Bibliometric rankings of journals based on Impact Factors: An axiomatic approach☆

Denis Bouyssou[a,b,*], Thierry Marchant[c]

[a] *CNRS–LAMSADE, FRE3234, F-75775 Paris Cedex 16, France*
[b] *Université Paris Dauphine, F-75775 Paris Cedex 16, France*
[c] *Ghent University, H. Dunantlaan, 1, B-9000 Gent, Belgium*

Abstract

This paper proposes an axiomatic analysis of Impact Factors when used as tools for ranking journals. This analysis draws on the similarities between the problem of comparing distribution of citations among papers and that of comparing probability distributions on consequences as commonly done in decision theory. Our analysis singles out a number of characteristic properties of the ranking based on Impact Factors. We also suggest alternative ways of using distributions of citations to rank order journals.

Keywords: Bibliometrics, Journal rankings, Impact Factor, Expected Utility, Decision Theory

---

1. Introduction

The traditional way to evaluate research is to rely on peer judgement. Because this evaluation technique is costly and may suffer from a number of problems (Campanario 1998a,b), the bibliometric literature has developed

☆Authors are listed alphabetically. During the preparation of this paper, Thierry Marchant benefited from a visiting position at the *Université Paris Dauphine*, Paris, and the *Indian Statistical Institute*, New Delhi. The work of Denis Bouyssou was partially supported by a grant from the *Agence Nationale de la Recherche* (Project ComSoc, ANR-09-BLAN-0305). This support is gratefully acknowledged. We would like to thank Ludo Waltman and an anonymous referee for their very useful comments on earlier drafts of this paper.

*Corresponding author

Email addresses: `bouyssou@lamsade.dauphine.fr` (Denis Bouyssou), `thierry.marchant@UGent.be` (Thierry Marchant)

alternative tools, mainly based on various ways of counting citations (Garfield 1955, 1972, 1979). As noted by van Raan (2005, p. 2), this "unavoidably introduces a bibliometrically limited view of a complex reality".

Among the numerous bibliometric indices that have been proposed in the literature, Impact Factors (IFs) of journals stand out as being among the oldest and the most widely used when it comes to evaluate journals (recent references using IFs for ranking journals include Bar-Ilan 2010, and Franceschet 2010). Glänzel and Moed (2002, p. 172) describe IFs as "a fundamental citation-based measure for significance and performance of scientific journals". We refer the reader to Glänzel and Moed (2002) for a thorough overview on IFs. Archambault and Larivière (2009) and Garfield (2006) detail the history and origins of IFs. Roughly speaking, the IF of a journal gives the mean number of citations received by papers published in this journal.

IFs are not the only bibliometric indices that have been introduced in the literature. Recent years have seen a flourishing of such indices, including, e.g., the h-index (Hirsch 2005) and the g-index (Egghe 2006). These new indices have almost immediately after their introduction been studied from an axiomatic standpoint (Marchant 2009a,b, Quesada 2009, Rousseau 2008b, Woeginger 2008a,b,c). Curiously, this axiomatic literature has not paid much attention to IFs, although they are much older and more widely used.

An exception is Bouyssou and Marchant (2010) who studied the problem of consistently ranking authors and journals. The study of consistent rankings of authors and journals requires a rather rich framework that includes three different sets (authors, papers, and journals) and three different binary relations (indicating authorship, citations, and publication media). Within this framework, we presented axioms implying that journals are ranked according to IFs and authors according to the number of citations received by the papers they have signed (with an eventual correction for co-authors). This analysis uses many axioms because of the richness of the framework. Furthermore, because axioms related to journals and axioms related to authors are linked by a consistency condition, they interact. It is therefore not easy to derive from the results in Bouyssou and Marchant (2010) what are the conditions needed to characterize the ranking of journals using IFs, independently of what happens with authors. This is the purpose of the present paper. We will use a simple framework that only involves journals and present a set of conditions implying that journals are ranked using IFs.

Our present treatment of the ranking of journals based on IFs rests on a fairly simple intuition. In order to compute the IF of a journal, it is only necessary to know how many papers published in this journal have received x citations, for all integer $x \in \mathbb{N}$, i.e., the *distribution of citations* for the journal. Comparing distributions of citations bears a striking resemblance with the problem of comparing probability distributions on consequences, a classical problem in decision theory. Exploiting this similarity will allow us to provide a simple axiomatic foundation to the ranking of journals based on IFs. This will also lead us to suggest alternative rankings that use generalizations of IFs. On the technical side, while Marchant (2009b) emphasizes the power of results on "extensive measurement" (Krantz, Luce, Suppes, and Tversky 1971) to deal with the ranking of authors, we will show that results on "expected utility" (Fishburn 1970) are much relevant to analyze the ranking of journals using IFs. We hope that this will shed a new light on the discussion concerning the pros and cons of this bibliometric ranking of journals by making explicit the conditions underlying it.

This paper is organized as follows. Section 2 introduces our framework and notation. Section 3 presents the main conditions used in the paper. Section 4 characterizes a class of bibliometric rankings of journals that includes the one based on IFs as a particular case. Section 5 specializes this analysis to characterize the ranking based on IFs. Section 6 concludes. Appendix A illustrates some of our findings using citation data for a small sample of economic journals.

2. Notation and definitions

A published paper may receive citations from other papers. We take here these citations as given and we do not discuss the various ways in which such citations can or should be computed (e.g., what is the set of journals that should be included in the database, what is the relevant time window to collect citations, how should we deal with self-citations, etc.). The only aspect of papers that is taken into account in our analysis is the number of citations (possibly zero) that they receive.

We view a journal as a set of papers, consistently with the approach used in Marchant (2009a,b) to deal with authors. For our purposes, we may model a journal a as a function from $\mathbb{N}$ to $\mathbb{N}$ where, for all $x \in \mathbb{N}$, $a(x)$ is interpreted as the number of papers published in journal a having received exactly x citations. Hence, we identify a journal with its distribution of citations. We will restrict our attention to journals for which there are only a finite number of $y \in \mathbb{N}$ such that $a(y) \neq 0$. We define the set of all journals

$\mathcal{J} = \{a, b, \dots\}$ as the set of all functions from $\mathbb{N}$ to $\mathbb{N}$ satisfying the above constraint.

Using the above notation, it is clear that journal $a \in \mathcal{J}$ has published a number of papers given by

$$N(a) = \sum_{x \in \mathbb{N}} a(x).$$

The number of citations received by the papers published in journal $a \in \mathcal{J}$ is simply given by:

$$C(a) = \sum_{x \in \mathbb{N}} x a(x).$$

By definition, the Impact Factor $IF(a)$ of journal $a \in \mathcal{J}$ represents the average number of citations received by the papers published in a, i.e.,

$$IF(a) = \frac{C(a)}{N(a)} = \frac{\sum_{x \in \mathbb{N}} x a(x)}{\sum_{x \in \mathbb{N}} a(x)}.$$

We will be concerned here with the ranking that is induced on the set $\mathcal{J}$ of all journals by the values IF, i.e., the binary relation $\succsim_{IF}$ defined letting, for all $a, b \in \mathcal{J}$,

$$a \succsim_{IF} b \Leftrightarrow IF(a) \geq IF(b). \tag{1}$$

Our main objective will be to characterize the relation $\succsim_{IF}$ within the class of all possible binary relations on $\mathcal{J}$.

