymbol{q}$ are well defined as respectively the normalized principal eigenvectors of the positive matrices $\boldsymbol{\pi}\widetilde{\boldsymbol{\pi}}$ and $\widetilde{\boldsymbol{\pi}}\boldsymbol{\pi}$.

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. Although the two sets of items and experts coincide, the Hits method is not qualified as a peers' method according to the definition introduced in Section 3.3. The reason is that item scores and expert's weights may widely differ and a main purpose of the Hits method is precisely to allow this distinction.

The **handicap-based** method, introduced in Demange (2012-b), is based on handicaps. 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 ℓ means that the score of i is half that of ℓ. So a handicap vector $\boldsymbol{h}$ is related to a ranking $\mathbf{r}$ through the relationship $h_i = 1/r_i$.

[9] Recall that the eigenvalues of a matrix and of its transpose are identical. The set of equations $\sum_i \pi_{i,j} = 1$ for each j implies that $\mathbb{1}_N$, the n-vector with components equal to 1, is a positive eigenvector of the transpose of $\boldsymbol{\pi}$ with eigenvalue 1.

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. This leads to a well defined method. Specifically, the handicap-based method H assigns the unique ranking $\mathbf{r}$ and unique experts weights $\boldsymbol{q}$ that satisfy

$$\sum_j (\pi_{i,j} q_j) \frac{1}{r_i} = 1 \text{ for each } i \text{ and } \sum_i (\frac{\pi_{i,j}}{r_i}) q_j = \frac{n}{m} \text{ for each } j. \tag{7}$$

The values for the handicap-based ranking and the associated experts' weights can be computed through an iterative process, similar to the iterative scaling algorithm used in the RAS model (Bacharach 1965). The algorithm assigns alternately handicaps so as to equalize the weighted counts across items and experts' weights so as to equalize the total of their distributed handicap points.

H is intensity-invariant and furthermore enjoys nice properties, as developed in Demange (2012-b). In particular the method is homogeneous, meaning that multiplying the relative statements to i by a positive scalar r_i multiplies i's score relative to other items by r_i. This follows from (see Demange (2012-b) or the proof of Lemma 1 in the appendix)

$$\begin{aligned} H_i([dg(\mathbf{r})\boldsymbol{\pi}]) &= \frac{r_i r_i^h}{\sum_\ell r_\ell r_\ell^h} \text{ for each } i,\ Q_j^H([dg(\mathbf{r})\boldsymbol{\pi}]) = \frac{q_j^h s_j}{\sum_\ell r_\ell r_\ell^h} \text{ for each } j \\ \text{where } \mathbf{r}^h &= H(\boldsymbol{\pi}),\ \boldsymbol{q}^h = Q^H(\boldsymbol{\pi}), s_j = \sum_i \pi_{i,j} r_i. \end{aligned} \tag{8}$$

According to (8), the ranking associated to $[dg(\mathbf{r})\boldsymbol{\pi}]$ is proportional to $(r_i r_i^h)$: this is the homogeneity property.

Transformed methods A method supported by weights can be transformed into another method by adjusting the weights through a function. Specifically, let F be supported by Q^F and g be a positive scalar function defined over $[0, 1]$. F is transformed by g into method G by assigning to each $\boldsymbol{\pi}$ experts' weights that are proportional to $g(Q_j^F(\boldsymbol{\pi}))$. This gives

$$\text{for each } \boldsymbol{\pi} \text{ in } \mathcal{M},\ G_i(\boldsymbol{\pi}) = \sum_{j \in M} \pi_{i,j} \frac{g(q_j)}{\sum_{k \in M} g(q_k)} \text{ for each } i \in N \text{ where } \boldsymbol{q} = Q^F(\boldsymbol{\pi}). \tag{9}$$

One may want to put restrictions on g, as we will see in next section. An interesting family of methods is obtained from F by letting g to be homogeneous and increasing: $g(x) = x^{1-\gamma}$ for some γ smaller than 1. This family contains the method F ($\gamma = 0$) and the counting method ($\gamma = 1$). This shows that there are many possible methods.

3 Dynamics

A premise of ranking methods is that statements are related to preferences. Citations or links are considered as positive votes. Even so, the absence of a citation to an article is not necessarily a negative vote because the paper might not have been read. In a context with many alternatives (potentially many relevant papers to read, many sites to visit) experts are not considering each alternative, or they are not devoting the same amount of attention to each one. We build a simple model in which a published ranking modifies attention independently of the preferences.

3.1 The influence model

The influence of a ranking is described by an 'influence function' that assigns attention intensities to a ranking. Let us first define attention intensities.

Attention intensities are described by a positive n-vector $\mathbf{b} = (b_i)$, where b_i represents the intensity spent on i. In the context of journals for example, b_i represents the proportion of the read articles that are in journal i. When attention differs across two journals, the articles in the journal with the higher attention intensity have more chances to be read, everything equal. In an electoral problem, in which statements are the results of an election, attention represents the time spent by voters on listening to the parties.

Attention intensities modify the relative evaluation of an expert, with a total[10] kept constant. These totals can be taken equal to 1 for each expert, without impact on the analysis. Let us interpret $\boldsymbol{\pi}_{.j} = (\pi_{i,j})_{i\in N}$ as j's 'true' preferences, that is the statement of j if j evaluates each i with equal attention. For example $\pi_{i,j}$ is the (average) proportion of the cites that authors in j make to an article in i *conditional* on reading all relevant articles (or alternatively of selecting them at random in an unbiased way). Bias in attention $\mathbf{b}$ results in statements proportional to $(\pi_{i,j}b_i)$. This results in a statement matrix given by

$$\pi'_{i,j} = \frac{\pi_{i,j}b_i}{\sum_i \pi_{i,j}b_i} \text{ or, in matrix notation, } \boldsymbol{\pi}' = [dg(\mathbf{b})\boldsymbol{\pi}]. \tag{10}$$

Of course $\boldsymbol{\pi}'$ coincides with $\boldsymbol{\pi}$ if attention is unbiased, b_i all equal.

An influence function describes how attention intensities are modified by a ranking. It is specified by a function B that assigns to each ranking $\mathbf{r}$ in Δ_N attention intensities $B(\mathbf{r})$. We will specify B later on, but this is unnecessary to describe the dynamics.

Let $\mathbf{r}^{(t)}$ be the ranking released at the beginning of period t. In period t, attention intensities are given by

$$\mathbf{b}^{(t)} = B(\mathbf{r}^{(t)}). \tag{11}$$

[10]This assumption is irrelevant if the method is intensity-invariant, as defined in Section 2.1.

These intensities result in new statements as described by (10):

$$\boldsymbol{\pi}^{(t)} = [dg(B(\mathbf{r}^{(t)}))\boldsymbol{\pi}]. \tag{12}$$

At the beginning of period $t+1$, the new ranking $\mathbf{r}^{(t+1)}$ based on matrix $\boldsymbol{\pi}^{(t)}$, $\mathbf{r}^{(t+1)} = F(\boldsymbol{\pi}^{(t)})$, is published. The ranking will determine the attention intensities of period $t+1$ hence the next matrix through (12) and so on. The joint process for the statement matrix and the ranking follows:

$$\boldsymbol{\pi}^{(t)} = [dg(B(\mathbf{r}^{(t)}))\boldsymbol{\pi}], \mathbf{r}^{(t+1)} = F(\boldsymbol{\pi}^{(t)}) \text{ for each } t. \tag{13}$$

Plugging the expression of $\boldsymbol{\pi}^{(t)}$ as a function of $\mathbf{r}^{(t)}$ yields the process followed by the sequence of rankings:

$$\mathbf{r}^{(t+1)} = F([dg(B(\mathbf{r}^{(t)}))\boldsymbol{\pi}]). \tag{14}$$

This section investigates the case where the influence function is linear, $B(x) = x$. Thus, spelling out the dynamics per item, the process followed by $\mathbf{r}^{(t)}$ writes

$$r_i^{(t+1)} = \sum_j \frac{\pi_{i,j} r_i^{(t)}}{\sum_{\ell \in N} \pi_{\ell,j} r_\ell^{(t)}} Q_j^F([dg(\mathbf{r}^{(t)})\boldsymbol{\pi})] \text{ for each } i \text{ in } N.$$

Given $\boldsymbol{\pi}$, the statement matrices along the process are all of the form $[dg(\mathbf{r})\boldsymbol{\pi}]$ for some positive $\mathbf{r}$. To simplify notation, denote $\boldsymbol{s}(\mathbf{r}) = (s_j(\mathbf{r}))$ the columns sums of $dg(\mathbf{r})\boldsymbol{\pi}$ and $\boldsymbol{q}(\mathbf{r}) = (q_j(\mathbf{r}))$ the weight vector associated to $[dg(\mathbf{r})\boldsymbol{\pi}]$:

$$s_j(\mathbf{r}) = \sum_i \pi_{i,j} r_i \text{ and } q_j(\mathbf{r}) = Q_j([dg(\mathbf{r})\boldsymbol{\pi}]). \tag{15}$$

The dynamics followed by $\mathbf{r}^{(t)}$ writes

$$r_i^{(t+1)} = \sum_j \frac{\pi_{i,j} r_i^{(t)}}{s_j(\mathbf{r}^{(t)})} q_j(\mathbf{r}^{(t)}) \text{ for each } i \text{ in } N. \tag{16}$$

Dividing (16) by $r_i^{(t)}$ gives an expression for the growth rate of i's score. The dynamics is trivial with identical preferences. The growth rate of the preferred item is strictly larger than for all other items, so its score converges to 1 and the others to zero. The interesting situation arises when preferences are diverse.

