Journal of Informetrics 11 (2017) 689–703

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Measuring the citation impact of journals with generalized Lorenz curves

Tommaso Lando[a,b], Lucio Bertoli-Barsotti[a,*]

[a] Department of Management, Economics and Quantitative Methods, University of Bergamo, via dei Caniana 2, 24127, Bergamo, Italy
[b] Department of Finance, VŠB-TU Ostrava, Sokolskà Trida 33, Ostrava, Czech Republic

ARTICLE INFO

Article history:
Received 7 September 2016
Received in revised form 23 February 2017
Accepted 16 May 2017

Keywords:
Journal impact factor
Stochastic dominance
Quantile function
Lorenz curve

ABSTRACT

To improve comparisons of journals, which are typically based on single-value indicators, such as the journal impact factor (JIF), this paper proposes a *functional* approach. We discuss interpretatively three progressively finer dominance relations. The first one corresponds to a comparison between the *quantile functions* of the citation distributions. The second one consists in comparing the integrals of the quantile functions—namely, the *generalized Lorenz curves* (GLCs). The third one consists in comparing the integrals of the GLCs, where the integration is designed to emphasize the role of the "central body" of the articles of the journal. Although dominance relations are generally not complete orders, we demonstrate with an empirical analysis that it is possible to increase significantly the proportion of pairs of journals that are comparable by moving from the first to the second criterion, and then from the second to the third.

Because, in practical applications, it may be convenient to reduce such a functional comparison to a scalar comparison between indicators, we follow an axiomatic approach to identify classes of indicators that are *isotonic* with the criteria introduced. We demonstrate that the established JIF may be usefully improved if it is corrected simply by multiplying it by one minus the Gini coefficient. The resulting index, defined as *stabilized-JIF*, has many attractive features and it is isotonic with all the dominance relations introduced.

1. Introduction

Research evaluation is a topic whose theoretical importance is widely recognized by the scientific community, and it has several practical implications of great interest for the policy and orientation of scientific research. In this present context, the ranking of scientific journals is a major issue. Authors and institutions are interested in quantifying the "impact" of a journal on the scientific community, and the most widely used impact measures are based on citation data. On one hand, authors generally aim to identify the journals that may provide the largest audience and, hence, (possibly) the highest number of citations of their papers. On the other hand, researchers or research institutions may be directly rewarded for publishing in highly ranked journals.

The indicator most widely used to evaluate the impact of journals is based on the average number of citations per paper. This simple indicator, generally referred to as the *journal impact factor* (JIF), is ascribable to Garfield and Sher (1963) (see also Garfield, 1972; Garfield, 2006). The JIF, basically the mean citedness, is not uniquely defined insofar as it may vary according

* Corresponding author.
E-mail address: lucio.bertoli-barsotti@unibg.it (L. Bertoli-Barsotti).

http://dx.doi.org/10.1016/j.joi.2017.05.005
1751-1577/

to different citation databases (e.g., Web of Science, Scopus), different publication and citation time windows; the JIF may also depend on the degree of overlap between these timeframes (leading to so-called synchronous and diachronous impact factors; see Bar-Ilan, 2010; Ingwersen, Larsen, Rousseau, & Davis, 2001; Ingwersen, 2012) and different document types to be considered as citing or citable items for the citation count (e.g., all documents versus articles alone versus reviews and articles alone, etc.). For this reason, the JIF is also referred to hereafter using the plural "journal impact factors" (JIFs).

The advantages and disadvantages of JIFs have been discussed from different perspectives, and many variants or adjustments have been proposed in order to account for several limitations, such as the lack of statistical significance (Vanclay, 2012; Stern, 2013), insensitivity to field differences (Moed, 2010), insensitivity to the "weight" of the citing articles (Ferrer-Sapena, Sánchez-Pérez, González, Peset, & Aleixandre-Benavent, 2015) and manipulability by editorial strategies (Moustafa, 2015), among others. In this paper, we do not consider these more advanced issues; rather, we focus on two related problems:

- poor representativeness of the citation distribution: the JIF alone says little about the shape of the citation distribution—i.e., the mean is not suitable to represent highly skewed distributions (see, e.g., Seglen, 1997); and
- poor robustness: the JIF may be strongly influenced by one highly cited publication (e.g., Foo, 2013, shows how one single article increased the JIF of *Acta Cristallographica A* from 2 to 50 in 2009) or by a few (Editorial, 2005).