It may be useful to stress a number of important points about our framework. Observe first that the set of journals $\mathcal{J}$ includes infinitely many journals since journals having all possible distributions of citations belong to $\mathcal{J}$. Such a rich set of journals $\mathcal{J}$ will allow us to keep the analysis simple (technically, we use here a "single profile" approach as in Marchant (2009a,b), contrary to the "multi-profile" approach used in Bouyssou and Marchant (2010)). Second, our framework is such that it makes sense to "add" two journals or to "multiply" a journal by an integer. Indeed, if $a, b \in \mathcal{J}$, it is clear that $a+b \in \mathcal{J}$ and $n \cdot a \in \mathcal{J}$, for all $n \in \mathbb{N}$ (i.e., for all $a, b \in \mathcal{J}$ and for all $n \in \mathbb{N}$, there are $c, d \in \mathcal{J}$ such that $c(x) = a(x) + b(x)$ and $d(x) = na(x)$, for all $x \in \mathbb{N}$). Third, our framework is such that we will be unable to deal with the important question of the field normalization of IFs (Adler, Ewing, and Taylor 2009, Zitt 2010). Let us finally note that this framework is inadequate to analyze models in which the "weight" of a citation could depend on the citing paper, e.g., through the journal in which the citing paper is published, as in Palacios-Huerta and Volij (2004). A framework that would allow us to deal with such models would be far more complex than the present one, while not being much useful to understand the characteristic properties of the relation $\succsim_{IF}$.

3. Axioms

Let $\succsim$ be a binary relation on the set $\mathcal{J}$ of all journals[1]. We interpret $\succsim$ as an "at least as good as" preference relation between journals (throughout the paper, we use the term "preference" to indicate what some readers may prefer to call "impact"). The relation $\succ$ denotes the asymmetric part of $\succsim$, i.e., the binary relation on $\mathcal{J}$ such that $a \succ b$ if $[a \succsim b$ and $Not[b \succsim a]]$. We interpret $\succ$ as a "strict preference" relation between journals. The relation $\sim$ denotes the symmetric part of $\succsim$ (i.e., $a \sim b$ if $[a \succsim b$ and $b \succsim a]$) and is interpreted as an "indifference" relation between journals.

We consider several conditions on a binary relation $\succsim$ on $\mathcal{J}$.

Axiom 1 (Weak order)**.** *The binary relation $\succsim$ on $\mathcal{J}$ is a weak order, i.e., a* complete *(for all $a, b \in \mathcal{J}$, $a \succsim b$ or $b \succsim a$) and* transitive *(for all $a, b, c \in \mathcal{J}$, $[a \succsim b$ and $b \succsim c]$ imply $a \succsim c$) binary relation.*

It is clear that the relation $\succsim_{IF}$ defined by (1) is a weak order.

In decision theory, starting with a preference relation that is a weak order is so classical that the above condition hardly needs an elaboration in this context. The situation is more complex here since we are dealing with a bibliometric comparison of journals. Basically, this condition states that we want a model able to compare *all* journals in a *consistent* way. This is a strong requirement. For instance, for rankings based on an index, one may consider that small differences in the value of the index should be neglected. This leads to intransitive relations $\sim$ and, hence, $\succsim$. Similarly, Adler et al. (2009) argue that a sensible evaluation of journals should use *several* indices. This will clearly lead to incomplete relations if using several indices means taking the intersection of the rankings induced by each of them.

Axiom 2 (Homogeneity)**.** *For all $a \in \mathcal{J}$ and all $n \in \mathbb{N}$ such that $n > 0$, $a \sim n \cdot a$.*

[1] i.e., a subset of $\mathcal{J} \times \mathcal{J}$. As is usual, we will write $a \succsim b$ instead of $(a, b) \in \succsim$.

It is clear that the relation $\succsim_{IF}$ defined by (1) is homogeneous.

Homogeneity wishes to capture the initial motivation of Eugene Garfield to create an index that would be *size-independent* (as explained, e.g., in Garfield 2006) through the normalization of the numbers of citations received by the size of the journal, measured by the number of papers it publishes. Without such a condition, a journal might well try to modify its ranking by asking authors to split all their papers into smaller units (e.g., requiring all papers to have a "part 1" and a "part 2"). In our framework, a journal is identified with its distribution of citations. When Homogeneity is admitted, two journals having the same distribution of relative frequencies of citations will be indifferent.

Axiom 3 (Independence)**.** *For all $a, b, c \in \mathcal{J}$ such that $N(a) = N(b) = N(c)$,*

$$a \sim b \Rightarrow a + c \sim b + c,$$
$$a \succ b \Rightarrow a + c \succ b + c.$$

It is easy to check that the above condition is satisfied by the relation $\succsim_{IF}$ defined by (1). Indeed, since $N(a) = N(c)$, we have

$$IF(a + c) = \frac{C(a) + C(c)}{N(a) + N(c)} = \frac{C(a) + C(c)}{2N(a)} = \frac{1}{2}IF(a) + \frac{1}{2}IF(c).$$

This condition says that the comparison of two journals of equal size (in terms of the number of papers that they publish) is not affected if each of them merges with the same journal having the same size. Intuitively, it is hard to imagine situations in which this condition would be undesirable. Nevertheless, this condition is far from being innocuous. For instance journal a may have no uncited papers, which may be the initial reason for stating that $a \succsim b$ in spite of the fact that b contains many highly cited papers. After the merger with c, the journal $a + c$, can have many uncited papers, which may clearly affect its comparison with $b + c$. The importance of this condition for bibliometric rankings has been stressed by Marchant (2009a,b), Rousseau (2008a) and Waltman and van Eck (2009a,b). Because of the field normalization that was used, the "crown indicator" of the Leiden group failed to satisfy it (see Waltman, van Eck, van Leeuwen, Visser, and van Raan 2010b, who use the word "consistency" instead of "independence"). This has lead the Leiden group to modify the definition of its crown indicator, using a different field normalization, so that the new indicator satisfies Independence (for more information on this new indicator, see also Waltman, van Eck, van Leeuwen, Visser, and van Raan 2010a).

Axiom 4 (Archimedean)**.** *For all* $a, b, c \in \mathcal{J}$ *such that* $N(a) = N(b) = N(c)$*,*

$$\left\{ \begin{array}{c} a \succ b \\ \textit{and} \\ b \succ c \end{array} \right\} \Rightarrow \left\{ \begin{array}{c} n \cdot a + m \cdot c \succ (n+m) \cdot b \\ \textit{and} \\ (n'+m') \cdot b \succ n' \cdot a + m' \cdot c \end{array} \right\}$$

for some $n, m, n', m' \in \mathbb{N}$ *that are strictly positive.*

It is easy to check that the above condition is satisfied by the relation $\succsim_{IF}$ defined by (1). Indeed, since $N(a) = N(c)$, we have

$$\begin{aligned} IF(n \cdot a + m \cdot c) &= \frac{nC(a) + mC(c)}{nN(a) + mN(c)} = \frac{nC(a) + mC(c)}{(n+m)N(a)} \\ &= \frac{n}{n+m} IF(a) + \frac{m}{n+m} IF(c). \end{aligned}$$

Hence, when $IF(a) > IF(b) > IF(c)$, it is clearly possible to find strictly positive $n, m \in \mathbb{N}$ such that $IF(n \cdot a + m \cdot c) = \frac{n}{n+m} IF(a) + \frac{m}{n+m} IF(c) > IF((n+m) \cdot b) = IF(b)$. A similar reasoning shows that the other part of the condition also holds.

This condition is technical and more difficult to interpret than the preceding ones. Take three journals of equal size $a, b, c \in \mathcal{J}$ with b being "between" a and c in terms of desirability. The condition introduces a new journal $n \cdot a + m \cdot c$ that consists of n times journal a plus m times journal c. This new journal "combines" the most preferred journal a with the least preferred one c. If n is large compared to m, the influence of the least desirable journal c in the journal $n \cdot a + m \cdot c$ will be small. Hence, when we compare it with journal $(n + m) \cdot b$, i.e., the intermediate journal scaled up to have the same size as the new journal, we should prefer the new journal. The interpretation of the other part of the condition is similar. The Archimedean condition would fail if some journals in $\mathcal{J}$ were "infinitely desirable", so that combining them with other ones in whatever manner would always lead to an infinitely desirable journal. Archimedean-like axioms are necessary as soon as one wishes to obtain a weak order that has a numerical representation (Krantz et al. 1971), which is clearly the case for $\succsim_{IF}$ defined by (1).