To analyze the dynamics and the possible convergence of the sequence $\mathbf{r}^{(t)}$ to a ranking $\mathbf{r}^*$, we need to consider the behavior of the method associated to statement matrices $[dg(\mathbf{r})\boldsymbol{\pi}]$ for $\mathbf{r}$ in a neighborhood of $\mathbf{r}^*$. Some continuity assumptions are needed. A minimal one is that the scores and the weights are continuous functions over the set of positive matrices. This is condition (a) below, which is satisfied by all methods introduced so far. An additional condition is needed because a limit point of the sequence may have null components (though each $\mathbf{r}^{(t)}$ is positive). For $\mathbf{r}^*$ with null components, the method F may not be defined on matrix $[dg(\mathbf{r}^*)\boldsymbol{\pi}]$ (because it has null rows). The additional continuity assumption (b) treats this case and bears on the behavior of the method close to $[dg(\mathbf{r}^*)\boldsymbol{\pi}]$ for any $\mathbf{r}^*$ in the simplex, possibly with null components.

Continuity assumption (C)

(a) F and Q^F are continuous functions over the set of positive matrices.

(b) $\boldsymbol{q}(\mathbf{r}) = Q^F([dg(\mathbf{r})\boldsymbol{\pi}])$ has a well defined limit when $\mathbf{r}$ tends to $\mathbf{r}^*$ where $\mathbf{r}^*$ is any vector in Δ_N.

Condition (a) does not deserve any comment.

Condition (b) bears on the set of nonnegative matrices of the form $[dg(\mathbf{r})\boldsymbol{\pi}]$, whose rows are either strictly positive (for $r_i > 0$) or null (for $r_i = 0$). The condition is of course satisfied by any method that is continuous over the whole set of all nonnegative matrices. The counting method is continuous over this larger set since experts' weights are constant, but very few methods are. For example, the invariant and the handicap-based methods are not[11] but they satisfy Condition (b) as shown in the appendix, Lemma 1.

From now on, we consider methods that satisfy the continuity assumption (C).

Rest points and their support The support of a ranking in the simplex is the set of indices with positive score. By continuity, any positive limit point $\mathbf{r}^*$ of the dynamics satisfies the fixed point condition

$$r_i^* = \sum_j \frac{\pi_{i,j} r_i^*}{s_j(\mathbf{r}^*)} q_j(\mathbf{r}^*) \text{ for each } i.$$

(For any point $\mathbf{r}^*$ in the simplex, each j's sum $s_j(\mathbf{r}^*) = \sum_\ell \pi_{\ell,j} r_\ell^*$ is strictly positive because of the positivity of $\boldsymbol{\pi}$) This equation is surely met for i with a null score since r_i^* can be factored out on the right hand side. So fixed points with some null components cannot be excluded. This reflects a self-enforcing mechanism: an item the score of which is null attracts no attention at all, hence is not assessed, which in turn justifies a null score. But such a mechanism may not be robust unless the point enjoys a minimum of stability. Recall that a fixed point $\mathbf{r}^*$ is locally asymptotically stable if the process converges to $\mathbf{r}^*$ for an open set of initial values for $\mathbf{r}^0$ around $\mathbf{r}^*$. Dividing (16) by $r_i^{(t)}$ gives an expression for the growth rate of i's score. If the sequence converges to a limit point

[11] Discontinuity arises when the matrix $\boldsymbol{\pi}$ is reducible and has several independent eigenvectors associated to the eigenvalue 1. Take for example the matrix $\boldsymbol{\pi} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. For the invariant method, any ranking is an eigenvector associated to the eigenvalue 1. The limit of the invariant ranking of $\boldsymbol{\pi} = \begin{pmatrix} 1-\epsilon & \epsilon^2 \\ \epsilon & 1-\epsilon^2 \end{pmatrix}$ when ϵ tends to zero is the transpose of $(0, 1)$; exchanging ϵ and ϵ^2 gives instead $(1, 0)$; The handicap-based method behaves similarly on this example, hence is not continuous as well.

with null i component, the limit growth rate must be less than 1. This gives the following necessary conditions for a point to be stable.

Definition 3 *Necessary conditions for* $\mathbf{r}^*$ *to be stable for the dynamics (16) are*

$$\sum_{j\in M} \frac{\pi_{i,j}}{s_j(\mathbf{r}^*)} q_j(\mathbf{r}^*) \leq 1 \text{ for each } i \text{ with an equality if } r_i^* > 0. \tag{17}$$

A point that satisfies (17) is called a **rest point***.*

It is useful to decompose the conditions (17), into two parts. First, a positive linear combination of the statements is equal to 1 on the support of the rest point and not greater than 1 outside it. Formally there is a non-negative m-vector $\boldsymbol{y}$ that satisfies the set of linear inequalities :

$$\sum_{j\in M} \pi_{i,j} y_j = 1 \text{ for each } i \text{ in } I \text{ and } \sum_{j\in M} \pi_{i,j} y_j \leq 1 \text{ for each } i \text{ not in } I. \tag{18}$$

Second, there is a ranking $\mathbf{r}^*$ with support I for which $(q_j/s_j)(\mathbf{r}^*) = y_j$ for each j.

A few properties on the support of a rest point are easily derived from the necessary conditions (18).

First, let us say that an item is 'dominated' by another item if it receives a strictly lower assessment from each expert than the other one. Quite naturally, a dominated item has a null score at a rest point. (The same result holds true if the item receives a strictly lower assessment than a convex combination of other items.)

Second, if the support is the whole set N, there is a positive linear combination of the statements under which all items' totals are equalized. To be possible, experts' preferences must be dispersed enough so as not to be in a clear way in favor of some items.

Third, when there are more items than experts, there are typically at most m items in the support of a rest point. To see this, observe that the m variables y_j must satisfy as many linear equations as the number of items in the support. If I has strictly more than m items, this is possible only in degenerate case since a small perturbation in the statements destroys the existence of the solution. Formally the non-degeneracy condition writes

(D1) Let I subset of N with cardinality strictly larger than m; then $\boldsymbol{\pi}_{I\times M}\boldsymbol{y} = \mathbb{1}_I$ has no solution.

To get an intuition of the dynamics, let us examine the impact on the scores of the modification of the statement matrix by $\mathbf{r}$. To simplify let us consider only two items. Let the statements for item 1 be increased relative to those of item 2. There is a direct effect, which tends to increase the score of item 1. The strength of this effect depends on the preferences, in particular on their diversity; the reason is that the proportions of the statements to item 1 are affected differently across experts, multiplied by $r_1/s_j(\mathbf{r})$ for expert j. An expert who likes item 1 more than another expert has a larger s_j thus its statement for item 1 is multiplied by less than for the other expert. This is more generally true with more than two items by considering the correlations of the experts' preferences with $\mathbf{r}$. As a result, the snowball mechanism due to the direct effect is mitigated by the diversity of preferences. This explains why several items may keep a positive score for the counting method. For other methods, there is also an indirect effect due to the variation in the weights with preferences. As seen in the next example, this indirect effect may be in either direction and reinforces or mitigates the direct effect, again depending on preferences.

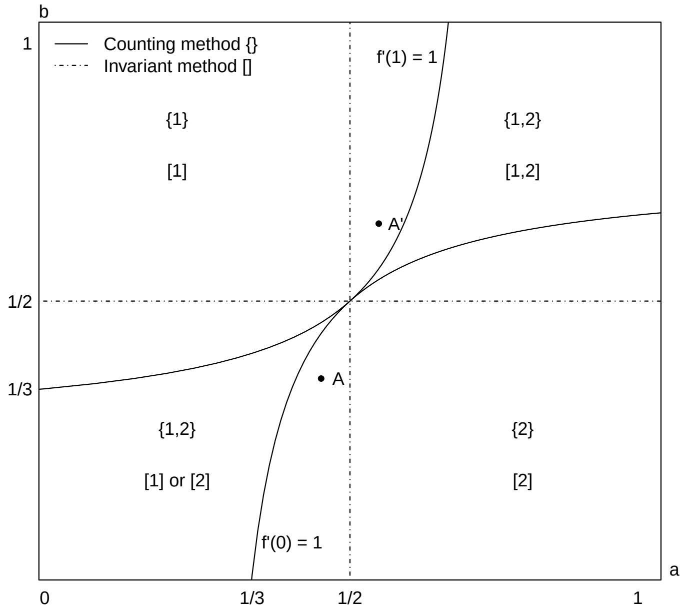

Figure 1: Supports depending on a, b for the invariant and counting methods

Example 1. Let us illustrate the dynamics when either the counting or the invariant method is used in the simple case with two items and two experts. For n and m equal to 2, a statement matrix writes

$$\boldsymbol{\pi} = \begin{pmatrix} 1-a & b \\ a & 1-b \end{pmatrix}.$$

The counting method assigns the ranking $\frac{1}{2}(1-a+b, 1+a-b)$ and the invariant method $(\frac{b}{a+b}, \frac{a}{a+b})$.

For the counting method, the process followed by r_1 is $r_1^{t+1} = f(r_1^{(t)})$, with

$$f(x) = \frac{1}{2}\left[\frac{(1-a)x}{(1-a)x + a(1-x)} + \frac{bx}{bx + (1-b)(1-x)}\right].$$

(the process for r_2 follows since $r_2 = 1 - r_1$). The sequence converges to a unique point whatever the initial value. It converges to 0 if $f'(0) = \frac{1}{2}(\frac{1-a}{a} + \frac{b}{1-b}) < 1$, to 1 if $f'(1) = \frac{1}{2}(\frac{a}{1-a} + \frac{1-b}{b}) < 1$, and to $r_1^* = \frac{1}{2} + \frac{b-a}{(1-2a)(1-2b)}$ otherwise. The regions are represented in Figure 1 where the numbers inside the {} represent the support of the ranking.