In particular, the JIF is often used (in a misleading way) to approximate the actual number of citations a paper might receive, although it is well documented in the literature that citation distributions generally show quite high concentration patterns, i.e., a few articles account for most of the citations, whilst most articles produce zero or only a few citations (see, e.g., Laband, 1986; Stern, 2013). This suggests that the issue of measuring citation impact should be related not only to the JIF but also to the measurement of inequality, or concentration, as we shall propose below.

Many authors agree that indicators of impact more representative and robust than the JIF should be employed. For instance, one may use the mode (Vanclay, 2012), the median (Wall, 2009) or another quantile of the distribution (Bornmann, Leydesdorff, & Mutz, 2013; Leydesdorff, 2012), although these indices have inferior discriminating power and may yield large numbers of ties. Alternatively, one can employ a trimmed mean (Seiler & Wohlrabe, 2014), a geometric mean corrected for uncited items (Zitt, 2012; Thelwall & Fairclough, 2015), or more generally (as we shall discuss in Section 3), a power mean of order a, where $0 < a < 1$. In all of these cases, the effect is that of "downsizing" the role of highly cited items and obtaining a more robust indicator of impact. In a more general framework, as discussed by Bouyssou and Marchant (2011), a family of generalized JIFs may be defined on the basis of the mean of an increasing function u of citations, where we may emphasize or downsize extreme values if we, respectively, choose u to be convex or concave.

The approach of Bouyssou and Marchant (2011) is, in turn, related to the concept of *stochastic dominance* for ranking citation distributions. In statistics, stochastic dominance relations establish preorders in the space of distribution functions that quantify the idea of one distribution being "preferable" to another (see, e.g., Marshall, Olkin, & Arnold, 2011). Such ranking criteria consist in a functional comparison that is generally much stronger (although not always verifiable) than a simple condition on a single-valued parameter (e.g., the mean) and leave very little room for ambiguity.

In a bibliometric context, the use of dominance rules for comparison of citation data has been proposed by Carayol and Lahatte (2009) and briefly discussed by Bouyssou and Marchant (2011) and Waltman and Van Eck (2009). In the paper, we develop and widen this approach, at both theoretical and applied levels. From a methodological point of view, it should be stressed that all of the aforementioned authors have used classic dominance relations—namely, first-order and second-order stochastic dominance, which are based on the distribution function and its integral, respectively. In contrast, we demonstrate that it is definitely more advantageous, in terms of ease of interpretation and computation, to express dominance relations by using the *quantile function* and, in particular, its integral—namely, the *generalized Lorenz curve* (*GLC*, Shorrocks, 1983). The GLC provides an attractive representation of the overall impact as well as the shape of the citation distribution. For a given journal with T publications, the GLC evaluated in $p \in [0, 1]$ determines an average of the quantiles, defined as the *partial JIF of order p*, i.e., the JIF corresponding to "the set of the 100p% less-cited papers", as will be shown in Section 2.3. Put another way, the GLC represents the "distribution" of the JIF within the papers of a journal.

It is well known that stochastic dominance relations are *preorders* (see Section 2) and, in particular, that they are not *total* (*complete*), because one may find pairs of distributions (journals) that cannot be ranked. In this case, it is possible to introduce some *finer* (or *weaker*, see Section 2) criteria that conform with our preferences and increase the number of comparable pairs of journals. In Sections 2.2 and 2.3, we analyse some *strong* preorders—namely, 1) *first order stochastic dominance (1-SD)*, which requires each quantile (i.e., generally, the citations of the Tp-th ranked paper) of the dominant journal to be higher compared to that of the dominated one; and 2) the *generalized Lorenz dominance (GLD)*, which basically requires that the condition for 1-SD holds *on average* or, in other words, that the "distribution of the JIF" (i.e., the GLC) of the dominant journal is uniformly higher. However, real data comparisons show that 1-SD and GLD are rarely verified.

Therefore, in Section 2.4, we introduce a new dominance relation for measuring the impact of journals that emphasizes the "body" of the citation distribution—namely, the *second-order outward generalized Lorenz dominance (2-OGLD)*. This is accomplished by cumulating the GLC, or "averaging" the values of the partial JIFs (as will become clear in the following explanation), from the "centre". This approach may serve a twofold objective: i) to rank the pairs of journals that are not ranked by 1-SD and GLD; and ii) to reward journals whose citations are mainly concentrated in the body, rather than in the tails, according to the principle that tails do not provide a good representation of the impact of a journal. Notably, the 2-OGLD does not require the dominant journal to have greater JIF, and it is especially suitable for ranking intersecting GLCs.