Axiom 5 (Increasingness)**.** *For all* $a, b \in \mathcal{J}$ *such that* $N(a) = N(b) = 1$ *and* $C(a) > C(b)$*,* $a \succ b$*.*

This condition is a very mild one. It says that when we compare journals having published a single paper, the one having received more citations should be preferred. This condition is clearly satisfied by the relation $\succsim_{IF}$ defined by (1).

4. Rankings using Generalized Impact Factors

In this section, we consider a family of bibliometric rankings that are very close to $\succsim_{IF}$ except that the value of a paper having received x citations is computed using an increasing function u. This leads to what we call Generalized Impact Factors (GIFs).

4.1. Result

The five conditions introduced in the preceding section are all that is needed to characterize relations comparing journals using GIFs.

Theorem 1. *A binary relation $\succsim$ on $\mathcal{J}$ satisfies Axioms 1, 2, 3, 4 and 5 if and only if (iff) there is an increasing real-valued function u on $\mathbb{N}$ such that, for all $a, b \in \mathcal{J}$,*

$$a \succsim b \Leftrightarrow GIF_u(a) = \frac{\sum_{x\in\mathbb{N}} u(x)a(x)}{\sum_{x\in\mathbb{N}} a(x)} \geq GIF_u(b) = \frac{\sum_{x\in\mathbb{N}} u(x)b(x)}{\sum_{x\in\mathbb{N}} b(x)}. \qquad (2)$$

The function u is unique up to scale and location, i.e., if there are two increasing real-valued functions u and v on $\mathbb{N}$ such that (2) holds, then $u = \gamma v + \delta$, for some $\gamma, \delta \in \mathbb{R}$ with $\gamma > 0$.

Going from model (2) to model (1) requires supposing that the function u is affine, i.e., that, for all $x \in \mathbb{N}$, $u(x) = \alpha x + \beta$ with $\alpha, \beta \in \mathbb{R}$ and $\alpha > 0$. We discuss below the interpretation that can be given to the function u in model (2).

Theorem 1 shows five conditions suffices to characterize the family of bibliometric ranking based on GIFs. Let us observe that if someone suggests to rank order journals using a new bibliometric index that is not a GIFs, the resulting ranking will necessarily violate at least one the above five conditions. This is one possible virtue of this result.

The proof of Theorem 1 exploits the similarity between our problem and classical results in expected utility theory. We first use the Homogeneity condition to transform the problem of comparing journals, i.e., functions from $\mathbb{N}$

to $\mathbb{N}$, into a problem of comparison of probability distributions on $\mathbb{N}$ having a finite support with rational components, i.e., using only probabilities in $\mathbb{Q}$. We then use a result in Shepherdson (1980) that gives necessary and sufficient conditions for an expected utility representation of a relation comparing probability distributions with finite support and rational components.

Proof of Theorem 1. Necessity is easily established. We prove sufficiency.

Let us associate to each journal $a \in \mathcal{J}$ a probability distribution π_a with finite support and rational components on $\mathbb{N}$ defining, for all $x \in \mathbb{N}$,

$$\pi_a(x) = \frac{a(x)}{N(a)},$$

i.e., $\pi_a(x)$ is the proportion of papers published by $a \in \mathcal{J}$ having received x citations.

Suppose that $a, b \in \mathcal{J}$ are such that $\pi_a = \pi_b$. Using the Homogeneity condition, we know that $a \sim a' = N(b) \cdot a$ and $b \sim b' = N(a) \cdot b$. It is clear that $a'(x) = b'(x)$, for all $x \in \mathbb{N}$, so that $a' = b'$ and, thus, $a' \sim b'$. Using the transitivity of $\sim$, we therefore conclude that $\pi_a = \pi_b \Rightarrow a \sim b$.

Let us define a relation $\succsim^*$ on the set Δ of all probability distributions with finite support and rational components on $\mathbb{N}$, letting, for all $p, q \in \Delta$,

$$p \succsim^* q \Leftrightarrow [a \succsim b, \text{ for all } a, b \in \mathcal{J} \text{ such that } p = \pi_a \text{ and } q = \pi_b].$$

We claim that, for all $a, b \in \mathcal{J}$,

$$a \succsim b \Leftrightarrow \pi_a \succsim^* \pi_b.$$

Indeed, suppose that $a \succsim b$ and that $Not[\pi_a \succsim^* \pi_b]$. This implies that there are $c, d \in \mathcal{J}$ such that $\pi_a = \pi_c$, $\pi_b = \pi_d$ and $Not[c \succsim d]$. Because $\pi_a = \pi_c$, we know that $a \sim c$. Similarly, $\pi_b = \pi_d$ implies $b \sim d$. Because $c \sim a$, $a \succsim b$, and $b \sim d$, the transitivity of $\succsim$ implies that $c \succsim d$, a contradiction. Furthermore, by construction, it is impossible that $\pi_a \succsim^* \pi_b$ and that $Not[a \succsim b]$. This proves the claim.

Suppose now that the relation $\succsim^*$ on Δ satisfies the following three conditions:

i– $\succsim^*$ is complete and transitive,

ii– for all $p, q, r \in \Delta$, $p \succ^* q$ implies $\frac{1}{2}p + \frac{1}{2}r \succ^* \frac{1}{2}q + \frac{1}{2}r$ and $p \sim^* q$ implies $\frac{1}{2}p + \frac{1}{2}r \sim^* \frac{1}{2}q + \frac{1}{2}r$, *iii*– for all $p, q, r \in \Delta$, $p \succ^* q$ and $q \succ^* r$ implies $\lambda p + (1-\lambda) r \succ^* q$ and $q \succ^* \mu p + (1-\mu) r$, for some $\lambda, \mu \in (0,1) \cap \mathbb{Q}$.

Using Shepherdson (1980, corollary 5.2, p. 106), we know that there is a real-valued function u on $\mathbb{N}$ such that, for all $p, q \in \Delta$,

$$p \succsim^* q \Leftrightarrow \sum_{x \in \mathbb{N}} p(x) u(x) \geq \sum_{x \in \mathbb{N}} q(x) u(x),$$

where u is unique up to scale and location.

Hence, we have, for all $a, b \in \mathcal{J}$,

$$\begin{aligned} a \succsim b &\Leftrightarrow \pi_a \succsim^* \pi_b \\ &\Leftrightarrow \sum_{x \in \mathbb{N}} \pi_a(x) u(x) \geq \sum_{x \in \mathbb{N}} \pi_b(x) u(x) \\ &\Leftrightarrow \frac{\sum_{x \in \mathbb{N}} u(x) a(x)}{\sum_{x \in \mathbb{N}} a(x)} \geq \frac{\sum_{x \in \mathbb{N}} u(x) b(x)}{\sum_{x \in \mathbb{N}} b(x)}. \end{aligned}$$

Let us show that, under Axioms 1, 2, 3, 4 on $\succsim$, conditions *i*, *ii*, and *iii* on $\succsim^*$ hold.

Observe first that, for all $p \in \Delta$, there is an $a \in \mathcal{J}$ such that $p = \pi_a$. Since we know that $a \succsim b \Leftrightarrow \pi_a \succsim^* \pi_b$, it is clear that $\succsim^*$ is complete and transitive.

Let $p, q, r \in \Delta$. There are $a, b, c \in \mathcal{J}$ such that $\pi_a = p$, $\pi_b = q$, $\pi_c = r$ and $N(a) = N(b) = N(c)$. Suppose now that $p \succ^* q$, so that $a \succ b$. Using the Independence condition, we know that $a + c \succ b + c$. Since it is clear that $\pi_{a+c} = \frac{1}{2}\pi_a + \frac{1}{2}\pi_c$, we obtain $\frac{1}{2}\pi_a + \frac{1}{2}\pi_c \succ^* \frac{1}{2}\pi_b + \frac{1}{2}\pi_c$, so that $\frac{1}{2}p + \frac{1}{2}r \succ^* \frac{1}{2}q + \frac{1}{2}r$. A similar reasoning shows that $p \sim^* q$ implies $\frac{1}{2}p + \frac{1}{2}r \sim^* \frac{1}{2}q + \frac{1}{2}r$.

