RESEARCH ARTICLE

WILEY

The unbearable lightness of scientometric indices

Matteo Migheli[1] | Giovanni B. Ramello[2] (iD)

[1]Università di Torino, Torino, Italy

[2]Università del Piemonte Orientale, Alessandria, Italy

Correspondence
Giovanni B. Ramello, Università del Piemonte Orientale, Alessandria, Italy.
Email: giovanni.ramello@uniupo.it

This paper advances the debate on scholarly publishing and the role of bibliometric indices in evaluating authors and their research, through a theoretical discussion and an empirical case study focused on economics and the impact factor. The rationale of the current bibliometric system is that reputation, assessed by citation figures, can be converted into an objective measure. We instead argue that it provides questionable results, because it fails to properly consider the meaning of indicators built for different purposes, as well as the psychological bias generated by the wrong interpretation of those indicators. However, the potential for abating these distortions exists.

1 | INTRODUCTION: REPUTATION IN SCIENCE

There is a maxim written in the first-century BC by Publilius Syrus, well known in the Latin world for his *sententiae*, asserting that: "A good reputation is more valuable than money." The dictum had a moral aim, as it was intended to encourage people to invest in their substance rather than material possessions. However, it also applies well to our current society, where quality is increasingly measured by external markers such as visibility or number of followers—what can roughly be defined as reputation—rather than objective traits.[1] Reputation is in fact becoming the core of the economy of the third millennium and is considered the main driver for many markets.[2] Even though a part of reputation is converted into money by the economic dimension, it remains central to most activities.

Science is no exception to this trend. Although multiple activities concur toward reputation-building in the scientific domain, one pivotal element is the "publish or perish" paradigm—in its turn heavily rooted in citations and author visibility.[3] This mechanism, whose origins can be dated back to the 18th century, has over time culminated in the central role of the peer-review system on the one hand and of bibliometric indices with the domination of the well-known impact factor on the other (Origgi, 2018). The current chapter of this story is a social construct—the modern research ecosystem—which on the one hand tries to foster progress in science for society, and on the other hand must provide individuals with private incentives for devoting time and energy to new research ventures (Clemens et al., 1995; Stephan, 2010).

It is worth noting that academia is an institution that naturally feeds on reputation. Though in the past symbolic rewards may have seemed more important than monetary proceeds, a scholar's prestige and notoriety are now instrumental to reaping economic benefits, albeit with some stratification of different systems. The contemporary scientific ecosystem has thus evolved in a way that makes it possible, on a regular basis, to convert effort and socially valuable results into tangible benefits—including money—for those who have produced it (Origgi & Ramello, 2015).

For specific reasons and path dependence, neither the traditional price system nor other allocative technologies were adopted in science, whose modern organization is instead an idiosyncratic institutional arrangement revolving around reputation as a means for fostering scientific advances while providing economic proceeds.[4] What has emerged is a complex institutional arrangement reliant upon bibliometrics—that is, numbers—which is today the core of scholars' and universities' reputational machinery. In the past decade, this has become a leitmotif of the research systems in many countries, to the point that terms familiar to industrial activities—such as *productivity* and *efficiency*—have become the focus of governments and agencies set up specifically to assess and foster scientific advances and technological change. All these systems rely primarily upon rankings, which, however, provide a specific overview of the reality that is somewhat distorted by their own construct and so may impact on the overall perspective given.

Against this backdrop, the aim of this paper is to show, through a theoretical discussion and an empirical case study focusing on economics, how the use of indices and rankings at the journal level may fail to capture the huge variance in quality across the different articles. This can in turn randomly reward some bad papers (i.e., with zero citations) and punish good ones (i.e., with many citations). We will argue that the current bibliometric system's failure to properly consider the meaning of indicators built for different purposes, as well as the psychological bias and the indirectly-induced psychological effect of numbers that anyone can easily order and then use as a metric, has produced a system that provides questionable results. However, the potential for lowering distortion exists.

The paper is organized as follows: Section 2 gives an overview of the current dynamics in scholarly publishing, outlining how this system developed and what kinds of idiosyncratic features it presents today. Section 3 further elaborates on the theoretical framework by disentangling the role of citations in asserting reputation, and how this works when applied to scholarly journals. Section 4 tests the ideas presented in the theoretical part through an empirical analysis, conducted on the most historical bibliometric dataset and its relative indices. Section 5 presents and discusses the findings, and Section 6 provides some further comments and extends the reasoning to the system. Finally Section 7 draws the conclusions.

2 | THE RISE OF BIBLIOMETRICS AS MEASURE OF REPUTATION

Whether the evaluation system that dominates science is the best possible model or not is outside the scope of the current paper. Whatever its merits, this system is today the pivotal element on which the scientific community's life relies. Indeed, the current scientific ecosystem is a stratification of multiple reputational logics, which taken together have in recent times led to rankings similar to those found in sports. Countries today pay substantial attention to their positioning in international research rankings, and research evaluations give great weight to scholarly publications. The dynamics of this system, and the structural consequences it has brought about, have been extensively discussed elsewhere (Egert & Scheufen, 2018; Origgi, 2018; Origgi & Ramello, 2015). Among other things, it has determined the prevalence of the "article" publication format, with an invasion of millions of articles per year whose exponential growth means that, instead of the scarcity familiar to many economic problems, scholars face a problem of "over choice," as abundance of information creates scarcity of attention (Migheli & Ramello, 2018; Origgi, 2018).

An abundant literature, including Hamermesh (2018) and Heckman and Mokten (2020), to cite just two recent contributions focusing on economics, shows that universities tend to base hiring and promotions more on the journals where candidates' articles are published, rather than on the performance of the articles themselves or the citations they receive.