For the invariant method, the process converges but the limiting value may depend on the initial one. It is easier to work on the ratio $\rho = \frac{r_1}{r_2}$. The process followed by $\rho^{(t)}$ is

$$\rho^{(t+1)} = g(\rho^{(t)}) \text{ where } g(\rho) = \frac{b}{a}\rho\frac{(1-a)\rho + a}{b\rho + (1-b)}.$$

The fixed points are 0, or ∞, or $\rho^* = \frac{a}{b}\frac{1-2b}{1-2a}$ (if positive), which correspond respectively to the ranking $(0,1)$, or $(1,0)$, or to an interior ranking. See Figure 1 where the numbers inside the [] represent the support of a limit ranking. Note that, for a and b smaller than 1/2, that is, when each cites more itself than the other, the process converges to $(1,0)$ or to $(0,1)$ depending on the initial values; the interior ranking is not stable.

A comparison between the two processes shows that the limits, especially their supports, substantially differ. The interesting cases arise when a and b both on the same side relative to 1/2. Otherwise, one of the item 'dominates' the other one, which explains why only item 1 survives for $a < 1/2$ and $b > 1/2$, and only 2 survives for $a > 1/2$ and $b < 1/2$ whatever the method.

Let $a < 1/2$ and $b < 1/2$. Consider the invariant method. If the score of item 1 is high enough at some stage, this supports a high weight for 1 as an expert, which, in turn, justifies the high score for 1 as an item since it receives high evaluations from expert 1 (because $a < 1/2$). This feedback induces a snowball effect which is sufficiently important to eliminate the other item if it starts with a low enough score. The same argument holds for item 2 since $b < 1/2$. This explains why both $(1,0)$ and $(0,1)$ are limit points and the interior ranking is not stable. In the counting method, there is a snowball effect since an item may end up with a null score, as at A for example where item 1 is eliminated; however this effect is independent of the initial ranking: it must be that item 2 has a sufficient relative advantage with respect to 1, namely that a is sufficiently large with respect to b.

Let $a > 1/2$ and $b > 1/2$. For the counting method, the behavior is basically the same as in the previous case, up to a permutation. To see this, consider point A' with coordinates larger than 1/2 and the symmetric point A with respect to $(1/2, 1/2)$, which has both coordinates smaller than 1/2. The associated matrices $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$ are simply obtained by permuting their rows. The counting method performs the same permutation on the scores. For the invariant ranking the permutation is not at all neutral, from its very definition. Indeed, the behavior drastically differs with the previous case where a and b are both smaller than $1/2$. Since $a > 1/2$, expert 1 cites more item 2 than himself. Thus, a low score for item 1, which gives a low weight for expert 1, implies a low score for item 2. This explains why both items end up with a positive score.

The next section displays a family of methods under which there is a unique rest point whatever preferences $\boldsymbol{\pi}$; thus if the process converges, it converges to that point. In contrast, Section 3.3 defines peers' methods and shows that there are surely multiple rest points for some preferences.

3.2 Generalized handicap-based methods

This section introduces a family of methods derived from the handicap-based method by transforming its weights through some function as explained in section 2.2. The family of generalized handicap-based methods is obtained by using the functions $g(x) = x^\gamma$ for γ non-negative. From (9) the method[12] G associated to γ writes as

$$\text{for each } \boldsymbol{\pi} \text{ in } \mathcal{M},\ G_i(\boldsymbol{\pi}) = \sum_{j \in M} \pi_{i,j} \frac{q_j^\gamma}{\sum_{k \in M} q_k^\gamma} \text{ for each } i \in N \text{ where } \boldsymbol{q} = Q^H(\boldsymbol{\pi}). \tag{19}$$

The handicap-based method obtains for $\gamma = 1$ and the counting method for $\gamma = 0$. Increasing the value of γ increases the dispersion of the weights when they differ.

The next proposition uses two non-degeneracy conditions on the preferences matrix $\boldsymbol{\pi}$. The first one is **(D1)**, introduced in the previous section, which ensures that the support of a rest point has at most m items (which binds only if $n > m$). The second non-degeneracy condition bears on the statements of the experts on subsets with cardinality not greater than m.

(D2) Let I subset of N with cardinality equal to or less than m. The statements vectors restricted to I are linearly independent, or alternatively the matrix $\boldsymbol{\pi}_{I \times M}$ is of rank $|I|$.

Proposition 1 *Consider a generalized handicap-based method as defined by (19) with the parameter γ in $[0, 1[$ and $\boldsymbol{\pi}$ that satisfies the non-degenerate assumptions* **(D1)** *and* **(D2)***. Then the process (16) admits a unique rest point.*

The proposition applies, in particular, to the counting method ($\gamma = 0$). The handicap-based method, which obtains for γ equal to one, is not covered by Proposition 1. The case for γ larger than 1 is considered at the end of this section.

Proposition 1 is proved by considering the following function L :

$$\text{for } \gamma = 0 \quad L(\mathbf{r}) = \sum_j \ln(\sum_i \pi_{i,j} r_i) \tag{20}$$

$$\text{for } \gamma > 0 \quad L(\mathbf{r}) = \frac{1}{1-\gamma} \sum_j (Q_j^H(\boldsymbol{\pi}))^\gamma (\sum_i \pi_{i,j} r_i)^\gamma. \tag{21}$$

[12]The method can be extended to any positive matrix by applying the formula to the normalized matrix.

Specifically the dynamics is in fact a gradient method under constraints associated to L. As a result, a rest point maximizes L over $\mathbf{r}$ in Δ_N.

Proof of Proposition 1. Let G be the handicap-based method transformed by x^γ. I first show that the weights satisfy

$$q_j(\mathbf{r}) = \frac{a_j s_j^\gamma(\mathbf{r})}{\sum_{k\in M} a_k s_k^\gamma}(\mathbf{r}) \text{ for each } j \in M \tag{22}$$

for some constant values a_j. This will allow us to use technical lemmas proved in the appendix. Expression (22) states that the modification of the statement matrix by $\mathbf{r}$ affects the experts' weights as a function of the totals s_j, that is the correlations of experts' statements with the adjustment vector $\mathbf{r}$. As each row i is multiplied by r_i, the ratio of two experts' weights, say j and k, is proportional to s_j^γ/s_k^γ: the weight of the expert whose statement is more correlated with $\mathbf{r}$ increases more than the other. The larger γ, the stronger the effect.

By definition, the weights Q^G are proportional to $(Q^H)^\gamma$ on normalized matrices where Q^H are the weights for the handicap-based method; in particular $Q^G([dg(\mathbf{r})\boldsymbol{\pi}]) \propto Q^H([dg(\mathbf{r})\boldsymbol{\pi}])^\gamma$. From (8) the weight vector $Q^H([dg(\mathbf{r})\boldsymbol{\pi}])$ is proportional to $(Q_j^H(\boldsymbol{\pi})s_j)$. Hence we have

$$Q^G([dg(\mathbf{r})\boldsymbol{\pi}]) \propto (Q_j^H(\boldsymbol{\pi})^\gamma s_j^\gamma). \tag{23}$$

This proves (22). Thus the assumption of Lemma 2 holds for functions ψ_j given by

$$\psi_j(s_j) = Q_j^H(\boldsymbol{\pi})^{1-\gamma} s_j^{\gamma-1}. \tag{24}$$

Let function L be defined by $L(\mathbf{r}) = \sum_j \Psi_j(\sum_i \pi_{i,j} r_i)$ where Ψ_j is a primitive of ψ_j. L coincides with the expressions (20). Thanks to Lemma 2, the rest points are the points that satisfy the first order conditions of the program $(\mathcal{P})$ of maximization of L over $\mathbf{r}$ in Δ_N.

Furthermore, for γ less than 1, the functions ψ_j are decreasing. The uniqueness of a rest point follows by application of Lemma 3 under the non-degeneracy assumptions. ∎

The case of γ larger than 1. I start with the handicap-based method, which obtains for γ equal to one (the function L used in the proof is linear so the uniqueness of a rest point is not guaranteed). Multiple rest points are possible but in degenerate situations. A direct proof shows the convergence towards rankings with supports on the items whose handicap-based score is maximal. Since, typically, there is only one item with maximal score, global convergence is guaranteed. The snowball effect is quite strong since all items but one end up with a null score. As we will see, as γ is increased, the snowball effect, which is increased, is large enough to justify several limit points, each one depending on the initial value.

Let $\mathbf{r}^h = H(\boldsymbol{\pi})$ and $\boldsymbol{q}^h = Q^H(\boldsymbol{\pi})$, the values assigned by the handicap-based method to $\boldsymbol{\pi}$. Recall that $\mathbf{r}^{(t+1)}$ and $\boldsymbol{q}^{(t+1)}$ are the rankings and the weights assigned by the handicap-based method to

$[dg(\mathbf{r}^{(t)})\boldsymbol{\pi}]$. Using (8), we have for each $t \geq 0$

$$r_i^{(t+1)} = \lambda^{(t+1)} r_i^{(t)} r_i^h \text{ and } q_j^{(t+1)} = \lambda^{(t+1)} q_j^h (\sum_\ell \pi_{\ell,j} r_\ell^{(t)}) \tag{25}$$

for some $\lambda^{(t+1)}$. Consider the growth rate of score i, $g_i^{(t+1)} = r_i^{(t+1)}/r_i^{(t)}$. (25) implies

$$\frac{g_i^{(t+1)}}{g_\ell^{(t+1)}} = \frac{r_i^h}{r_\ell^h},$$

that is, the ratio of the growth rate of i over that of ℓ stays constant, equal to the ratio of their handicap-based scores. Denote by I the set of items for which the handicap-based score is maximal. In general I is a singleton, say $\{1\}$. In that case, 1's growth rate is strictly larger than that of each other item. Since the score of 1 is bounded, the scores of each other item must converge to 0. The argument extends to the case where I is not a singleton: any ranking that has I as a support is a limit point.[13] ∎

When γ is larger than 1, several rest points are possible. (Part of the proof of Proposition 1 is still valid. In particular the rest points still satisfy the first order conditions of the maximization of L over the set of rankings. But now L is convex.) As proved in the appendix, the process always converges and the limit ranking has all its components but one equal to zero. The convergence is local: the limit point may depend on the initial ranking. The convergence and the multiplicity are due to the high sensitivity of the weights to differences in preferences; once an item has a high enough score, it induces the following period a large weight for the experts who like that item, and then the snowball effect makes its score converge to 1.