Suppose now that $p \succ^* q$ and $q \succ^* r$. There are $a, b, c \in \mathcal{J}$ such that $\pi_a = p$, $\pi_b = q$, $\pi_c = r$ and $N(a) = N(b) = N(c)$. We know that $a \succ b$ and $b \succ c$. Using the Archimedean condition, we know that there are strictly positive $n, m \in \mathbb{N}$ such that $n \cdot a + m \cdot c \succ (n+m) \cdot b$. Let $\lambda = \frac{n}{n+m}$. We have $\lambda \in (0,1) \cap \mathbb{Q}$. It is easy to check that $\pi_{n \cdot a + m \cdot c} = \lambda \pi_a + (1-\lambda)\pi_c$ and $\pi_{(n+m) \cdot b} = \pi_b$. Hence, we have $\lambda \pi_a + (1-\lambda)\pi_c \succ^* \pi_b$, for some $\lambda \in (0,1) \cap \mathbb{Q}$. The proof of the other part of the condition is entirely similar.

Noting that Axiom 5 clearly implies that u must be increasing completes the proof. $\square$

4.2. Interpretation and uses of GIFs

Computing GIFs involves using an increasing function u. This raises the problem of the choice and interpretation of this function. We first examine the case in which GIFs are used with a *single* function u. We then consider using GIFs together with a whole family of functions u.

4.2.1. Attitudes towards dispersion

Consider, e.g., the case of two journals, $a, b \in \mathcal{J}$. Journal a has published two hundred papers, each of them being cited once. Journal b has also published two hundred papers. One hundred of them have never been cited. The other one hundred have all been cited twice. Comparisons based on IFs require that two distributions of relative frequencies of citations having the same mean must be judged indifferent. Since $IF(a) = IF(b) = 1$, it is impossible to distinguish the above two journals, although they have quite different distributions of citations. This is no more true using GIFs instead of IFs. Using the function $u(x) = \sqrt{x}$, we obtain $GIF_u(a) > GIF_u(b)$. Using the function $u(x) = x^2$, we have $GIF_u(b) > GIF_u(a)$. Concave functions u tend to favor distributions of citations that are centered around their mean, while an opposite effect occurs with convex functions. This should not be a surprise to readers familiar with expected utility theory (Fishburn 1970, Ch. 8) and the literature on "risk aversion" (Arrow 1965, Pratt 1964). In this framework, the concavity (resp. convexity) of u models "risk aversion" (resp. "risk proneness").

Distinguishing several "attitude towards risk" makes perfect sense in decision theory. But is this so in the area of bibliometrics? This paper does not wish to be a plea for the use of GIFs instead of IFs when comparing journals. Nevertheless, we think that the extra flexibility offered by GIFs when compared to IFs has much to offer.

Seglen (1997, p. 499) warns us that "Citation rate of article determines journal impact, but not vice versa". Nevertheless, the past distribution of citations of papers published in a journal may give *some* information on the selection process of papers in this journal and its visibility in the scientific community. Hence, it may give some information about the future citations of newly published papers in this journal. The empirical evidence about the reasonableness of such an hypothesis is mixed. The analysis in Seglen (1994) would tend to reject it. However, more recent studies lead to the reverse conclusion. Vieira and Gomes (2010, p. 10) mention that IFs influence future citations. The most impressive evidence in this direction was presented in

Larivière and Gingras (2010) who study the fate of identical papers published in different journals. This shows that IFs may have an influence on future citations. We think likely that the influence of past distributions of citations on future citations is not only through their means. Their entire shape may also matter. Notice that this would be in line with the famous *Matthew effect* described in Merton (1968). Authors willing to try to take advantage of the Matthew effect will tend to have a preference for distributions that are much dispersed around their mean: having one highly cited paper will attract many future citations and it makes sense to take risks to have one highly cited papers via a submission policy to journals having an extremely skewed distribution of citations. Other authors, e.g., because they are anxious to produce a relatively stable profile of citations over time, will tend to have a preference for distributions that are centered around their means. Such preferences are easily accounted using GIFs: convex (resp. concave) functions u will be able to model the preferences of authors having a favorable attitude towards dispersion (resp. concentration). More generally, the literature on risk aversion in expected utility theory can easily be transposed to our problem of comparing distributions of citations respectively replacing "risk attitudes" by the "attitude towards dispersion". This literature (Arrow 1965, Pratt 1964) has devised various indices to compare "degrees of risk aversion" and various remarkable classes of utility functions. It may therefore be of much help to guide the choice of an adequate function u in our context.

Observe finally that Axiom 5 can be weakened to cope with *nondecreasing* functions u. It suffices to modify it stating that, for all $a, b \in \mathcal{J}$ such that $N(a) = N(b) = 1$, $C(a) > C(b) \Rightarrow a \succsim b$. This modification clearly leaves the statement of Theorem 1 unchanged except that "increasing" is replaced by "nondecreasing". In this relaxed setting, GIFs contain as a particular case indices based on the proportion of "highly cited papers" published by a journal, where highly cited papers are interpreted as papers having received at least $k \in \mathbb{N}$ citations, where k is a fixed threshold. This is obvious considering a nondecreasing function u that is stepped, i.e., $u(x) = 0$, for all $x < k$ and $u(x) = 1$, for all $x \geq k$. As pointed out to us by a referee, indices based of the proportion of highly cited papers deserve attention since they are frequently used. This is particularly true if k is taken to be 1: in this case, the proportion of "highly cited papers" becomes one minus the proportion of uncited papers.

4.2.2. Stochastic dominance

Bibliometric distributions are highly skewed (Seglen 1992) and the distribution of citations to papers published in a journal are no exception (van Leeuwen and Moed 2005). However, there is no universal agreement in the bibliometric literature concerning the type of distributions that is prevalent (see, e.g., Bar-Ilan 2008, sect. 4.1 and Vieira and Gomes 2010, p. 2. One may also compare Egghe 2005, Seglen 1992, and van Raan 2001). Indeed, the data presented in Appendix A show that, for similar journals in the same field, rather different distributions can be observed. Comparing such diverse distributions taking only their mean into account may involve a huge loss of information. Hence, it may be wise to look for more robust ways to compare them. GIFs offer a simple way to do so, via the idea of *stochastic dominance* (the potential interest of this concept for bibliometrics was already stressed in Carayol and Lahatte 2009 and Waltman and van Eck 2009b). The underlying idea is quite simple. Instead of comparing journals using GIFs computed using a single function u (as is implicitly the case when using IFs), we might look for comparisons that are valid for a much larger set of functions u. Let us illustrate how such comparisons are performed using first and second order stochastic dominance (Fishburn and Vickson 1978, Levy 1992, 1998).

Let, for all $x \in \mathbb{N}$,

$$F_a^{(1)}(x) = \sum_{i=0}^{x} \pi_a(i),$$

i.e., $F_a^{(1)}(x)$ is the proportion of papers published in journal a having received at most x citations. First order stochastic dominance corresponds to the case in which $F_a^{(1)}(x) \leq F_b^{(1)}(x)$, for all $x \in \mathbb{N}$. Intuitively, the proportion of paper having less than x citations is greater for journal b than for journal a, for all x. In such a case, we know that $GIF_u(a) \geq GIF_u(b)$, for all increasing functions u on $\mathbb{N}$ (Fishburn and Vickson 1978).

Similarly defining, for all $x \in \mathbb{N}$,

$$F_a^{(2)}(x) = \sum_{i=0}^{x} F^{(1)}(i),$$

second order stochastic dominance corresponds to the case in which $F_a^{(2)}(x) \leq F_b^{(2)}(x)$, for all $x \in \mathbb{N}$. When this happens, we know that $GIF_u(a) \geq GIF_u(b)$, for all increasing and concave functions u on $\mathbb{N}$ (Fishburn and Vickson 1978).