The use of bibliometric indices—including the best-known one, the impact factor (IF)—as a measure of quality for journals and the articles published in them has been widely criticized (e.g., Larivière & Sugimoto, 2019; Pendlebury & Adams, 2012; Seglen, 1997), leading to a sort of manifesto, the *Declaration on Research Assessment* (https://sfdora.org/). These critiques contend that the quality of publications and research cannot be assessed based solely on bibliometric performance and on IF in particular. Yet such measures continue to be highly relevant, both for authors deciding where to submit their work and for employers and research evaluators in hiring, promotions, and assessment of scholars and research centers. For this reason, these indices cannot be disregarded, and there are worth analyzing to help better understand what they really measure and the weight they should be given in research evaluation. Naturally, focusing on these metrics does not mean defending their pre-eminence; rather, deepening our understanding of them will help the scientific community to improve the quality of evaluations and, hence, of scientific research.[5]

The current state of the art, despite the dissenting voices mentioned previously, is that bibliometric indices are quite pervasively used in academic life and research evaluation and today serve a twofold purpose. First, they provide semiotic markers of usefulness or relevance that help direct researchers' attention toward specific articles and journals. Second, they facilitate ranking and classification by virtue of their numeric nature—based simply on the number of citations received by articles and journals in a given timespan (Migheli & Ramello, 2018). These dynamics have not only profoundly affected the ethos of science, and the way in which scientific knowledge is produced and transmitted. They have also shifted a large part of scientific activity toward an industrial-production model in which the key variables for measuring output are quality and quantity. Whereas quantity is easy to measure once we define the unit of the output (in this case the article), quality instead calls for specific metrics—and bibliometrics was there to offer an arms-length solution (Biagioli & Lippman, 2020; Origgi, 2018). On the whole, if building a solid reputation is today the major tenet of a scientific career, the publishing system and citations provide the tool arbitrarily chosen for "quantifying" reputation (Woolston, 2015). What is the meaning of the numeric metrics thus obtained is of course the key question.

In short, citations have become the unit of the metric through which reputation is measured, and their numerical magnitude stands for the overall visibility of a scholar. Following Merton (1988, p. 620), citations can thus be likened to a coinage system that makes it possible to capitalize "pellets of peer recognition that aggregate into reputation wealth" for individual scholars. This model is also very convenient for acting in the real world: The reputational capital that a scientist acquires through papers and citations can be converted into rewards such as career advancement, recruitment by better institutions, access to research grants, consulting, and similar activities (Heckman & Mokten, 2020; Migheli & Zotti, 2020; Origgi & Ramello, 2015).

However, data on an individual's citations are difficult to collect, because it tends to be unstable, changes over time, and also depends on the source considered. Counting citations in the pre-Internet era was a task not easily accomplished, and even today is non-trivial and potentially subject to fluctuations, when the focus is on a single

researcher.[6] Accordingly, another simple yet idiosyncratic rule was introduced: namely, the custom of using a static snapshot of a journal's reputational standing as a conversion benchmark that entitles the scholar to the same reputational cachet. The assumption behind this habit is simple: If a journal is esteemed by the community—and "esteemed" today means "has a high bibliometric index"—then the new articles landing in this journal via peer review must be of the same quality. In consequence, bibliometric indices—which essentially provide a weighted citation count of a journal—became the practical tool for tackling the problem of individual reputation. They are much easier to calculate, they are stable, and they operate at the journal level rather than at the author level.

So today, the easy way out for settling an author's reputation is to take the bibliometric index of the journal where an article was published and pass it on to its authors, according a sort of syllogism: If the authors can afford a journal of (bibliometric) quality X, then the paper must be of the same quality X, and thus, authors too are of quality X. Many national research evaluation agencies have largely endorsed this logic, classifying journals into categories based on bibliometric indices, and distributing benefits to researchers and institutions accordingly (Karpik, 2011; Origgi & Ramello, 2015; Osterloh & Frey, 2020).

The core of this system is well understood among practitioners: Bibliometrics is governed by a handful of indices, among which the long-time king is *impact factor* (IF), calculated as the total citations in a given year of the articles published in the previous 2 years, divided by the number of those articles. How this can give a strong measure of scientific quality is quite obscure, especially since its inventor, the linguist Eugene Garfield, was actually trying to create a useful instrument for studying the propagation of scientific thinking, rather than a measure of quality (Garfield, 1955). Notwithstanding Garfield's original intentions, impact factor and the ensuing bibliometric indices have become important for establishing the value of scholarly publications and, indirectly, the reputation of researchers.[7]

The anomalies—or at least the idiosyncratic traits—of the overall system are many, and certain oddities emerge. For example, all these indices are by construction backward-looking, meaning they are built on what happened in the past, whereas by definition a new article is forward looking, since it is part of the future and determines the future indices. In other terms, the rate of return of authors at time t is determined by what other authors did at time $t - 1$ (or even before that), and the authors at time t will in their turn determine return for authors at $t + 1$ and the subsequent generations.

If we consider the analogy with financial markets, the journal IF represents the remuneration of the authors. However, it is worth noting that in scholarly publishing, the remuneration of the current investor, the author, is based on what other investors have done in the past. Only a perfect selection process would align quality. No comparable situation exists in markets, and given that this system is central to the remuneration of scholars, its ability to fairly repay investments must depend on there being a steady correlation, in terms of quality, among papers published in a given journal. Unfortunately, as will be shown later in the paper, this appears questionable not only across different issues of a journal, but even within the same issue of a journal.

Another anomaly is the wide scope for endogenous manipulation. Biagioli (2016) and Biagioli and Lippman (2020) highlight how evaluations based on bibliometric indicators, and IFs in particular, have continuously engendered manipulation of quality on the part of both journals and authors. According to these researchers, authors and editors, sometimes even specific associations, engage in a "game" of trying to boost the citations of papers published in the journals, which they edit, or where their articles have appeared. There are many different ways to accomplish this, from citing those journals in their own articles, to suggesting references to papers published in particular journals when refereeing somebody else's paper. To continue the analogy with financial markets, this would be tantamount to allowing extensive use of insider trading to boost the yield of financial assets. But as we know, financial authorities strictly monitor markets to avoid this.