Example of multiplicity Consider a generalized handicap-based method G with $g(x) = x^2$ (which corresponds to $\gamma = 2$). This example shows that there are multiple rest points for a whole open set of parameters a and b: Multiplicity is a robust phenomena. Let us consider the 2×2 matrix $\boldsymbol{\pi} = \begin{pmatrix} 1-a & b \\ a & 1-b \end{pmatrix}$. Easy computation gives that the handicap-based ranking and the supporting weights are

$$\mathbf{r}^h \propto \left(\sqrt{(1-a)b}, \sqrt{a(1-b)}\right), \boldsymbol{q}^h \propto \left(\sqrt{(1-b)b}, \sqrt{a(1-a)}\right). \tag{26}$$

The computation performed in the proof of Proposition 1 is still valid so that the weights $\boldsymbol{w} = Q^G([dg(\mathbf{r})\boldsymbol{\pi}])$ are proportional to $((q_1^h s_1)^2, (q_2^h s_2)^2)$. We look for values of a and b under which the

[13] By the same argument, the scores of the items outside I must converge to zero. To show that conversely any ranking with null components on $N - I$ is a limit point, observe that the growth rates of the scores of elements in I are equal : for i and ℓ both in I, the ratio r_i^h/r_ℓ^h is equal to one, thus $g_i^{(t+1)} = g_\ell^{(t+1)}$. This implies that the ratios $r_i^{(t)}/r_\ell^{(t)}$ stay constant equal to their initial values $r_i^{(0)}/r_\ell^{(0)}$. Thus the sequence $\mathbf{r}^{(t)}$ converges towards the ranking whose components are proportional on I to their initial values and are null outside.

ranking (1,0) is a rest point. The conditions (17) write

$$(1-a)\frac{w_1}{s_1} + b\frac{w_2}{s_2} = 1 \text{ and } a\frac{w_1}{s_1} + (1-b)\frac{w_2}{s_2} \leq 1$$

where $s_1 = 1-a$ and $s_2 = b$. The first condition writes $w_1 + w_2 = 1$, so it is surely satisfied. As for the second, it can be rewritten as $1 \leq \frac{1}{2}(\frac{1-a}{a} + \frac{b}{1-b})$ by easy computation.[14] Similarly, (0,1) is a rest point if $1 \leq \frac{1}{2}(\frac{a}{1-a} + \frac{1-b}{b})$. The set of values for which both rankings (1,0) and (0,1) are rest points is an open set. These values are exactly the one for which the counting method converges to a ranking with full support, represented by the lens in Figure 1 in section 3.1, but there is no general reason for why it is true.

3.3 Peers' methods

This section gives a definition to a peers' method and analyzes the associated dynamics. A peers' method requires the experts to coincide with the items but not only. The rationale underlying a peers' method is that the ability to provide correct expertise is positively related with the performance as an item. This makes sense in a setting in which items are ordered by a single 'ability' parameter that drives their capacity both to perform and to judge others. For a method supported by weights, a minimal requirement to be qualified as a peers' method is that an item that is assigned a small score is also assigned a small expert's weight and vice-versa. The definition follows. A peers' method is one for which the weight as an expert is bounded relative to the score as an item.

Definition 4 *Let $N = M$. A method F supported by Q^F is a peers' method if the ratio Q_i^F/F_i is bounded on $\mathcal{M}$: there are positive k and k' such that*

$$k' \leq \frac{Q_i^F}{F_i}(\boldsymbol{\pi}) \leq k \text{ for each positive } \boldsymbol{\pi} \text{ in } \mathcal{M}. \tag{27}$$

The lower bound condition is automatically satisfied if the weights assume a minimum positive value, as is the case for the counting method: since the scores are bounded above by 1, the ratio Q_i^F/F_i is larger than the minimum of the weights. Thus the lower bound is relevant only for methods with arbitrarily small experts weights. The upper bound condition on Q_i^F/F_i is always relevant because scores can be arbitrarily small: for a method supported by weights (or for any reasonable method), an item's score is not greater than the maximum value of its assessments.

The counting method (when applied to $N = M$) is not a peers' method since an item's score can be arbitrarily small while its expert's weight is constant. The invariant method is a peers'

[14] We have $w_1 = \frac{{q_1^h}^2 s_1^2}{{q_1^h}^2 s_1^2 + {q_2^h}^2 s_2^2}$ so plugging the expression (26) of $\boldsymbol{q}^h$ yields $w_1 = \frac{b(1-b)s_1^2}{b(1-b)s_1^2 + a(1-a)s_2^2}$ and, replacing the value of s_1 and s_2, $\frac{w_1}{s_1} = \frac{(1-b)}{(1-a)(1-b)+ab}$; similarly $\frac{w_2}{s_2} = \frac{a}{(1-a)(1-b)+ab}$; the condition thus writes $2a(1-b) \leq (1-a)(1-b) + ab$ or equivalently $1 \leq \frac{1}{2}(\frac{1-a}{a} + \frac{b}{1-b})$.

method since F and Q^F coincide. The Hits method is not a peers' method, as the following example illustrates. Let the matrix

$$\boldsymbol{\pi}(\epsilon) = \begin{pmatrix} 2\epsilon & \epsilon & \epsilon \\ 1/2-\epsilon & \epsilon & 1-2\epsilon \\ 1/2-\epsilon & 1-2\epsilon & \epsilon \end{pmatrix}.$$

As ϵ tends to 0, the ranking assigned by the Hits method converges to $(0, 1/2, 1/2)$ and the experts' weights converge to $(1/3, 1/3, 1/3)$: while the score of item 1 vanishes, the weight does not.[15] For sake of comparison, the same limit ranking $(0, 1/2, 1/2)$ is obtained for the invariant method, which is also, by definition, the limit of the experts' weights.

Quite strong results are obtained using the following property of a peer's method. the weight $q_i(\mathbf{r})$ of expert i associated to matrix $[dg(\mathbf{r})\boldsymbol{\pi}]$ is null if and only if r_i is null. To show this, observe that for a peers' method the weight of an expert is null if only if its score is. It suffices then to observe that $F_i([dg(\mathbf{r})\boldsymbol{\pi}])$ writes as $\sum_j \frac{\pi_{i,j} r_i}{s_j(\mathbf{r})} q_j(\mathbf{r})$, a convex combination of the values $\frac{\pi_{i,j} r_i}{s_j(\mathbf{r})}$. Hence the score of i for $[dg(\mathbf{r})\boldsymbol{\pi}]$ is null if and only if r_i is null.

The two following propositions extend the results obtained for the invariant method according to which the dynamics may have multiple rest points, or even multiple locally stable rest points (Demange (2012-a)). It turns out that this multiplicity is bound to occur with peers' methods.

The first proposition characterizes the supports of the rest points. Notice that the characterization is independent of the peers' method under consideration.

Proposition 2 *Consider a peers' method. Given $\boldsymbol{\pi}$, a subset I of N is the support of a rest point if and only if there is a positive vector indexed by I, $\boldsymbol{x} = (x_i)_{i \in I}, x_i > 0$ for each i, such that*

$$\sum_{j \in I} \pi_{i,j} x_j = 1 \text{ for each } i \text{ in } I \text{ and } \sum_{j \in I} \pi_{i,j} x_j \leq 1 \text{ for each } i \text{ not in } I. \tag{28}$$

or, in matrix form $\boldsymbol{\pi}_{I \times I} \boldsymbol{x} = \mathbb{1}_I$, $\boldsymbol{\pi}_{N-I \times I} \boldsymbol{x} \leq \mathbb{1}_{N-I}$.

It is easy to understand why (28) is necessary for I to be the support of a rest point. For a peers' method, $q_j(\mathbf{r}^*)$ is null whenever r_i^* is null. Hence it suffices to define $\boldsymbol{x}$ by $x_i = (q_i/s_i)(\mathbf{r}^*)$ on I and to use the conditions (17) on a rest point. For a general method, we used a similar argument to prove the linear inequalities (18); the difference here is that the weights are null outside I. As a

[15]

$$\widetilde{\boldsymbol{\pi}}(\epsilon)\boldsymbol{\pi}(\epsilon) = \begin{pmatrix} 0 & 1/2 & 1/2 \\ \epsilon & \epsilon & 1-2\epsilon \\ \epsilon & 1-2\epsilon & \epsilon \end{pmatrix} \begin{pmatrix} 0 & \epsilon & \epsilon \\ 1/2 & \epsilon & 1-2\epsilon \\ 1/2 & 1-2\epsilon & \epsilon \end{pmatrix} \text{ converges to } \begin{pmatrix} 1/2 & 1/2 & 1/2 \\ 1/2 & 1 & 0 \\ 1/2 & 0 & 1 \end{pmatrix}.$$

The limit matrix has three independent positive eigenvectors: $(1/3, 1/3, 1/3)$, $(0, 1, 0)$, and $(0, 0, 1)$. Since, by symmetry, the principal eigenvector of $\widetilde{\boldsymbol{\pi}}(\epsilon)\boldsymbol{\pi}(\epsilon)$ puts equal weights on 2 and 3, it converges to $(1/3, 1/3, 1/3)$, the eigenvector of the limit matrix that has this property.

result, the obtained conditions, in particular the linear equation $\boldsymbol{\pi}_{I\times I}\boldsymbol{x} = \mathbb{1}_I$, are more restrictive hence gives more information on the possible candidates for a support.

To show that conversely (28) guarantees the existence of a rest point with support in I, we need to find $\mathbf{r}^*$ positive on I and null outside I that satisfies (17). This is proved by building a correspondence on the set of rankings with support I whose fixed points satisfy $x_i = (q_i/s_j)(\mathbf{r}^*)$ (see the details in the section proof).