The intuition behind this property is related to the fact that when two journals have the same IF, if $F_a^{(2)}(x) \leq F_b^{(2)}(x)$, for all $x \in \mathbb{N}$, then one can go from a to b through a series of "mean-preserving spreads" (Rothschild and Stiglitz 1970), i.e., transformations that increase the dispersion of the distribution around its mean.

Using the idea of stochastic dominance leads to robust journal comparisons based on bibliometric data. Doing so, we may however end up with many journals that remain incomparable. The results presented in Appendix A nevertheless show that these robust comparisons are not as incomplete as could have been expected.

4.3. Independence of the axioms in Theorem 1

The following five examples show that none of the conditions used in Theorem 1 is redundant.

Example 1 (Not Axiom 1). Let $\succsim^1$ and $\succsim^2$ be two binary relations on $\mathcal{J}$ satisfying the conditions of Theorem 1. Define the relation $\succsim = \succsim^1 \cap \succsim^2$. It is easy to check that $\succsim$ satisfies Axioms 2, 3, 4, and 5. It violates Axiom 1 as soon as there are $a, b \in \mathcal{J}$ such that $a \succ^1 b$ and $b \succ^2 a$, since, in this case, the relation $\succsim$ is not complete.

Example 2 (Not Axiom 2). Define the relation $\succsim$ on $\mathcal{J}$ letting, for all $a, b \in \mathcal{J}$,

$$a \succsim b \Leftrightarrow C(a) \geq C(b),$$

i.e., a ranking based on the number of citations received. It is easy to check that Axioms 1, 3, 4 and 5 are satisfied. Axiom 2 is clearly violated.

Example 3 (Not Axiom 3). Let u and v be two increasing real-valued functions on $\mathbb{N}$, with $v > 0$, that are not related by a positive affine transformation. Define the relation $\succsim$ on $\mathcal{J}$ letting, for all $a, b \in \mathcal{J}$

$$a \succsim b \Leftrightarrow \frac{GIF_u(a)}{GIF_v(a)} \geq \frac{GIF_u(b)}{GIF_v(b)}.$$

It is easy to check that Axioms 1, 2, 4, and 5 are satisfied. Axiom 3 is violated (Fishburn 1988, page 62).

Example 4 (Not Axiom 4). Let u and v be two increasing real-valued functions on $\mathbb{N}$ that are not related by a positive affine transformation. Define the relation $\succsim$ on $\mathcal{J}$ letting, for all $a, b \in \mathcal{J}$

$$a \succsim b \Leftrightarrow [GIF_u(a), GIF_v(a)] \geq^L [GIF_u(b), GIF_v(b)],$$

where $\geq^L$ is the lexicographic order on $\mathbb{R}^2$. It is easy to check that Axioms 1, 2, 3, and 5 are satisfied. Axiom 4 is violated. Indeed, suppose that $GIF_u(a) = GIF_u(b) > GIF_u(c)$ and $GIF_v(a) > GIF_v(b) > GIF_v(c)$. We have $a \succ b$ and $b \succ c$. But we have $(n+m) \cdot b \succ n \cdot a + m \cdot c$, for all strictly positive $n, m \in \mathbb{N}$.

Example 5 (Not Axiom 5)*.* Let u be a *decreasing* real-valued function on $\mathbb{N}$. Define $\succsim$ on $\mathcal{J}$ through (2). Axioms 1, 2, 3, and 4 are clearly satisfied. Axiom 5 is violated by construction.

5. Ranking using Impact Factors

This section investigates what must be added to the conditions used in Theorem 1 to guarantee that the function u becomes an affine function, implying that model (2) reduces to model (1).

Consider a journal $a \in \mathcal{J}$. Suppose that a paper published in a receives an additional citation. When journals are compared using IFs, it does not matter which of the papers published in a receives the additional citation. An additional citation to any paper published in a will increase $IF(a)$ by $1/N(a)$. This is characteristic of the ranking of journals using IFs. Since our framework is static, it is not flexible enough to analyze the case of a journal receiving an additional citation, we will consider different journals in $\mathcal{J}$ that will allow for such an interpretation.

Let $a \in \mathcal{J}$. Let $I(a) = \{x \in \mathbb{N} : a(x) \neq 0\}$, i.e., $I(a)$ is a set of integers x such that a has papers having been cited x times. For each $x \in I(a)$, we can build a distribution corresponding to a journal $a^{\uparrow x}$ that is identical to that of a except that a paper published by a that was cited x times has now received an additional citation. More precisely we have $a^{\uparrow x}(y) = a(y)$, for all $y \in \mathbb{N} \setminus \{x, x+1\}$, $a^{\uparrow x}(x) = a(x) - 1$ and $a^{\uparrow x}(x+1) = a(x+1) + 1$. Clearly, for all $x \in I(a)$, the journal $a^{\uparrow x}$ has the same number of papers as a and has one additional citation.

The next condition says that all journals built in this way should be indifferent, i.e., an additional citation has always the same impact whatever the paper that receives it.

Axiom 6 (Equal Impact of Additional Citations)**.** *For all $a \in \mathcal{J}$ and for all $x, y \in I(a)$, $a^{\uparrow x} \sim a^{\uparrow y}$*

It is clear this additional condition is satisfied by the ranking defined by (1). When added to the conditions used in Theorem 1, it implies that u on $\mathbb{N}$ is affine so that models (2) and (1) coincide.

Theorem 2. *A binary relation $\succsim$ on $\mathcal{J}$ satisfies Axioms 1, 2, 3, 4, 5 and 6 iff, for all $a, b \in \mathcal{J}$,*

$$a \succsim b \Leftrightarrow IF(a) \geq IF(b).$$

Proof. Necessity is clear. We show sufficiency. Using Theorem 1, we know that there is an increasing real-valued function u on $\mathbb{N}$ such that, for all $a, b \in \mathcal{J}$,

$$a \succsim b \Leftrightarrow GIF_u(a) \geq GIF_u(b).$$

Let $x \in \mathbb{N}$. Consider the journal $a_x \in \mathcal{J}$ such that such that, for all $y \in \mathbb{N}$, $a_x(y) = 0$ except that $a_x(x) = 1$ and $a_x(x+1) = 1$. Hence the journal a_x has two papers: one with x citations and the other with $x + 1$ citations. Let us apply Axiom 6 to a_x, implying that the following two journals are indifferent. The first journal, $a_x^{\uparrow x}$, has two papers, each one with $x + 1$ citations. The second journal, $a_x^{\uparrow x+1}$, has two papers, the first one having x citations and the second having $x + 2$ citations. This implies that the function u must be such that:

$$\frac{1}{2}u(x+1) + \frac{1}{2}u(x+1) = u(x+1) = \frac{1}{2}u(x) + \frac{1}{2}u(x+2).$$

Because this relation must hold for all $x \in \mathbb{N}$, it is clear that u must be affine in x. Because we know that u is increasing, it must be such that $u(x) = \alpha x + \beta$ with $\alpha, \beta \in \mathbb{R}$ and $\alpha > 0$. This completes the proof since ranking journals using such a function u is clearly equivalent to ranking journals according to their IFs. □

If someone suggests to rank order journals using a new bibliometric index that is not the IF and that uses the same information as in our model, we know that at least one of the six conditions used in Theorem 2 will be violated. A person willing to completely rank order journals will surely agree the Weak order condition. It is hard to imagine that someone could disagree with Increasingness in the context of ranking journals on the basis of bibliometric data. Moreover, we think that the need to use a size-independent index for comparing journals, as modelled by Homogeneity, is rather uncontroversial. Hence, if it is admitted that the Archimedean condition is mainly technical, the use of IFs seems to rest on the combined use of Independence and Equal Impact of Additional Citations.

The independence of the conditions used in Theorem 2 raises subtle questions. Since this is not central for our purpose, we will not detail this point here.