3 | JOURNALS, BIBLIOMETRIC INDICES, AND THE HALO EFFECT

In light of the preceding discussion, a deeper understanding is required of the theoretical framework governing the use of citations for assessing quality and values.

In the patent domain, there is a long-lasting tradition of using citations to elicit the perceived importance of an invention, based on the notion that the scientific relevance of an existing patent is disclosed by the number of citations it receives in subsequent patent applications. The underlying assumption is that citations in new patents of the prior art are a proxy for the importance or value of the previous scientific contribution (Ippoliti et al., 2021; Moser et al., 2015). More precisely, Trajtenberg (1990) and Griliches (1990) show citation counts to be positively correlated with the estimated social surplus, and subsequent work has shown that citations are also correlated with changes in the stock market value of U.S. firms (Hall et al., 2005) and with inventors' valuation of patents (Harhoff et al., 1999). The basic idea is that citations serve as a measure in two ways. First, they measure the impact of the cited knowledge on follow-on ideas (as an input), and this is a measure of a potentially important economic externality. Moreover, citations refer to the state of the art, so that the number of citations might be used as a proxy for economic value (Griliches, 1990).

An approach along similar lines, though without empirical verification, has been adopted in science, with a number of works asserting that the number of citations received by a given scientific article is proportional to the research quality and hence to its value (Davis, 2009). This reasoning which we outlined previously somehow harks back to the Mertonian model of the citation (Merton, 1988) as a unit of remuneration, with the assumption that the token of attention represented by the citation can thus be converted into reputation by the researcher and thence into money. Even assuming this relation to be true, there are further problems that arise. As we have pointed out,

counting citations at the article level or author level is impractical and problematic, and so the difficulty is commonly overcome by using the readily available bibliometric indices instead. However, the issue with bibliometric indices is that they aggregate citations and thereby provide an average measure that does not reflect the individual importance of single articles.

To clarify this point, consider two journals, both with a 2-year IF equal to 1, and assume that each journal published 40 articles during the 2 years preceding the assessment. The information conveyed by the IF is that, for each journal, the 40 articles received a total of 40 citations in the reference period. However, suppose that the first journal had one article that was cited 40 times, while in the second journal, each article was cited once. In both journals, the number of citations and therefore the IF is the same, but in the first journal, 39 articles did not attract any attention at all, which was instead all concentrated on a single article. In the second journal, all the items published attracted some attention. In terms of production results, this means that the first journal was able to "produce" one champion and 39 underdogs, while the second journal produced 40 articles that attracted some moderate attention. Nevertheless, in the first journal, the 39 underdogs still share part of the attention obtained by the champion, as they enjoy the IF of the journal.

If we assimilate scientific papers to general consumption goods, and journals to firms producing such goods, then the first journal invested resources to produce a bestseller and 39 unsold products, a situation that no firm would consider sustainable under standard circumstances. In addition, the perceived quality of the journal may seem the same if scholars and evaluators limit their attention to the IF, whereas in reality, publishing in the first journal yields a much lower probability of receiving at least one citation than publishing in the second journal. So, the central issue is not just the IF figure itself, but also the distribution of figures determining the IF, which in turn indicates how much of a journal's IF depends on the publication of champions.

A skewed distribution in citations would imply that articles within the same journal are different in terms of quality. In this case, the IF metric will not rightly capture the relevance of each individual contribution, as it might in fact either over-value or under-value a single paper depending on whether its number of citations is below or above the average value for the journal. This means there is a sort of default value, a standard "yield" for papers published in a given journal, that has an important impact in defining a researcher's value in the scientific community, yet is only scantly correlated to the true value of the contribution. Since scientific products are in general ranked according to the journal IF, for purposes ranging from evaluations of the quality of research institutions to hiring and promoting scholars in universities, this implies that there may be systematic distortions (Migheli & Zotti, 2020; Osterloh & Frey, 2020).

The described dynamics are not strange if we consider the psychology literature and the dynamics of evaluation that occur in many situations, including markets. There is a well-known and ubiquitous phenomenon called the "halo effect," which links the current evaluation to previous judgments.[8] When the halo effect operates in markets, for example, the good reputation of a company becomes a driver for consumer choice, even in terms of price sensitivity (Burke et al., 2018). In marketing, it has been extensively exploited, to the point that several commercial practices and strategies rely upon it.

Brands, in particular, are a way for capturing reputation and thus encapsulating the halo effect into signs that can be recognized by consumers. The power is such that this behavioral inertia can be transferred across products and markets. This is the basis for all strategies of brand extension—whereby reputational inertia is transferred to similar products—or brand stretching—whereby reputational inertia is transferred to very distinct products (Pepall & Richards, 2002). Such practices are essentially connected with the increasing returns to scope in the use of a brand. Once a signal conveying reputation has been created, the firm can in fact use it to convey an equivalent halo about other, distinct products, and this in fact proves helpful for managing multi-product production (Ramello, 2006).

However, from a psychological point of view, the halo effect is a cognitive bias—that is, a systematic error in thinking—since it causes evaluators to be influenced by their prior beliefs or judgments of a previous performance and thus distorts behaviors. This is also because, according to social psychology, the human mind tends to attribute the general characteristics of a class to each of its individual members (Origgi, 2018). Daniel Kahneman (2010) has further explained that this interpretation of the world is due to our inbuilt preference for causal explanations. Even when it might happen that an observation is randomly included in a class, we still tend to believe that there is a causal relation.