Proposition 2 can be interpreted as follows. Consider first the whole set N. The conditions (28) state the existence of a positive n-vector $\boldsymbol{x}$ for which $\boldsymbol{\pi x} = \mathbb{1}_N$, as is also the case for any method, peers or not. This is equivalent to the non-negativity of the vector $\boldsymbol{\pi}^{-1}\mathbb{1}_N$ when the matrix $\boldsymbol{\pi}$ is invertible. As already seen, this is possible if experts' preferences are dispersed enough.

Consider now a subset I of N. Now there is a positive linear combination of the statements of the experts in I that is equal to 1 across the items of I and smaller outside. Within I, the experts should not be too much in favor of a restricted subset of them. Furthermore they should give low enough assessments on the outsiders.

Let us relate the proposition with the direct analysis in the 2×2 case for the invariant method. There is a rest point with support $N = \{1, 2\}$ if $\boldsymbol{\pi x} = \mathbb{1}_N$ has a positive solution. Assuming $\boldsymbol{\pi}$ invertible, $1 \neq a + b$, the solution is $\boldsymbol{x}^* = (\frac{1-2b}{1-a-b}, \frac{1-2a}{1-a-b})$. To be positive a and b must be both on the same side relative to 1/2, which is the condition we found where an interior ranking is a fixed point of the function g describing the dynamics. As we saw, when both a and b are smaller than 1/2 there are multiple rest points, and only the singletons are stable.

From Proposition 2, the supports of the rest points are independent of the peers' method. The rest points, that is the precise values assumed by the scores on such a support, are not.

The next proposition bears on convergence to zero of some scores.

Proposition 3 *Consider a peers' method. Let I be a subset of N and a matrix $\boldsymbol{\pi}$ such that for some $k < 1$,*

$$\frac{\beta}{\alpha} \leq k \text{ where } \beta = \max_{(i,j)\in N-I\times I} \pi_{i,j} \text{ and } \alpha = \min_{(i,j)\in I\times I} \pi_{i,j}. \tag{29}$$

If the scores on $N - I$ are small enough at some stage, their scores converge to zero.

Observe that the conditions (29) bear on the columns of $\boldsymbol{\pi}$ indexed by I only, namely on the statements of experts I. They requires experts I to state small enough values on the items in $N - I$ relative to the items in I. The intuition is clear: for a peer's method, if the score on $N - I$ are low enough, then their weights are also low and the statements of experts mostly matter. Since these experts do not cite much items in $N - I$, the effect is amplified. Formally, the condition ensures that the growth rate of each score of items in $N - I$ is bounded by a value smaller than 1 if the ranking

$\mathbf{r}$ has low enough components on $N - I$. What is meant by 'low' enough depends on the method and the values of the statements for the experts on $N - I$.

Many matrices satisfy (29). The simplest example is

$$\begin{pmatrix} \alpha & & \alpha & \times,\times \\ .. & .. & .. & \times,\times \\ \alpha & .. & \alpha & \times,\times \\ \beta & .. & \beta & \times,\times \\ \beta & .. & \beta & \times,\times \end{pmatrix}$$

where $\alpha|I| + \beta|N - I| = 1$, and β small enough so that $\beta/\alpha < 1$.

Corollary 1 *Consider a peers' method. There are matrices* $\boldsymbol{\pi}$ *for which the dynamics (16) admit several locally stable points.*

It is easy to understand why a problem may admit multiple stable rest points. Let preferences be sufficiently antagonistic in the following way. Take I a subset and choose preferences from I to $N-I$ small enough and the same for $N - I$ to I, that is the $\pi_{i,j}$ small enough for i in I and j not in I or the reverse. First I or a subset of I is the support of a rest point, and similarly for $N - I$. The key point is that the stability of a rest point null on $N - I$ is independent of the values of $\pi_{i,j}$ for j not in I, namely the preferences of $N - I$.

4 Extensions and concluding remarks

We first analyze the impact of non linear influence functions.

4.1 Non linear influence functions

So far we have assumed a specific form, linear, for the influence of rankings. This section considers more general influence functions B, $B(r) = r^\alpha$, for a positive parameter α: j's statements are proportional to $\pi_{i,j} r_i^\alpha$ given ranking $\mathbf{r}$. The dynamics follow, as described in Section 3.1.

With a slight abuse of notation let $\mathbf{r}^\alpha$ denote the vector (r_i^α). Thus the process (13) followed by the rankings writes

$$r_i^{(t+1)} = \sum_j \frac{\pi_{i,j} r_i^{(t)\alpha}}{s_j(\mathbf{r}^{(t)\alpha})} q_j(\mathbf{r}^{(t)\alpha}) \text{ each } i. \tag{30}$$

in which the function $\boldsymbol{q}$ and $\boldsymbol{s}$ are unchanged given by (15). A rest point $\mathbf{r}^*$ satisfies

$$\sum_j \frac{\pi_{i,j} r_i^{*\alpha-1}}{s_j(\mathbf{r}^{*\alpha})} q_j(\mathbf{r}^{*\alpha}) \leq 1 \text{ for each } i \text{ with an equality if } r_i^* > 0. \tag{31}$$

As α increases, the impact of a difference in the scores of two items induces larger differences in attention. Furthermore, for α larger than 1, the marginal gain in attention is increased with the score, whereas it is the opposite for α smaller than 1. This explains why the analysis differs sensibly depending on the value of α with respect to 1, the boundary case $\alpha = 1$ being the case studied in the previous sections.

Diminishing marginal impact ($\alpha < 1$) For $\alpha < 1$, (30) implies that the growth rate of i's score is strictly larger than 1 for r_i small enough. As a result, no item's score converges to zero. Furthermore the possible rest points are necessarily strictly positive and satisfy

$$\sum_j \frac{\pi_{i,j} {r_i^*}^{\alpha-1}}{s_j(\mathbf{r}^{*\alpha})} q_j(\mathbf{r}^{*\alpha}) = 1 \text{ for each } i. \tag{32}$$

We can say more for generalized handicap-based methods. The proof of Proposition 1 extends by considering the function $L(\mathbf{r}^\alpha)$ (see the details at the end of the proof of Lemma 2). It follows that the process (30) admits a unique rest point for any generalized handicap-based method adjusted by $g(x) = x^\gamma$ and influence function $B(r) = r^\alpha$ for which $\gamma \geq 0$ and $\alpha \leq 1$ with at least one strict inequality.

Increasing marginal impact ($\alpha > 1$) For $\alpha > 1$, the marginal gain in attention is increasing with the score so that one may expect a limit ranking to be concentrated on few items. We show here that this is indeed true for the invariant method: any ranking concentrated on a single point is a stable point.

The proof is as follows. Starting with a score for item i low enough, we show that i's score will decrease exponentially to 0. Thus the sequence $\mathbf{r}^{(t)}$ converges to $\mathbb{1}_{\{\ell\}}$ for initial rankings in a neighborhood of $\mathbb{1}_{\{\ell\}}$, that is rankings for which all scores except that of ℓ are small enough. The proof relies on the following inequality : For some C

$$\forall t > 0, r_i^{(t+1)} \leq C {r_i^{(t)}}^\alpha. \tag{33}$$

Assuming (33), iteration from 0 up to t implies

$$r_i^{(t+1)} \leq C^{1+\alpha+\cdots+\alpha^t} {r_i^{(0)}}^{\alpha^{t+1}}, \text{ or } r_i^{(t+1)} \leq C^{\frac{1}{1-\alpha}} \left[C^{\frac{1}{\alpha-1}} r_i^{(0)}\right]^{\alpha^{t+1}}.$$

Hence if $C^{\frac{1}{\alpha-1}} r_i^{(0)} < 1$ then the sequence $r_i^{(t)}$ converges to 0 as t tends to ∞. As a consequence, if all $r_i^{(0)}$ for $i \neq \ell$ are small enough, then the rankings converge to $\mathbb{1}_{\{\ell\}}$.

It remains to show (33). For the invariant method, the dynamics is given by

$$\frac{r_i^{(t+1)}}{{r_i^{(t)}}^\alpha} = \sum_j \pi_{i,j} \frac{r_j^{(t+1)}}{\sum_{\ell \in N} \pi_{\ell,j} {r_\ell^{(t)}}^\alpha}.$$

We have to show that the right hand side is bounded above. Let us denote by π_{max} and π_{min} respectively the maximum and minimum of the elements in matrix $\boldsymbol{\pi}$. We first provide a lower bound to $\sum_{\ell \in N} \pi_{\ell,j} {r_\ell^{(t)}}^\alpha$. Observe that

$$\pi_{min} \sum_{\ell \in N} {r_\ell^{(t)}}^\alpha \leq \sum_{\ell \in N} \pi_{\ell,j} {r_\ell^{(t)}}^\alpha. \tag{34}$$

To bound $\sum_{\ell \in N} {r_\ell^{(t)}}^\alpha$, we apply Holder inequality[16] to the vectors $\mathbf{r}^{(t)}$ and $\mathbb{1}_N$ with the parameters $p = \alpha$ and $q = \alpha/(\alpha - 1)$ (q is positive since $\alpha > 1$). This yields

$$\sum_{\ell \in N} r_\ell^{(t)} \leq (\sum_{\ell \in N} {r_\ell^{(t)}}^\alpha)^{\frac{1}{\alpha}} n^{\frac{(\alpha-1)}{\alpha}}.$$

Since $\sum_{\ell \in N} r_\ell^{(t)} = 1$, this writes $1 \leq (\sum_{\ell \in N} {r_\ell^{(t)}}^\alpha)^{\frac{1}{\alpha}} n^{\frac{(\alpha-1)}{\alpha}}$ or $n^{(1-\alpha)} \leq \sum_{\ell \in N} {r_\ell^{(t)}}^\alpha$. Using inequality (34) gives

$$\pi_{min} n^{(1-\alpha)} \leq \pi_{min} \sum_{\ell \in N} {r_\ell^{(t)}}^\alpha \leq \sum_{\ell \in N} \pi_{\ell,j} {r_\ell^{(t)}}^\alpha.$$

So together with $\pi_{i,j} \leq \pi_{max}$ we obtain

$$\frac{r_i^{(t+1)}}{{r_i^{(t)}}^\alpha} \leq \sum_j \pi_{max} \frac{r_j^{(t+1)}}{\pi_{min} n^{(1-\alpha)}} = \frac{\pi_{max}}{\pi_{min} n^{(1-\alpha)}}.$$

Inequality (33) holds for C the value on the right hand side. ∎

4.2 Concluding remarks

This paper analyzes the influence of rankings based on the premise that rankings induce a coordination on attention. It shows that, for an identical influence mechanism, the interplay between preferences and the ranking method results in a variety of different outcomes. While a class of methods (the class based on handicaps) ensures the uniqueness of a rest point, self-enforcing mechanisms at play in peers' methods are strong enough to sustain multiple equilibria.