6. Discussion

This paper has analyzed the ranking of journals based on IFs using an axiomatic approach. We have given necessary and sufficient conditions (Theorem 2) for a binary relation comparing journals to coincide with the relation induced by IFs. We have also analyzed a general class of indices, called Generalized Impact Factors (GIFs), in which distributions of citations having equal means are not necessarily considered as indifferent (Theorem 1). This family of indices uses a function u that allows to capture various attitudes towards the dispersion of citations, as a von Neumann-Morgenstern utility function allows to capture several attitudes towards risk.

The axiomatic approach followed in this paper aims at characterizing a bibliometric ranking of journals by the conjunction of a small number of conditions. When these conditions have a clear intuitive content, we think that such an analysis may enlighten the discussion concerning the pros and cons of a bibliometric ranking. Indeed, discussions centered around axioms are often more precise than discussions based on the "intuitive reasonableness" of an index. A central condition in the characterization of the class of bibliometric rankings based on GIFs is Independence. This condition says that if two journals a and b of equal size both merge with a third journal c having the same size, the merger should not alter the comparison of a and b. Although this condition might sometimes be seen as being too strong (see Section 3), it appears reasonable in many cases. Yet, as observed by Marchant (2009a), Rousseau (2008a), and Waltman and van Eck (2009b), it is violated by the bibliometric ranking based on the h-index (Hirsch 2005) (the same is true for most bibliometric rankings based on variants of the h-index, see Waltman and van Eck 2009b). This gives a precise basis for comparing rankings based on GIFs with the ones based on the h-index or its variants. Our hope is that axiomatic analyses of the type exposed here, combined with systematic empirical studies of indices, will give a sound basis to compare and evaluate bibliometric rankings.

We would like to conclude with the mention of a few directions for future research and the indication of a number of important limitations of the present study.

Once it is recognized that the problem of comparing distributions of citations is very similar to the problem of comparing probability distributions, it is clear that any model used in decision making under risk can be transposed to the case of comparing journals, if Homogeneity is admitted. Some of these models are quite subtle, involving consistent violations of the Independence condition (see Gilboa 2009, Schmidt 2004, Sugden 2004, Wakker 2010, for overviews). In particular models stemming from Expected Utility with Rank Dependent Probabilities can offer interesting approaches to cope with outliers in the distribution of citations. This will be the subject of future research.

Although importing models for decision making under risk in the area of bibliometrics may give rise to new and interesting concepts, such a transposition also raises problems and questions. In decision making under risk, we model the preferences of a decision maker having to make a choice between risky prospects. The model wishes to be a useful guide to elicit her preferences and to help her choose. No such decision maker is present in bibliometrics. What is sought here are "indices" leading to rankings that would make sense to several persons (authors, editors, publishers, evaluators, etc) having potentially different perceptions and preferences. This seems to call for much caution in the above transposition. For instance, the weak order condition is almost universally accepted in decision theory (see Fishburn 1991, however). Its reasonableness is far less obvious when it comes to compare journals (Adler et al. 2009).

Another important limitation of our study is the static and rather rough character of our framework. It clearly does not allow to analyze the, often sweeping, consequences of publicizing a measure of impact of journals on the behavior of authors, editors, publishers or evaluators. It is clear that the use of IFs has had consequences on academia that go far beyond the formal properties of the ranking that it induces on journals (Adler et al. 2009, Monastersky 2005, Seglen 1997, Weingart 2005). These consequences are an essential part of the "Impact Factor debate" evoked in Bar-Ilan (2008, p. 22) and are left untouched by the analysis presented here.

References

Adler, R., Ewing, J., Taylor, P., 2009. Citation statistics. A report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the International Council of Institute of Mathematical Statistics (IMS). Statistical Science 24 (1), 1–14.

Archambault, É., Larivière, V., 2009. History of the journal impact factor: Contingencies and consequences. Scientometrics 79 (3), 635–649.

Arrow, K. J., 1965. Aspects of the theory of risk-bearing. Academic Bookstore, Yrjö Jahnsson Foundation, Helsinki, reprinted under the title "Essays in the theory of risk-bearing", 1971, North Holland, Amsterdam.

Bar-Ilan, J., 2008. Informetrics at the beginning of the 21st century—A review. Journal of Informetrics 2 (1), 1–52.

Bar-Ilan, J., 2010. Ranking of information and library science journals by JIF and by h-type indices. Journal of Informetrics 4 (2), 141–147.

Bouyssou, D., Marchant, T., 2010. Consistent bibliometric rankings of authors and journals. Journal of Informetrics 4 (3), 365–378.

Campanario, J. M., 1998a. Peer review for journals as it stands today, part 1. Science Communication 19 (3), 181–211.

Campanario, J. M., 1998b. Peer review for journals as it stands today, part 2. Science Communication 19 (4), 277–306.

Carayol, N., Lahatte, A., 2009. Dominance relations and universities ranking. Cahier du GREThA 2009-02, GREThA, Université Montesquieu Bordeaux IV.

Egghe, L., 2005. Power laws in the information production process: Lotkaian informetrics. Elsevier, Amsterdam.

Egghe, L., 2006. Theory and practice of the g-index. Scientometrics 69 (1), 131–152.

Fishburn, P. C., 1970. Utility theory for decision-making. Wiley, New York.

Fishburn, P. C., 1988. Nonlinear preference and utility theory. Johns Hopkins University Press, Baltimore.

Fishburn, P. C., 1991. Nontransitive preferences in decision theory. Journal of Risk and Uncertainty 4, 113–134.

Fishburn, P. C., Vickson, R. G., 1978. Theoretical foundations of stochastic dominance. In: Whitmore, G. A., Findlay, M. C. (Eds.), Stochastic Dominance: An Approach to Decision-Making Under Risk. Heath, Lexington, Massachusetts, pp. 37–113.

Franceschet, M., 2010. Journal influence factors. Journal of Informetrics 4 (3), 239–248.

Garfield, E., 1955. Citation indexes to science: A new dimension in documentation through the association of ideas. Science 122, 108–111.

Garfield, E., 1972. Citation analysis as a tool in journal evaluation. Science 178, 471–479.

Garfield, E., 1979. Is citation analysis a legitimate evaluation tool? Scientometrics 1, 359–375.

Garfield, E., 2006. The history and meaning of the journal impact factor. Journal of the American Medical Association 295, 90–93.

Gilboa, I., 2009. Theory of decision under uncertainty. No. 45 in Econometric Society Monographs. Cambridge University Press, Cambridge.

Glänzel, W., Moed, H. F., 2002. Journal impact measures in bibliometric research. Scientometrics 53 (2), 171–193.

Hirsch, J. E., 2005. An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences 102, 16569–16572.

Krantz, D. H., Luce, R. D., Suppes, P., Tversky, A., 1971. Foundations of measurement. Vol. 1: Additive and polynomial representations. Academic Press, New York.

Larivière, V., Gingras, Y., 2010. The impact factor's Matthew effect: a natural experiment in bibliometrics. Journal of the American Society for Information Science and Technology 61 (2), 424–427.

Levy, H., 1992. Stochastic dominance and expected utility: Survey and analysis. Management Science 38 (4), 555–593.

Levy, H., 1998. Stochastic dominance. Investment decision making under uncertainty. Kluwer, Boston.

Marchant, T., 2009a. An axiomatic characterization of the ranking based on the h-index and some other bibliometric rankings of authors. Scientometrics 80 (2), 327–344.

Marchant, T., 2009b. Score-based bibliometric rankings of authors. Journal of the American Society for Information Science and Technology 60, 1132–1137.

Merton, R. K., 1968. The Matthew effect in science. Science 159 (3810), 56–63.

Monastersky, R., 2005. The number that's devouring science. The Chronicle of Higher Education 52 (8), A12–A17.

Palacios-Huerta, I., Volij, O., 2004. The measurement of intellectual influence. Econometrica 72, 963–978.