The halo effect provides an economical heuristic for inferring information and taking decisions under different constraints. However, when applied to something like bibliometrics that aspires to scientific legitimacy, it is epistemologically wrong because it does not consider important second-order information, such as why an item is included in a list and what are the mechanisms for inclusion. It also gives rise to a self-perpetuating rigidity and inertia in the signal representing the prestige of journals (see Table 2), which can only be broken by some external intervention, random shock, or purpose-designed policy that has the effect of altering individuals' perceptions and hence their decisions (Migheli & Ramello, 2013). Yet no such external intervention is currently happening, despite the many critical voices.

4 | EMPIRICAL ANALYSIS: DATA AND METHODOLOGY

The analysis presented here uses data from the Clarivate Journal Citations Report (JCR henceforth) released in mid-2020, focusing on 371 economics journals, ranked with respect to the citations received in 2019. The reasons for focusing on economics and the JCR are manifold. JCR is a milestone of bibliometrics: It was started by Eugene Garfield, the founding father of bibliometrics and inventor of the impact factor, and boasts a longer tradition than the many other indices available today. Another reason is tied to Garfield's insight that each discipline represents a distinct citations ecosystem, which sometimes even needs to be further disentangled into subdisciplines to

provide accurate measures (Garfield, 2006; Hamermesh, 2018). In this vein, our empirical example focuses on a single "citations ecosystem"—economics—which has already been extensively studied as such.[9] Finally, as Rousseau and Rousseau (2021) point out, the IF is (rightly or wrongly) extremely relevant and widely used in economics—the field on which the present paper focuses. This again confirms that bibliometric indices are a good match with this discipline.[10]

Thus, we examine a number of bibliometric indicators provided by the JCR report, which represent some of the metrics commonly employed today in the research evaluation:

- The 2- and 5-year impact factor (i.e., the standard IF calculated on citations for articles published in the previous 2 years and the one calculated on the longer timespan of 5 years)
- The AIS (Article Influence Score), a composite index based on an algorithm that, among other things, attempts to locate citations in highly cited journals[11]
- The total number of articles published in the last 2 years
- The total number of citations obtained in the same period
- The number of citations obtained by the five most-cited articles of each journal in the last 2 years

For each journal, we computed the weight of the citations obtained by the most-cited articles on total citations and thus their contribution to the 2-year IF. This is not a measure of the variance within each journal, because Clarivate does not disclose the number of citations for all the articles published. However, it does provide the citations count for the seven articles that received the most citations during the reference period for calculating the 2-year IF. This information is enough to understand whether champions are present—that is, if few articles account for a large share of the total citations received—and enable us to compute the 2-year IF that the journal would have obtained if it had not published the most-cited articles. The IF adjusted in this way is a measure of the attention attracted by the other papers in the journal.

On this basis, we determined the contribution of the most-cited article, the three most-cited articles, and the five most-cited articles to the 2-year IF, both in absolute terms (i.e., citations/articles in the reference period) and as the share of 2-year IF ascribable to the most-cited articles.

The IF and AIS are in some respects useful indicators of the average attention obtained by the articles published in a journal in a given timespan, but the distribution of citations across different articles, even published in the same issue, can be very skewed. Different items in the same journal may attract widely varying attention (measured as number of citations) from the scientific community. Such variability is not captured by any of the bibliometric indices traditionally provided, such as those examined here, yet the quality of a journal assessed through these indices spills over to all the papers published in it. This essentially means that the cognitive bias implied by the halo effect may be in place. The publication of champions in journals may bias the perceived quality of the outlet itself, distorting several evaluation processes based on comprehensive bibliometric indicators. To evaluate this, our empirical strategy is to present rankings of scientific journals in economics based on the 2-year IF adjusted by subtracting the most cited articles.

Table 1 shows the descriptive statistics of the Clarivate data used in the empirical investigation. The figures suggest that large variability

TABLE 1 Descriptive statistics (standard deviations in parentheses)

	Average	Maximum	Minimum	Journal(s) with highest value of the indicator	Journal(s) with lowest value of the indicator
Impact factor	1.805 (1.485)	11.375	0.143	Quarterly Journal of Economics	Review of Network Economics
Citations	252.53 (528.17)	7.008	4	Energy Policy	Review of Network Economics, Korean Economic Review
Number of articles published in the last 2 years	113.40 (130.50)	1390	15	Energy Policy	NBER Macroeconomics Annual
Article influence score	1.257 (2.092)	22.091	0.019	Quarterly Journal of Economics	Custos e Agronegocio Online

TABLE 2 Correlations between different indicators

	Articles published in the last 2 years	Citations received in the last 2 years	2-year impact factor	5-year impact factor	Article influence score
Articles published in the last 2 years	1				
Citations received in the last 2 years	0.854***	1			
2-year impact factor	0.245***	0.493***	1		
5-year impact factor	0.237***	0.478***	0.962***	1	
Article influence score	0.035	0.225***	0.747***	0.824***	1

*0.05 < p value ≤ 0.1. **0.01 < p value ≤ 0.05. ***p value ≤ 0.01.

exists among journals, in terms of both citations received and number of articles published. This last piece of information is particularly relevant: Journals clearly have different strategies concerning the number of articles to publish annually. Obviously, such a number may depend on the aims and scope of each journal, with generalist outlets likely eliciting more submissions than subject-specific journals. The number of papers published yearly is important in determining how a journal attracts the attention of scholars for at least two reasons. On the one hand, the more articles are published, the larger the number of scholars potentially interested in at least one of them. On the other hand, if bibliometric indices measure academic attention, a large number of published articles may dilute such measures, especially when—as in the case of IF—the average number of citations per article is considered. Table 2 shows the correlation matrix between the variables listed in Table 1.

While there is a positive and statistically significant correlation between the number of articles published during the 2 years that precede the computation of the 2-year IF and the other variables, it is far from perfect. This suggests that publishing many papers does not necessarily entail receiving more attention from the academic community. As mentioned before, there might be champion articles on which much attention is focused, while other published papers attract no interest at all.