Several developments could be contemplated. It would be interesting to analyze more closely the support of the rest points and to understand better their links with the method at hand. Also, the paper concentrates on the influence of a single ranking. Although assuming a single ranking is appropriate for modeling the Web environment given the dominance of the use of Google Search, the assumption should be relaxed. A natural issue is to relate the number of rankings with the preferences, thereby making rankings endogenous.

[16] Holder inequality is $\sum_\ell x_\ell y_\ell \leq (\sum_\ell x_\ell^p)^{1/p} (\sum_\ell y_\ell^q)^{1/q}$, for p and q positive related by $1/p + 1/q = 1$.

5 Proofs

Lemma 1 *The invariant and the handicap-based methods satisfy the continuity assumption (C) (b).*

Proof of Lemma 1 We need to show that $Q^F([dg(\mathbf{r})\boldsymbol{\pi}])$ has a well defined limit when $\mathbf{r}$ tends to a ranking $\mathbf{r}^*$ that has null components. Let I be the support of $\mathbf{r}^*$. The matrix $dg(\mathbf{r}^*)\boldsymbol{\pi}$ has its rows indexed by $N-I$ null.

Consider the invariant method. Let a sequence of positive $(\mathbf{r}^k)$ that tend to $\mathbf{r}^*$ as k increases to infinity and $(\boldsymbol{\xi}^k)$ be the corresponding sequence of weights, $\boldsymbol{\xi}^k = Q([dg(\mathbf{r}^k)\boldsymbol{\pi}])$. We show that the sequence $(\boldsymbol{\xi}^k)$ converges to the vector $(\boldsymbol{q}_I^*, \mathbf{0}_{N-I})$ where $\boldsymbol{q}_I^*$ is the normalized principal eigenvector of $[dg(\mathbf{r}^*)\boldsymbol{\pi}_{I\times I}]$. Since each $\boldsymbol{\xi}^k$ belongs to the simplex, it suffices to show that any limit point $\overline{\boldsymbol{\xi}}$ of a subsequence is the vector $(\boldsymbol{q}_I^*, \mathbf{0}_{N-I})$. The experts' weights are equal to the ranking for a normalized matrix, hence $\boldsymbol{\xi}^k$ is the vector in Δ_N that satisfies

$$\xi_i^k = \sum_j \frac{\pi_{i,j} r_i^k}{\sum_{\ell\in N} \pi_{\ell,j} r_\ell^k} \xi_j^k \text{ for each } i \text{ in } N.$$

Taking the limit on a converging subsequence, $\overline{\boldsymbol{\xi}}$ satisfies (since $\sum_{\ell\in I} \pi_{\ell,j} r_\ell^*$ is positive)

$$\overline{\xi}_i = \sum_{j\in I} \frac{\pi_{i,j} r_i^*}{\sum_{\ell\in I} \pi_{\ell,j} r_\ell^*} \overline{\xi}_j \text{ for each } i \text{ in } I \text{ and } \overline{\xi}_i = 0 \text{ for each } i \text{ not in } I.$$

Since $\overline{\boldsymbol{\xi}}$ is in the simplex, $\overline{\boldsymbol{\xi}}_I$ is the unique normalized eigenvector of the matrix $[dg(\mathbf{r}^*)\boldsymbol{\pi}_{I\times I}]$ (which coincides with $[dg(\mathbf{r}^*)\boldsymbol{\pi})]_{I\times I}$ the submatrix of $[dg(\mathbf{r}^*)\boldsymbol{\pi}]$ formed by deleting the rows and the columns not indexed by I). This proves the result.

For the handicap-based method, we show the expression 8 for each $\mathbf{r} \gg \mathbf{0}$

$$\begin{aligned} H_i([dg(\mathbf{r})\boldsymbol{\pi}]) &= \frac{r_i r_i^h}{\sum_\ell r_\ell r_\ell^h} \text{ for each } i,\ Q_j^H([dg(\mathbf{r})\boldsymbol{\pi}]) = \frac{q_j^h s_j}{\sum_\ell r_\ell r_\ell^h} \text{ for each } j \\ \text{where } \mathbf{r}^h &= H(\boldsymbol{\pi}),\ \boldsymbol{q}^h = Q^H(\boldsymbol{\pi}), s_j = \sum_i \pi_{i,j} r_i. \end{aligned}$$

Denote $\pi'_{i,j} = \frac{\pi_{i,j} r_i}{s_j} = [dg(\mathbf{r})\boldsymbol{\pi}]_{i,j}$ where $s_j = \sum_i \pi_{i,j} r_i$. From (7) we have

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

Plugging in the value of $\pi_{i,j}$ as a function of $\pi'_{i,j}$ we obtain

$$\sum_j (\pi'_{i,j} q_j^h s_j) \frac{1}{r_i r_i^h} = 1 \text{ for each } i \text{ where } \sum_i (\frac{\pi'_{i,j}}{r_i r_i^h}) q_j^h s_j = \frac{n}{m} \text{ for each } j. \tag{35}$$

Thus the handicap-based ranking of $\boldsymbol{\pi}' = [dg(\mathbf{r})\boldsymbol{\pi}]$ is the ranking proportional to the vector $(r_i r_i^h)$ with weights proportional to $(q_j^h s_j)$. This yields (8). When $\mathbf{r}$ tends to a vector $\mathbf{r}^*$ with possibly null components, the limits are well defined (since $\sum_\ell r_\ell^* r_\ell^h$ is positive). ■

Lemma 2, 3 and 4.

We state and prove here three lemmas used in Section 3. They all assume the following assumption on the weights Q. Given $\boldsymbol{\pi}$ in $\mathcal{M}$, there are some functions ψ_k, k in M, defined from the set of positive scalar numbers, $\Re_+$, to itself, continuous such that

$$q_j(\mathbf{r}) = Q_j([dg(\mathbf{r})\boldsymbol{\pi}]) = \frac{s_j\psi_j(s_j)}{\sum_{k\in M} s_k\psi_k(s_k)} \text{ for each } j \in M \text{ where } s_j = \sum_{i\in N} \pi_{i,j} r_i. \tag{22}$$

Recall that $\boldsymbol{s} = (s_j)$ assigns to $\mathbf{r}$ the columns' sums $s_j(\mathbf{r}) = \sum_i \pi_{i,j} r_i$. Let Ψ_j denote a primitive of ψ_j and L the function

$$L(\mathbf{r}) = \sum_j \Psi_j(\sum_i \pi_{i,j} r_i) \text{ or } L(\mathbf{r}) = \sum_j \Psi_j(s_j(\mathbf{r})).$$

When there is no confusion, we simply write s_j for $s_j(\mathbf{r})$. Consider the program $(\mathcal{P})$

$$(\mathcal{P}) : \max_{\mathbf{r}} L(\mathbf{r}) \text{ over } \mathbf{r} \geq 0, \sum_i r_i \leq 1.$$

Lemma 2 *The points that satisfy the first order conditions of program* $(\mathcal{P})$ *coincide with the rest points.*

Proof of Lemma 2. Let μ denote the multiplier associated to the constraint $\sum_i r_i \leq 1$. The first order conditions satisfied by a solution $\mathbf{r}$ of $(\mathcal{P})$ are

$$\frac{\partial L}{\partial r_i}(\mathbf{r}) \leq \mu \text{ for each } i \text{ with } = \text{ for } r_i > 0. \tag{36}$$

Multiplying by r_i, using $\sum_i r_i = 1$, we obtain the value of μ: $\mu = \sum_{\ell\in N} r_\ell \frac{\partial L}{\partial r_\ell}(\mathbf{r})$. Now

$$\frac{\partial L}{\partial r_i}(\mathbf{r}) = \sum_j \pi_{i,j}\psi_j(s_j). \tag{37}$$

Exchanging the sums yields

$$\sum_{\ell\in N} r_\ell \frac{\partial L}{\partial r_\ell}(\mathbf{r}) = \sum_{j\in M} s_j(\mathbf{r})\psi_j(s_j). \tag{38}$$

Thus the first order conditions (36) write

$$\sum_j \pi_{i,j}\psi_j(s_j) \leq \sum_k s_k\psi_k(s_k) \text{ for each } i \text{ with} = \text{for } r_i > 0. \tag{39}$$

Using the expression (22) for $\boldsymbol{q}(\mathbf{r})$, i.e. $q_j(\mathbf{r}) = \frac{s_j\psi_j(s_j)}{\sum_k s_k\psi_k(s_k)}$, the first order conditions (39) write

$$\sum_j \frac{\pi_{i,j}}{\sum_{\ell\in N} \pi_{\ell,j} r^*_\ell} q_j(\mathbf{r}^*) \leq 1 \text{ for each } i \text{ with an equality if } r^*_i > 0$$

which are the conditions (17) on a rest point. ■

The intuition for the result can be derived by observing that the dynamics is in fact a gradient method under constraints associated to L. To see this using both (37) and (38), the dynamics (16) followed by $\mathbf{r}^{(t+1)}$, $r_i^{(t+1)} = \sum_j \frac{\pi_{i,j} r_i^{(t)}}{\sum_{\ell \in N} \pi_{\ell,j} r_\ell^{(t)}} q_j(\mathbf{r}^{(t)})$ each i, can be written as

$$r_i^{(t)} = \frac{\frac{\partial L}{\partial r_i}(\mathbf{r}^{(t-1)})}{V^{(t-1)}} r_i^{(t-1)} \text{ each } i \text{ where } V^{(t-1)} = \sum_{\ell \in N} r_\ell^{(t-1)} \frac{\partial L}{\partial r_\ell}(\mathbf{r}^{(t-1)}). \tag{40}$$

According to these equations, the growth rates of the components of $\mathbf{r}$ are proportional to the gradient of L.