Pratt, J. W., 1964. Risk aversion in the small and in the large. Econometrica 32, 122–136.

Quesada, A., 2009. Monotonicity and the Hirsch index. Journal of Informetrics 3 (2), 158–160.

van Raan, A. F. J., 2005. Measurement of central aspects of scientific research: performance, interdisciplinarity, structure. Measurement: Interdisciplinary Research and Perspectives 3 (1), 1–19.

Rothschild, M., Stiglitz, J. E., 1970. Increasing risk: I. A definition. Journal of Economic Theory 2, 225–243.

Rousseau, R., 2008a. Luckily, Science focuses on achivements. Some thoughts related to the h-index. ISSI Newsletter 4 (3), 49–50.

Rousseau, R., 2008b. Woeginger's axiomatisation of the h-index and its relation to the g-index, the $h^{(2)}$-index and the r^2-index. Journal of Informetrics 2 (4), 335–340.

Schmidt, U., 2004. Alternatives to expected utility: Some formal theories. In: Barberà, S., Hammond, P. J., Seidl, C. (Eds.), Handbook of Utility Theory. Vol. 2. Kluwer, Dordrecht, pp. 757–837.

Seglen, P. O., 1992. The skewness of science. Journal of the American Society for Information Science 43 (9), 628–638.

Seglen, P. O., 1994. Causal relationship between article citedness and journal impact. Journal of the American Society for Information Science 45 (1), 1–11.

Seglen, P. O., 1997. Why the impact factor of journals should not be used for evaluating research. BMJ 314, 498–502.

Shepherdson, J., 1980. Utility theory based on rational probabilities. Journal of Mathematical Economics 7, 91–113.

Sugden, R., 2004. Alternatives to expected utility: Foundations. In: Barberà, S., Hammond, P. J., Seidl, C. (Eds.), Handbook of Utility Theory. Vol. 2. Kluwer, Dordrecht, pp. 685–755.

Thomson Reuters, 2009a. Journal Citation Reports, 2008. `http://www.isiknowledge.com`, last accessed 1 December 2009.

Thomson Reuters, 2009b. Web of Science. `http://www.isiknowledge.com`, last accessed 1 December 2009.

van Leeuwen, T. N., Moed, H. F., 2005. Characteristics of journal impact factors: The effects of uncitedness and citation distribution on the understanding of journal impact factors. Scientometrics 63 (2), 357–371.

van Raan, A. F. J., 2001. Competition amongst scientists for publication status: Toward a model of scientific publication and citation distributions. Scientometrics 51 (1), 347–357.

Vieira, E. S., Gomes, J. A. N. F., 2010. Citation to scientific articles: Its distribution and dependence on the article features. Journal of Informetrics 4 (1), 1–13.

Wakker, P. P., 2010. Propect theory. Cambridge University Press, Cambridge, forthcoming.

Waltman, L., van Eck, N. J., 2009a. A simple alternative to the h-index. ISSI Newsletter 5 (3), 46–48.

Waltman, L., van Eck, N. J., 2009b. A taxonomy of bibliometric performance indicators based on the property of consistency. Technical Report ERS-2009-014-LIS, Erasmus University Rotterdam, Erasmus Research Institute of Management, Rotterdam, The Netherlands, presented at the 12th International Conference on Scientometrics and Informetrics, Rio de Janeiro, July 2009.

Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., van Raan, A. F. J., 2010a. Towards a new crown indicator: An empirical analysis. Tech. rep., Center for Science and Technology Studies, Leiden University, The Netherlands, arXiv:1004.1632v1.

Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., van Raan, A. F. J., 2010b. Towards a new crown indicator: Some theoretical considerations. Tech. rep., Center for Science and Technology Studies, Leiden University, The Netherlands, arXiv:1003.2167v1.

Weingart, P., 2005. Impact of bibliometrics upon the science system: Inadvertent consequences? Scientometrics 62 (1), 117–131.

Woeginger, G. J., 2008a. An axiomatic analysis of Egghe's g-index. Journal of Informetrics 2 (4), 364–368.

Woeginger, G. J., 2008b. An axiomatic characterization of the Hirsch-index. Mathematical Social Sciences 56, 224–232.

Woeginger, G. J., 2008c. A symmetry axiom for scientific impact indices. Journal of Informetrics 2 (4), 298–303.

Zitt, M., 2010. Citing-side normalization of journal impact: A robust variant of the Audience factor. Journal of Informetrics 4 (3), 392–406.

A. Citation data for some economic journals

We have studied citation data for eight economic journals: *American Economic Review* (AER), *Econometrica* (Eca), *Journal of Economic Literature* (JEL), *Journal of Economic Perspectives* (JEP), *Journal of Financial Economics* (JFE), *Journal of Political Economy* (JPE), *Quarterly Journal of Economics* (QJE), *Review of Economic Studies* (RES). These eight journals all have high IFs in the 2008 edition of the JCR (Thomson Reuters 2009a). AER, Eca, JPE, QJE and RES are often considered as "top journals" in Economics. They publish papers in all areas in Economics. JFE is a renowned but more specialized journal. JEL and JEP publish mainly "review" articles.

We have based our analysis on citations extracted from Thomson Reuters (2009b) (as in van Leeuwen and Moed 2005). For each journal, we have decided to select only documents of the "article" and "review" types (as given by Thomson Reuters 2009b) and published in 2006–2007 [2]. Table 1 gives, for each of the 8 journals, the number of citations received in 2008 by the "article" and "review" items published in 2006–2007. Dividing this number by the number of "article" and "review" items gives "our" 2008 Impact Factor for these journals. Table 2 and Figure 1 give, for each of the 8 journals, the number of items having received x citations.

We would like to emphasize three main points.

1. Figure 1 reveals that all distributions of citations are highly skewed, as was expected (Seglen 1992). Although we are dealing with similar journals in the same field, all distributions do not have the same shape: QJE and Eca have distributions that are much less skewed than the other ones in our sample, while the distribution for JEL seems to be very particular. Only taking the mean of these distribution of citations into account to compare journals may involve a huge loss of information. GIFs offer a simple mean to take more information into account.

2. The idea of stochastic dominance, i.e., using of GIFs with several functions u, leads to comparisons that are not as incomplete as could have been expected. As is easily checked, Table 2 contains several cases of first order stochastic dominance between journals: QJE, Eca, JEP, and

[2] The exclusion of "proceedings papers" was mainly motivated by the inclusion of AER in our sample. This journal publishes "proceedings issues" that are rather different from its normal issues

JFE dominate RES. More cases of second order of stochastic dominance occur. QJE dominates Eca, JEL, JEP, JFE, JPE, AER, and RES. Eca dominates JEL, JEP, JFE, JPE, AER, and RES. JEP, JFE, JPE, and AER dominate RES. These comparisons are valid for all rankings based on GIFs using an increasing and concave function u. It is interesting to notice that JEL does not first or second order stochastically dominate RES although the difference in their IFs is large and several journals with lower IFs than JEL second order stochastically dominate RES. Again, this is due to the very particular distribution of citations for JEL. Using the mean of this distribution to compare JEL with other journals may not lead to very robust comparisons.

3. Journals having close IFs can be compared differently when using GIFs with different functions u. For instance the IF of JEP is 3.420 while the IF of JFE is 3.394. These two journals are not related by first or second order stochastic dominance. Using GIFs together with a function u that is close to being affine will therefore lead to prefer JEP to JFE. Introducing more aversion to dispersed distributions of citations, may reverse this comparison. Consider the family of CARA (Constant Absolute Risk Aversion) utility functions (Arrow 1965, Pratt 1964) for which we have $u(x) = -e^{-\tau x}$ with $\tau > 0$. When $\tau = 0.03$, it can be checked that JFE becomes preferable to JEP.