More specifically, these correlations show that the number of citations received by a journal increases—unsurprisingly—with the number of articles it publishes. Yet no such correlation exists between AIS and number of published articles, suggesting that whether a journal's impact is above or below the mean does not depend on how many articles it has published. On the other hand, there is a positive, high and statistically significant correlation between 5-year IF and AIS: Journals that receive more citations over a 5-year period are also more likely to have an above-the-mean impact. Finally, the AIS of an

outlet depends positively on the number of citations (i.e., 5-year IF) of the journals that cite papers published in that outlet. Thus, the table suggests that a high number of citations does not correlate with the quality of the citing journals, when this quality is measured in terms of 5-year IF.

As previously written, the choice of the bibliometric index used in this paper is arbitrary, and there exist other comparable indicators (such as SJR or Cite Score) that are used in rating journals. However, hierarchical and principal component analyses show that these three indicators all provide information that is very similar to each other (Bollen et al., 2009), so we can expect the results presented in the next section to be generalizable.

5 | RESULTS AND DISCUSSION

The first step of the analysis is to present the contribution of the one, three, and five most-cited articles published in a journal to the total numbers of citations received by the same outlet during the 2 years that precede the calculation of the 2-year IF.

Table 3 presents that information in both absolute and relative terms. In the first case, the average number of citations obtained by the one, three, and five most-cited articles is shown. In the second case, the same citation numbers are presented as shares of total citations received in the reference period. The variability is again large, with articles that obtained between 1 and 241 citations, and cases of journals (the *Review of Network Economics* and the *Korean Economic Review*) that concentrated all the citations in only four articles, despite having published many more (28 and 23 respectively for the two mentioned journals). In addition, four journals (the *Hitotsubashi Journal of Economics*, the *Review of Network Economics*, the *Korean Economic Revue*, and *Estudios de Economía*), which published, respectively,

TABLE 3 Contribution of the most cited articles to total cites (standard deviations in parentheses)

	Average	Maximum	Minimum	Journal(s) with highest value of the indicator	Journal(s) with lowest value of the indicator
Absolute values (citations)					
Most cited	15.52 (20.65)	241	1	*Journal of Economic Perspectives*	*Rev. of Network Econ., Korean Econ. Rev., Econ. J. Watch, FinanzArchiv, Rev. Hist. Indust.*
Three most cited	33.24 (36.42)	368	3	*Journal of Economic Perspectives*	*Rev. of Network Econ., Korean Econ. Rev., Econ. J. Watch, FinanzArchiv, Rev. Hist. Indust.*
Five most cited	46.12 (47.37)	412	4	*Journal of Economic Perspectives*	*Rev. of Network Econ., Korean Econ. Rev.*
Percentage values					
Most cited	10.79 (7.58)	45.45	0.73	*Hitotsubashi Journal of Economics*	*Energy Policy*
Three most cited	23.84 (13.87)	81.82	1.92	*Hitotsubashi Journal of Economics*	*Energy Policy*
Five most cited	33.07 (17.52)	100.00	2.48	*Hitotsubashi J. of Econ., Rev. of Network Econ., Korean Econ. Rev., Est. de Econ.*	*Energy Policy*

18, 28, 23, and 22 papers in the reference period, gathered all their citations with just five articles, leaving about three quarters of the published articles uncited. The most cited paper of the *Hitotsubashi Journal of Economics* accounts for almost half of the total citations received. *Energy Policy*, instead, shows the most uniform distribution of citations in the sample, with the top five articles accounting for only 2.48% of the received citations.

The impact of the most-cited papers on the 2-year IF therefore differs among journals, with some that would have obtained the same result with only four or slightly more papers, and others with many more papers that attracted the attention of the scientific community. The impact of these differences between journals on their 2-year IF is the object of the following part of the analysis.

Journals were first ranked according to the official IF bibliometric index. We then computed how their absolute positions in the ranking would change with the adjusted IFs (i.e., excluding the one, three and five most-cited articles). Table 4 presents the journals whose position improved or worsened the most, along with the positions gained or lost in the ranking.

The changes reported in the table are large. Excluding the single most-cited article has a notable impact: It lowers the ranking of the *Journal of Financial Econometrics* by 75 positions, from 157 to 232, while it improves that of the *Journal of Economic Behavior & Organization* by 20 positions, from 159 to 139. Obviously, the impact of excluding the three and five most-cited papers is even greater, as the second and third rows of the table show. This result suggests that, when journals are ranked according to the 2-year IF, their position is largely affected by the presence of champion articles. The presence of such abundantly cited papers allows the other papers, which receive much fewer citations, to anyhow enjoy a high IF, thereby engendering a subsidy in terms of figures and a halo effect in terms of reputation. The distortive effects on perception and evaluation are quite evident. Papers that do not attract any attention may appear to be highly rated, and vice versa. For example, an article with substantial citations but published in a journal whose IF is 0.5 will appear to be of worse quality than a never-cited article published in a journal whose IF is 1.0. The halo effect in this case causes papers of low interest—and by extension their authors—to benefit from the proximity of champions in the same journal.

Following a similar approach, Table 5 presents the top 30 journals ranked according to the 2-year IF provided by Clarivate and then according to the recalculated IFs after excluding the one, three, and five most-cited papers. The data shown are limited to the first 30 journals for the sake of brevity and clarity; however, similar results hold for the journals not listed here.[12]

As in the case of Table 4, negative changes in Table 5 represent improvements in the position held by a journal, while positive numbers stand for the opposite. We can see that the changes in the first 30 positions are limited, though some are not negligible. *The Quarterly Journal of Economics* holds the top position, irrespective of the correction applied to its 2-year IF, suggesting that the halo effect enjoyed by the articles published in it is limited. The same result seems to hold for the second and third journals in the ranking, although the application of the proposed corrections reverses the positions of these two journals. The largest variation present in Table 5 is that of the *Brookings Papers on Economic Activity*, which loses 17 positions when the five most-cited articles are excluded from the computation of the 2-year IF. Such a result means that the champion articles published in it relevantly pull the 2-year IF of this outlet.