Lemma 3 *Let* $\boldsymbol{\pi}$ *satisfy* **(D1)** *and* **(D2)** *and* Q *satisfy (22) with* ψ_k *strictly decreasing. Then there is a unique rest point.*

Proof of Lemma 3. If the ψ_j are decreasing, then the functions Ψ_j are strictly concave. Thus L is concave in $\mathbf{r}$. If L is strictly concave, the result is obvious. However L is not strictly concave if the function $\boldsymbol{s}$ is not one-to-one, which is surely true if $n > m$. So we work on the solutions to $(\mathcal{P})$.

Let $\mathbf{r}$ and $\mathbf{r}'$ be two solutions. Observe first that the values of the columns' sums are identical across solutions: $\boldsymbol{s}(\mathbf{r}) = \boldsymbol{s}(\mathbf{r}')$. Otherwise, a convex combination of $\mathbf{r}$ and $\mathbf{r}'$ would strictly increase L. $\boldsymbol{s}(\mathbf{r}) = \boldsymbol{s}(\mathbf{r}')$ requires the vector $\boldsymbol{\rho} = \mathbf{r} - \mathbf{r}'$ to satisfy $\sum_i \pi_{i,j} \rho_i = 0$ for each j, hence to be orthogonal to each experts' statements. We prove that $\boldsymbol{\rho}$ is null. Let us distinguish two cases according to the cardinalities of N and M.

Let $n \leq m$. Under **(D2)** the dimension of the statements (in $\Re^n$) is n, and no non-null vector is orthogonal to all of them. (Note that the same argument shows that $\boldsymbol{s}$ is one-to-one.)

Let $n > m$. The argument involves an additional step. Let I be the set of indices for which the inequality (36) is satisfied as an equation. We have, using (37), $\sum_j \pi_{i,j} \psi_j(s_j) = \mu$ for each i in I and $\sum_j \pi_{i,j} \psi_j(s_j) < \mu$ for each i not in I. Under **(D1)**, the cardinality of I is not larger than m. The support of any solution to $(\mathcal{P})$ is included in I. Thus the vector $\boldsymbol{\rho}$ is null outside I, hence satisfies $\sum_{i \in I} \pi_{i,j} \rho_i = 0$ for each j: the vector ρ_I is orthogonal to the statements vectors restricted to I. **(D2)** implies that these vecors are linearly independent, so $\boldsymbol{\rho}_I$ must be null; this ends the proof. ∎

Lemma 4 *Let* $n \leq m$ *and assume* **(D2)**. *Let the* ψ_k *be strictly increasing. The support of a rest point is a singleton. The process always converges but the limit point may depend on the initial ranking.*

Proof of Lemma 4. The assumptions $n \leq m$ and **(D2)** imply that L is one-to-one, as seen in the proof of Lemma 3. Hence L is strictly convex in $\mathbf{r}$. This implies that any solution to $(\mathcal{P})$ has all its components but one equal to zero. (An interior point satisfying the first order conditions corresponds to a minimum of L.)

To prove convergence, we show that the sequence $L(\mathbf{r}^{(t)})$ strictly increases with t as long as $\mathbf{r}^{(t)}$ differs from $\mathbf{r}^{(t-1)}$. Consider the difference $L(\mathbf{r}^{(t)}) - L(\mathbf{r}^{(t-1)})$. The strict convexity of L implies

$$L(\mathbf{r}^{t}) - L(\mathbf{r}^{(t-1)}) \geq \sum_i \frac{\partial L}{\partial r_i}(\mathbf{r}^{(t-1)})[r_i^{(t)} - r_i^{(t-1)}] \text{ with } > \text{ if } \mathbf{r}^{(t)} \neq \mathbf{r}^{(t-1)}. \tag{41}$$

We first show that

$$\frac{\partial L}{\partial r_i}(\mathbf{r}^{(t-1)})[r_i^{(t)} - r_i^{(t-1)}] \geq r_i^{(t-1)}[\frac{\partial L}{\partial r_i}(\mathbf{r}^{(t-1)}) - V^{(t-1)}]. \tag{42}$$

By (40) we have

$$\frac{\partial L}{\partial r_i}(\mathbf{r}^{(t-1)})[r_i^{(t)} - r_i^{(t-1)}] = \frac{r_i^{(t-1)} \frac{\partial L}{\partial r_i}(\mathbf{r}^{(t-1)})}{V^{(t-1)}}[\frac{\partial L}{\partial r_i}(\mathbf{r}^{(t-1)}) - V^{(t-1)}].$$

The term $[\frac{\partial L}{\partial r_i}(\mathbf{r}^{(t-1)}) - V^{(t-1)}]$ is positive (resp. negative) if the partial derivative $\frac{\partial L}{\partial r_i}$ is larger (resp. smaller) than $V^{(t-1)}$; hence we always have

$$\frac{\frac{\partial L}{\partial r_i}(\mathbf{r}^{(t-1)})}{V^{(t-1)}}[\frac{\partial L}{\partial r_i}(\mathbf{r}^{(t-1)}) - V^{(t-1)}] \geq [\frac{\partial L}{\partial r_i}(\mathbf{r}^{(t-1)}) - V^{(t-1)}].$$

This proves inequality (42). Summing these inequalities over i yields

$$\sum_i \frac{\partial L}{\partial r_i}(\mathbf{r}^{(t-1)})[r_i^{(t)} - r_i^{(t-1)}] \geq \sum_i r_i^{(t-1)} \frac{\partial L}{\partial r_i}(\mathbf{r}^{(t-1)}) - V^{(t-1)} = 0$$

Hence from (41), the value of L strictly increases as long as $\mathbf{r}^{(t)}$ differs from $\mathbf{r}^{(t-1)}$, that is as long as a rest point is not reached: the sequence converges to a rest point. ■

Extension to influence functions $B(x) = x^\alpha$, $\alpha < 1$. Let L be defined by $L(\mathbf{r}) = \sum_j \Psi_j(\sum_i \pi_{i,j} r_i^\alpha)$ where Ψ_j is a primitive of ψ_j. Consider as above the program $\mathcal{P}$ of maximization of L over Δ_N. By similar computation as for the linear influence function, one checks that the first order conditions of $\mathcal{P}$ are

$$\sum_j \pi_{i,j} r_i^{\alpha-1} \frac{\psi_j(s_j)}{\sum_k s_k \psi_k(s_k)} = 1 \text{ each } i \text{ with } s_j = \sum_i \pi_{i,j} r_i^\alpha. \tag{43}$$

Under the assumption on q_j these conditions write $\sum_j \frac{\pi_{i,j} r_i^{*\,1-\alpha}}{\sum_\ell \pi_{\ell,j} r_\ell^{*\alpha}} q_j(\mathbf{r}^{*\alpha}) = 1$ for each i, which coincide with the conditions (32) on a rest point. For $\alpha < 1$ L is strictly concave in $\mathbf{r}$ because the sum functions $s_j(\mathbf{r}^\alpha) = \sum_i \pi_{i,j} r_i^\alpha$ are. Thus there is a unique rest point. ■

Proof of Proposition 2. We proved in the text that conditions (28) are necessary for the existence of a rest point with support I. Let us show the converse.

Assume that $\boldsymbol{x}$ satisfies (28). Given a vector $\mathbf{r}_I$ in Δ_I let $\widehat{\mathbf{r}}_I = (\mathbf{r}_I, \mathbf{0}_{N-I})$. $\widehat{\mathbf{r}}_I$ is in Δ_N with null components outside I. Observe that by the peers' property $\boldsymbol{q}(\widehat{\mathbf{r}}_I)$ is also null outside I. We prove the existence of a positive vector $\mathbf{r}_I$ in Δ_I that satisfies (17):

$$x_j = \frac{q_j}{s_j}(\widehat{\mathbf{r}}_I) \text{ for each } j \in I.$$

Define the correspondence Φ from Δ_I to itself by

$$\Phi(\mathbf{r}_I) = \{\mathbf{r} \in \Delta_I \text{ s.t. } r_k = 0 \text{ for each } k \text{ that does not minimize } \frac{q_j(\widehat{\mathbf{r}}_I)}{x_j s_j(\widehat{\mathbf{r}}_I)} \text{ over } j \in I\}. \quad (44)$$

It is easy to check that the correspondence Φ is convex-valued and that the continuity of the function $\boldsymbol{q}$ implies that Φ is upper hemi-continuous. Therefore, Φ has a fixed point by Kakutani theorem, $\mathbf{r}_I^* \in \Phi(\mathbf{r}_I^*)$. We prove that $\mathbf{r}_I^*$ is positive and satisfies (17).