Journals	# items published in 2006–2007	# of citations received in 2008	IF
QJE	83	367	4.422
Eca	100	386	3.860
JEL	37	135	3.649
JEP	88	301	3.420
JFE	175	594	3.394
JPE	69	227	3.290
AER	191	577	3.021
RES	90	226	2.511

Table 1: Data for some economic journals. Source: Thomson Reuters (2009a,b). Items in analysis are of the "article" and "review" types published in 2006 and 2007. Note: Extraction performed 1 December 2009.

	QJE		Eca		JEL		JEP		JFE		JPE		AER		RES	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
0	2	0.024	15	0.150	16	0.432	18	0.205	43	0.246	16	0.232	40	0.209	27	0.300
1	13	0.157	12	0.120	5	0.135	20	0.227	33	0.189	16	0.232	51	0.267	13	0.144
2	14	0.169	21	0.210	2	0.054	16	0.182	21	0.120	8	0.116	25	0.131	19	0.211
3	18	0.217	11	0.110	1	0.027	11	0.125	25	0.143	6	0.087	18	0.094	9	0.100
4	9	0.108	11	0.110	3	0.081	4	0.045	5	0.029	6	0.087	13	0.068	6	0.067
5	4	0.048	8	0.080	0	0.000	4	0.045	13	0.074	3	0.043	12	0.063	2	0.022
6	2	0.024	6	0.060	3	0.081	2	0.023	8	0.046	2	0.029	9	0.047	4	0.044
7	7	0.084	2	0.020	1	0.027	3	0.034	7	0.040	3	0.043	6	0.031	1	0.011
8	2	0.024	3	0.030	1	0.027	0	0.000	3	0.017	1	0.014	2	0.010	3	0.033
9	2	0.024	1	0.010	0	0.000	0	0.000	4	0.023	1	0.014	2	0.010	3	0.033
10	3	0.036	1	0.010	2	0.054	3	0.034	4	0.023	2	0.029	1	0.005	2	0.022
11	0	0.000	2	0.020	1	0.027	1	0.011	1	0.006	1	0.014	1	0.005	0	0.000
12	4	0.048	3	0.030	0	0.000	1	0.011	0	0.000	1	0.014	5	0.026	1	0.011
13	2	0.024	2	0.020	0	0.000	1	0.011	1	0.006	0	0.000	1	0.005	0	0.000
14	0	0.000	0	0.000	0	0.000	0	0.000	2	0.011	1	0.014	2	0.010	0	0.000
15	0	0.000	0	0.000	0	0.000	1	0.011	0	0.000	1	0.014	1	0.005	0	0.000
16	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	1	0.014	1	0.005	0	0.000
17	1	0.012	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000
18	0	0.000	0	0.000	0	0.000	1	0.011	2	0.011	0	0.000	0	0.000	0	0.000
19	0	0.000	2	0.020	0	0.000	0	0.000	1	0.006	0	0.000	0	0.000	0	0.000
20	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000
21	0	0.000	0	0.000	0	0.000	0	0.000	1	0.006	0	0.000	0	0.000	0	0.000
22	0	0.000	0	0.000	0	0.000	1	0.011	0	0.000	0	0.000	0	0.000	0	0.000
23	0	0.000	0	0.000	1	0.027	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000
24	0	0.000	0	0.000	1	0.027	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000
25	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000
26	0	0.000	0	0.000	0	0.000	1	0.011	0	0.000	0	0.000	0	0.000	0	0.000
27	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	1	0.005	0	0.000
28	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000
29	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000
30	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000
31	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000
32	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000
33	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000
34	0	0.000	0	0.000	0	0.000	0	0.000	1	0.006	0	0.000	0	0.000	0	0.000
35	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000
	83	1.000	100	1.000	37	1.000	88	1.000	175	1.000	69	1.000	191	1.000	90	1.000

Table 2: Distribution of citations. Source: Thomson Reuters (2009b). For each journal, the column "#" gives the number of items published in 2006–2007 having received the corresponding number of citations in 2008. The column "%" gives the corresponding proportion.

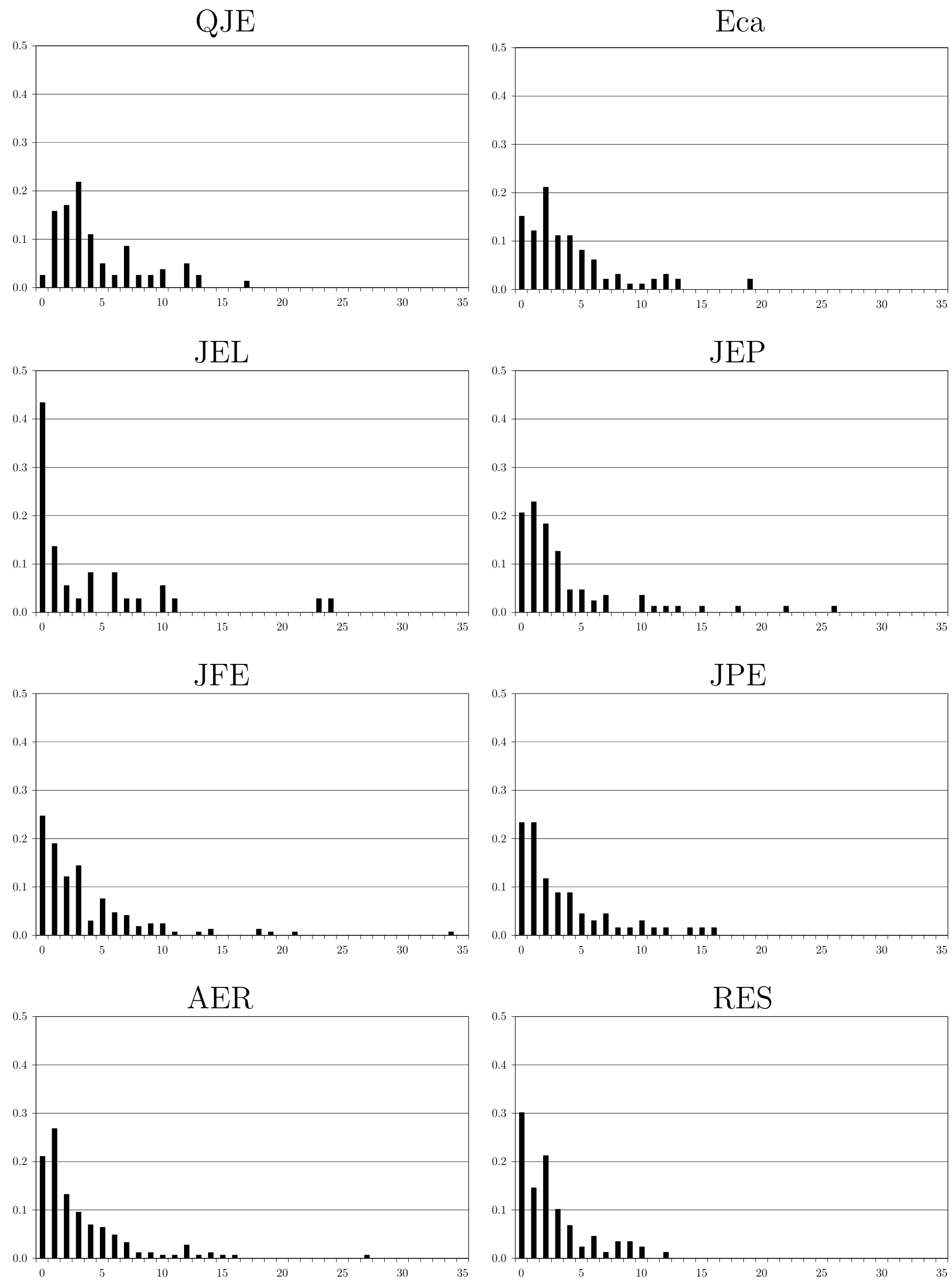

Figure 1: Distribution of citations for some economic journals. Source: Table 2. Note: The scale of both axes is identical for each journal: 0 to 35 (citations) on the horizontal axis, 0.00 to 0.50 (proportion of papers) on the vertical axis.