It is worth noting that this section examines the *rankings* of journals according to IF, rather than only discussing the IF itself. The reason for choosing this approach is that IFs by themselves are not very informative if they are not compared to those of other journals in the same category. Indeed, maximum and mean IFs differ across fields and between years, so that an IF of 9 may be very high in economics (where the maximum value in 2019 was 11.375), but relatively low, for example, in oncology (where the maximum IF in 2019 was 292.278). Thus, the relative position of a journal in a particular field is more informative about its quality (conditional on the limitations represented by the measure used), as Bradshaw and Brook (2016) highlight.

The correlations reported in Table 2 might suggest that the IFs adjusted to exclude the most-cited papers have no different meaning than the official IF. In other words, the halo effect evinced for some journals by the recalculations of Tables 4 and 5 may continue to exist—though reduced in magnitude—even when the most-cited papers are excluded from the IF computation. To understand whether this is the case, a principal component analysis (PCA) was performed on the official and recalculated measures. If they represent the same

TABLE 4 Variations in ranking when the most cited articles are excluded from IF (standard deviations in parentheses)

	Maximum gain	Maximum loss	Journal(s) with maximum gain	Original rank/ new rank	Journal(s) with maximum loss	Original rank/ new rank
Most cited excluded	−20	75	*Journal of Economic Behavior and Organization*	159/139	*Journal of Financial Econometrics*	157/232
Three most cited excluded	−39	133	*Journal of Economic Behavior and Organization*	159/120	*Econometrics Journal*	157/290
Five most cited excluded	−49	151	*Journal of Economic Behavior and Organization*	159/110	*Econometrics Journal*	102/253

Note: Negative variations indicate gains in the ranking, as lower numbers represent higher ranks. Analogously, positive variations indicate losses in the ranking.

TABLE 5 Top-30 journals using different indicators and their variation in the ranking position with respect to the ranking based on 2-year IF as calculated by Clarivate analytics

Rank	Usual impact factor	Impact factor excluding the most cited article	Variation	Impact factor excluding the three most cited articles	Variation	Impact factor excluding the five most cited articles	Variation
1	Quarterly Journal of Economics	Quarterly Journal of Economics	-	Quarterly Journal of Economics	-	Quarterly Journal of Economics	-
2	Journal of Economic Perspectives	Economic Geography	−1	Economic Geography	−1	Economic Geography	−1
3	Economic Geography	Journal of Economic Perspectives	1	Journal of Economic Perspectives	1	Journal of Economic Perspectives	1
4	Brookings Papers on Economic Activity	Journal of Finance	−1	Journal of Finance	−1	Journal of Finance	−1
5	Journal of Finance	Review of Environmental Economics and Policy	−2	Journal of Financial Economics	−3	American Economic Review	−4
6	Journal of Economic Literature	Journal of Financial Economics	−2	American Economic Review	−3	Journal of Financial Economics	−2
7	Review of Environmental Economics and Policy	American Economic Review	−2	Journal of Political Economy	−3	Energy Economics	−4
8	Journal of Financial Economics	Journal of Economic Literature	2	Energy Economics	−3	Energy Policy	−5
9	American Economic Review	Journal of Political Economy	−1	Energy Policy	−4	Journal of Political Economy	−1
10	Journal of Political Economy	Energy Economics	−1	Review of Environmental Economics and Policy	3	Transportation Research Part B-Methodological	−8
11	Energy Economics	Energy Policy	−2	Journal of Economic Literature	5	Transportation Research Part E-Logistics and Transportation Review	−9
12	Journal of the Association of Environmental and Resource Economists	Brookings Papers on Economic Activity	8	Transportation Research Part B-Methodological	−6	Value in Health	−7
13	Energy Policy	Transportation Research Part B-Methodological	−5	Value in Health	−6	Journal of Economic Literature	7
14	American Economic Journal-Applied Economics	Value in Health	−5	Transportation Research Part E-Logistics and Transportation Review	−6	Ecological Economics	−12
15	Journal of Policy Analysis and Management	Transportation Research Part E-Logistics and Transportation Review	−5	Small Business Economics	−2	Small Business Economics	−2
16	Review of Economic Studies	Small Business Economics	−1	Ecological Economics	−10	Review of Economic Studies	-
17	Small Business Economics	Review of Economic Studies	1	Review of Economic Studies	1	American Economic Journal-Applied Economics	3
18	Transportation Research Part B-Methodological	American Economic Journal-Applied Economics	4	Brookings Papers on Economic Activity	14	Review of Environmental Economics and Policy	11

TABLE 5 (Continued)