Let us show that $\mathbf{r}_I^*$ is positive. By contradiction, assume $r_i^* = 0$ for some i in I. The peers' property implies that $q_i(\widehat{\mathbf{r}}_I^*)$ is null, which gives a null value for the minimum of the $\frac{q_j(\widehat{\mathbf{r}}_I^*)}{x_j s_j(\widehat{\mathbf{r}}_I^*)}$ over j in I. Now since $\sum_{k \in I} q_k(\widehat{\mathbf{r}}_I^*) = 1$, there must be k in I with $q_k(\widehat{\mathbf{r}}_I^*) > 0$. Such a k surely does not achieve the minimum of $\frac{q_j(\widehat{\mathbf{r}}_I^*)}{x_j s_j(\widehat{\mathbf{r}}_I^*)}$, so any $\mathbf{r}$ in $\Phi(\mathbf{r}_I^*)$ has $r_k = 0$. Since $q_k(\widehat{\mathbf{r}}_I^*) > 0$, applying the peers' property again, r_k^* must be positive, hence $\mathbf{r}_I^* \in \Phi(\mathbf{r}_I^*)$ cannot hold, which gives the contradiction.

It follows that $\Phi(\mathbf{r}_I^*)$ contains a strictly positive vector $(\mathbf{r}_I^*)$. The definition (44) implies that the ratios $\frac{q_j(\widehat{\mathbf{r}}_I^*)}{x_j s_j(\widehat{\mathbf{r}}_I^*)}$ are equalized across j: there is some λ such that $q_j(\widehat{\mathbf{r}}_I^*) = \lambda x_j s_j(\widehat{\mathbf{r}}_I^*)$ for each j. Summing over j yields $\sum_{j \in I} q_j(\hat{\mathbf{r}}_I^*) = \lambda \sum_{j \in I} x_j s_j(\widehat{\mathbf{r}}_I^*)$. The left hand side is equal to 1; using $s_j(\widehat{\mathbf{r}}_I^*) = \sum_{i \in I} \pi_{i,j} r_i^*$ and exchanging the order of summation in the right hand side, we obtain

$$1 = \lambda \sum_{i \in I} r_i^* (\sum_{j \in I} \pi_{i,j} x_j).$$

By (28), each sum $\sum_{j \in I} \pi_{i,j} x_j$ is equal to 1 for i in I. Using $\sum_{i \in I} r_i^* = 1$ yields $1 = \lambda$, which proves $\frac{q_j}{s_j}(\widehat{\mathbf{r}}_I^*) = x_j$ for each j, as desired. ■

Proof of Proposition 3. Let a peers' method. The proof is divided in two steps.

The first step states conditions on a matrix $\boldsymbol{\pi}$ under which the scores on $N - I$ converge to zero if their initial values are low enough, whatever the behavior of the scores on I. The second step shows that these conditions are surely satisfied under the assumptions (29) of the Proposition.

Step 1. Let a matrix $\boldsymbol{\pi}$ be such that for some $k < 1$,

$$\sum_j \frac{\pi_{i,j}}{s_j(\mathbf{r})} q_j(\mathbf{r}) \leq k \text{ for each } i \notin I \text{ and each } \mathbf{r} \text{ null on } N - I. \quad (45)$$

We show that, starting from rankings with small enough values outside I, the growth rate of the scores on $N - I$ is strictly smaller than 1, hence the scores converge to zero.

Recall that the growth rate of i's score satisfies

$$\frac{r_i^{(t+1)}}{r_i^{(t)}} = \sum_j \frac{\pi_{i,j}}{s_j(\mathbf{r}^{(t)})} q_j(\mathbf{r}^{(t)}). \quad (46)$$

The continuity of $\boldsymbol{q}$ ensures that similar inequalities to (45) hold for $\mathbf{r}$ with components that are small enough on $N - I$. Formally, given $\epsilon > 0$ let $\mathcal{V}(\epsilon)$ denote the subset of Δ_N composed with the rankings whose components on $N-I$ are smaller than ϵ. For k' with $k<k'<1$, there is $\epsilon>0$ such that

$$\sum_j \frac{\pi_{i,j}}{s_j(\mathbf{r})} q_j(\mathbf{r}) \leq k', i \notin I, \text{ for each } r \in \mathcal{V}(\epsilon). \tag{47}$$

Using (46), this inequality implies that the growth rates of all components on $N-I$ are less than k'. Assume that the ranking belongs to $\mathcal{V}(\epsilon)$ at some date t. Since k' is strictly smaller than 1, the ranking at date $t+1$ also belongs to $\mathcal{V}(\epsilon)$. By induction, the sequence stays in $\mathcal{V}(\epsilon)$ at any further date. Furthermore, the components on $N-I$ converge to zero because their growth rates are smaller than k'.

Step 2. Let us prove that the conditions (29) imply (45). Recall (29): for some $k<1$

$$\frac{\beta}{\alpha} \leq k \text{ where } \beta = \max_{(i,j)\in N-I\times I} \pi_{i,j} \text{ and } \alpha = \min_{(i,j)\in I\times I} \pi_{i,j}$$

Let $\mathbf{r}$ with null components on $N-I$. Consider i not in I. We want to bound $\sum_j \frac{\pi_{i,j}}{s_j(\mathbf{r})} q_j(\mathbf{r})$. Note first that $q_j(\mathbf{r})=0$ for any j not in I by definition of a peers' method. The inequality $\pi_{i,j} \leq \beta$ for j in I thus yields

$$\sum_j \frac{\pi_{i,j}}{s_j(\mathbf{r})} q_j(\mathbf{r}) = \sum_{j\in I} \frac{\pi_{i,j}}{s_j(\mathbf{r})} q_j(\mathbf{r}) \leq \beta \sum_{j\in I} \frac{q_j(\mathbf{r})}{s_j(\mathbf{r})}.$$

Since $\mathbf{r}$ has null components on $N-I$, $s_j(\mathbf{r}) = \sum_{\ell\in I} \pi_{\ell,j} r_\ell$; the inequality $\pi_{i,j} \geq \alpha$ for each pair j and ℓ in I thus implies that for each j in I, $s_j(\mathbf{r}) \geq \alpha \sum_{\ell\in I} r_\ell = \alpha$. Hence, $\sum_{j\in I} \frac{q_j(\mathbf{r})}{s_j(\mathbf{r})} \leq 1/\alpha \sum_{j\in I} q_j(\mathbf{r})$. $\boldsymbol{q}$ is in the simplex, so we finally obtain

$$\sum_j \frac{\pi_{i,j}}{s_j(\mathbf{r})} q_j(\mathbf{r}) \leq \frac{\beta}{\alpha} \text{ for each } i \notin I$$

This proves that (29) implies (45) by taking $k = \beta/\alpha$. ■

Proof of Corollary 1 There exist matrices with several locally stable points. To show this, note that the columns indexed by I in matrix $\boldsymbol{\pi}$ can be chosen so that there is a locally stable point with support I, independently of the columns indexed by $N-I$. To see this, observe that conditions (45), which ensure local convergence to zero of the components on $N-I$, only bear on the columns indexed by I. Furthermore the values of these columns on rows I can be chosen freely given the lower bound m so they can be chosen so that there is convergence within I (that is starting with a ranking null on $N-I$). Combining these two conditions, which bear only on the columns indexed by I, gives a locally stable point with support I. The values of the matrix on $N-I$ can also be chosen so that the same result hold on $N-I$: this ensures the existence of another stable point with support included in $N-I$. Clearly, the argument can be extended so as to show that several (more than two) stable points may exist. ■

References

Altman, A. and M. Tennenholtz (2005): “On the axiomatic foundations of ranking systems,” in Proc. 19th International Joint Conference on Artificial Intelligence, 917–922.

Amir, R. (2002): “Impact-adjusted Citations as a measure of Journal quality,” CORE discussion paper 74.

Barabasi, A.-L., R. Albert, and H. Jeong (1999): “Mean-field theory for scale-free random networks,” *Physica A: Statistical Mechanics and its Applications*, 272(1-2), 173–187.

Bacharach, M. (1965): “Estimating Nonnegative Matrices from Marginal Data,” *International Economic Review*, 6(3), 294–310.

Bonacich, P. (1987): “Power and centrality: a family of measures,” *American Journal of Sociology*, 92, 1170–1182.

Brin, S., and L. Page (1998): “The anatomy of large-scale hypertextual web search engine,” *Computer Networks and ISDN Systems*, 30(1-7), 107–117.

Cho, J., R. Roy, and R. Adams (2005): “Page quality: in search of an unbiased Web ranking,” Proceedings of the 2005 ACM SIGMOD, 551–562.

de Clippel, G., H. Moulin, and N. Tideman (2008): “Impartial division of a dollar,” *Journal of Economic Theory*, 139(1), 176–191.

DeGroot, M. H. (1974): “Reaching a Consensus,” *Journal of the American Statistical Association*, 69(345), 118–121.

Demange, G. (2012): “On the influence of a ranking system,” forthcoming in *Social Choice and Welfare*, DOI 10.1007/s00355-011-0631-5.

Demange, G. (2012): “A ranking method based on handicaps,” PSE Working paper.

DeMarzo, P.M., D. Vayanos, and J. Zwiebel (2003): “Persuasion Bias, Social Influence, and Unidimensional Opinions,” *Quarterly Journal of Economics*, 118(3), 909–968.

Fortunato, S., A. Flammini, F. Menczer, and A. Vespignani (2006): “The egalitarian effect of search engines,” Proc. Natl. Acad. Sci. USA, 103(34), 12684–12689,

Golub, B. and M. Jackson (2008): “How Homophily Affects Communication in Networks,” Stanford working paper.

Goyal, S. (2005): “Learning in networks,” in *Group Formation in Economics: Networks, Clubs and Coalitions* eds. Demange, G. and M. Wooders, Cambridge University Pres.

Katz, L. (1953): “A new status index derived from sociometric analysis,” *Psychometrika*, 18, 39–43.

Kleinberg, N. (1999): “Authoritative sources in a hyperlinked environment,” *Journal of the ACM*, 46(5).

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,” VLDP Conference, 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.

Slutzki, G. and O. Volij (2006): “Scoring of Web pages and tournaments-axiomatizations,” *Social choice and welfare* 26, 75–92.