Rank	Usual impact factor	Impact factor excluding the most cited article	Variation	Impact factor excluding the three most cited articles	Variation	Impact factor excluding the five most cited articles	Variation
19	Value in Health	Review of Financial Studies	−2	American Economic Journal-Applied Economics	5	Food Policy	−9
20	Transportation Research Part E-Logistics and Transportation Review	Ecological Economics	−6	Review of Financial Studies	−1	Review of Financial Studies	−1
21	Review of Financial Studies	Journal of Policy Analysis and Management	6	Food Policy	−7	Brookings Papers on Economic Activity	17
22	Journal of Economic Growth	Review of Economics and Statistics	−5	Socio-Economic Planning Sciences	−7	Socio-Economic Planning Sciences	−7
23	NBER Macroeconomics Annual	Food Policy	−5	Review of Economics and Statistics	−4	Transportation Research Part A-Policy and Practice	−7
24	Economic Policy	Socio-Economic Planning Sciences	−5	Transportation Research Part A-Policy and Practice	−6	World Development	−9
25	Cambridge Journal of Regions, Economy and Society	Transportation Research Part A-Policy and Practice	−5	World Development	−8	Review of Economics and Statistics	−2
26	Ecological Economics	American Economic Journal-Economic Policy	−6	American Economic Journal-Economic Policy	−6	Journal of Transport Geography	−8
27	Review of Economics and Statistics	World Development	−6	Journal of Transport Geography	−7	American Economic Journal-Economic Policy	−5
28	Food Policy	Econometrica	−2	Econometrica	−2	Econometrica	−2
29	Socio-Economic Planning Sciences	NBER Macroeconomics Annual	6	Pharmacoeconomics	−11	Pharmacoeconomics	−11
30	Econometrica	Economic Policy	6	NBER Macroeconomics Annual	7	NBER Macroeconomics Annual	7

TABLE 6 Correlations between the principal components and the variables used in the analysis

	Share of the most cited article on total IF	Share of the three most cited articles on total IF	Share of the five most cited articles on total IF	Article influence score	2-year impact factor
First component	0.91***	0.97***	0.97***	−0.34***	−0.45***
Second component	0.32***	0.22***	0.17***	0.87***	0.82***

*0.05 < p value ≤ 0.1. **0.01 < p value ≤ 0.05. ***p value ≤ 0.01.

phenomenon, only one component will have an eigenvalue larger than 1; otherwise, more than one component will pass the threshold.

The PCA results yielded two components with eigenvalues larger than one; these two factors are therefore retained and correlated with the variables used in the analysis. Table 6 presents these results.

The figures in the table reveal that the two components represent two different dimensions of the data set. The first component correlates positively with the IFs recalculated after excluding the most cited papers, while their association with the official IF and the AIS is negative. The second component correlates positively with all five variables, but the coefficients of correlation with the official bibliometric indices are much higher than those with the adjusted measures. In other words, the first component represents the adjusted measures, while the second component represents the unadjusted one, suggesting that the measures falling into these two groups represent different, though correlated, phenomena.

This piece of evidence highlights that the AIS and the 2-year IF are not the same as the share of IF represented by the most-cited articles. In the case of the AIS, which depends not only on the number of citations, but also on the quality of the citing journal, the result suggests that highly cited articles are referenced more frequently by papers published in non-highly cited journals, compared with articles that received fewer citations in the 2 years after publication. This phenomenon may depend on the fact that writing and publishing papers in top-ranked journals requires more time than writing and publishing papers in less prestigious outlets. In other words, during the 2 years following their publication, top-cited articles influence mainly the research published in non-top journals. The correlations between the second component on the one side and 2-year IF and AIS on the other suggests that these two measures capture the same piece of information: The most cited journals are, on average, referenced by top-quality outlets. Taken together with the previous result, this suggests that journals without champions, that is, those whose articles are on average uniformly well-cited, contribute to research published in top journals more than outlets that publish champions and underdogs. Given that the IF provides scholars with a comprehensive measure of citations, this last result means that two journals having equal IF, but a different number of champions, will go on to influence research published in journals of differing quality (measured in terms of IF and AIS).

6 | COMMENTS AND FURTHER IMPLICATIONS

Overall, the results presented in this section show that, in many journals, the articles published benefit from the presence of champions which attract the majority of citations, generating a sort of halo effect for the less-cited ones. This type of halo effect occurs and is utilized in many markets and activities and has no negative implications per se. However, its legitimization via bibliometrics—which renders it somehow "scientific," and often relied upon for research assessment—is problematic since the halo effect is based essentially upon a cognitive bias, as well-illustrated by psychology. Since all journals feature some papers with more citations than others, and indices are calculated at the journal level, most indices are systematically biased. In particular, non-champion articles and their authors benefit from the citations and the reputation obtained by the most-cited papers; this effect is much stronger in the case of some journals than others.

Any evaluation based on these indices incorporates all the distortions arising from the described bias. In practice, this has two different detrimental effects: The first, quite obvious one, is that constructing a metric with structural shortcomings yields systematic measurement errors. The second is that, precisely because of these shortcomings, the outputs of this metric are not comparable with each other. Yet their numeric character prompts observers to treat the outputs as if they were "pure" and reliable. This is an additional psychological effect, resulting from the human habit of ordering numbers and tending to attribute more objectivity to them, because the mathematical order drives what has been termed "mathematical intimidation" (Ewing, 2011).

It could be argued that, if a journal accepts an article, it is because its overall quality is similar to that of others previously published in that journal. And by extension, so is the quality of the author. However, such an assertion is questionable because, either citations have no meaning at all—in which case bibliometrics are meaningless too—or a paper that gets few or no citations must be less relevant to the scientific debate. It simply landed—by luck, or other reasons—in a journal where there are more interesting (better) papers. If citations are votes on the quality of an article, as all bibliometric and research evaluations seem to claim, then papers with few or no citations that are published in good—sometimes excellent—journals pose a conundrum. This raises the question of possible mistakes in paper selection or bias in the procedure, which opens up a separate avenue of investigation that is beyond the scope of this paper.

One example of bias is self-citation, which the JCR bibliometric index already takes into account, recognizing that these might signify a lesser impact on scientific research than citations from other journals. In point of fact, self-citations may indicate that the cited paper is of interest to a narrow scientific community (that focuses on the aims and scopes of that particular journal), while the latter testifies to interest from (and therefore impact on) a broader community.

That said, the questions posed here remain crucial to bibliometrics, also in light of the endogenous effect in science that tends to over-reward successful publication, such as the Matthew effect that enhances the outcomes of scholars who already have a pre-existing reputation (Migheli & Zotti, 2020). The implication might be that a bad paper (that is, with zero citations) landing in a good journal might be worth (and pay back) much more than a good paper (highly cited) published in a more modest journal.

7 | CONCLUSIONS

Reputation and its value are at the core of many activities in society today. Science and research are no exception. Most governments try to assess the research productivity and quality of their national system, universities, and research institutions much like sports teams and endeavor to recruit and reward productive "players"—that is to say, highly reputed scholars. This is the core of the modern research ecosystem.

This paper tries to advance the debate on scholarly publishing and the impact of bibliometric indices on the evaluation of research and its authors. Bibliometric indices were invented for a different and more limited purpose than evaluation, but since they provide figures based upon citations, they have become the basis for measuring the importance of scientific contributions, and in turn the overall value of research, somehow mimicking what has been done in the patent domain.

The rationale of the system is that reputation, based in this case on citation figures, can be converted into an objective measure. However, there is no theory supporting this assumption and, on the contrary, the

psychology literature suggests that cognitive biases such as the 'halo effect' can easily hamper evaluation of quality in scholarly publishing. If a paper is considered good because it was published in a good journal, its author gains the prestige associated to this journal even if the paper does not gain any citations at all. But zero citations is a clearer assessment of relevance to the scientific community: If a paper has zero citations, by the very rationale on which the entire system is based, it should not be considered a good paper.

The presence of the halo effect, and the subsidy of citations and reputation coming from other papers, would be straightforwardly confirmed if we discover that the papers published in a journal have skewed figures in term of citations. To test this, we conducted an empirical investigation using the oldest citations database and scientific report, Clarivate JCR for economics journals. Our findings confirm the skewness of citations within journals and the fact that a halo effect is in place.

Bibliometric indices in fact systematically incorporate the halo effect, and it is only on occasion absent, namely, when articles show a more uniform distribution in terms of citations. In most cases, though, the citations' distribution is highly skewed, in some journals more than in others, and a few champions may distort the value of the index used. In such cases, the exclusion of the champions might at least partially redress the picture, helping to more clearly evaluate the quality of the other papers published in the journal. For these reasons, the results presented in this paper recommend a cautious use of the indices—JCR in our case, and reasonably any bibliometric index—for evaluating the quality of scholars' research.

In some case, indices may loosely capture some quality of a journal. However, this work shows that the outward metrics may conceal the presence of "champions," which might artificially increase the average perceived quality of all the articles published in a journal, owing to the halo effect. This kind of shortcoming can be partially corrected by disentangling the analysis at the author level or—better yet—at the article level. However, such a task is resource-intensive and also has limitations.

ACKNOWLEDGEMENT

DATA AVAILABILITY STATEMENT

Data subject to third party restrictions.

ORCID

Giovanni B. Ramello 🄳 https://orcid.org/0000-0002-6510-1748

ENDNOTES

[1] This trend is further enhanced by the transition from an Information Society to a Big Data economy, which has made it possible to create new and very broad datasets on virtually anything (Marciano et al., 2020).

[2] Reputation is indeed central to many activities of our society: According to practitioners, it is today one of the major drivers of corporate value, and many studies have shown how—at the end of the day—brands can be conceived of as a way for capturing and capitalizing market value (among other ref. Ramello, 2006).

[3] The metamorphosis of "publish or perish" into "impact or perish," whereby citations and publication somehow merge, is extensively discussed by Biagioli and Lippman (2020).

[4] According to Karpik (2011), it is the structural lack of direct connection between demand and supply that makes it impossible to set a price for a scientific article, which of course justifies the rise of an indirect complex of evaluation

[5] In addition to academic community's inertia in using journals and their ranking as a measure for evaluating research, a number of scholars still believe that these are a good reference for assessing overall quality. For example, Waltman and Traag (2021) propose a formal model which shows that the impact factor of a journal is a good proxy for the quality of the individual articles published in it, thus providing further grounds for the usefulness of continuing to debate the use and relevance of such an indicator.

[6] Difficult of course does not mean impossible. As a reviewer correctly points out, today many databases exist to make the life of evaluators easier. Yet it is a delicate operation that requires accuracy, since not all citations merit equal "weight." Even within a given discipline, this may depend on where the citation appears, whether it is positive or negative, by how much it should be discounted based on the number of co-authors, and so on (Hamermesh, 2018). This is why many national research assessments continue to rely on rankings based upon bibliometric indices.

[7] It must be said that Garfield did understand the economic potential of his invention and, as early as 1956, he founded the Institute for Scientific Information, ISI and soon became a major part of the scientific division of Thomson Reuters (interview with William Roberts, July 29, 1997, available at http://www.garfield.library.upenn.edu/papers/oralhistorybywilliams.pdf)

[8] For a general overview on the halo effect, ref., e.g., Nisbett and Wilson (1977).

[9] Just to cite the most recent, see, for example, Hamermesh (2018), Heckman and Mokten (2020), and Van Dalen (2021).

[10] Additionally, it must be said that, since the audience of this journal consists mainly of economists and the authors themselves are also economists, there were practical reasons for studying the economics citation ecosystem.

[11] The AIS relies on the eigenfactor, an algorithm that claims—according to its developers and owner—to better capture the importance of a scholarly journal (Bergstrom et al., 2008).

[12] Available upon request. Generally, scholars are concerned especially with the outlets considered most important and relevant both in terms of reputation and from the viewpoint of their salary and career. This is an additional reason to focus on the top journals.

REFERENCES

Bergstrom, C. T., West, J. D., & Wiseman, M. A. (2008). The eigenfactor metrics. *Journal of Neuroscience, 28*, 11433–11434. https://doi.org/10.1523/JNEUROSCI.0003-08.2008

Biagioli, M. (2016). Watch out for the cheats in citation game. *Nature, 535*, 201. https://doi.org/10.1038/535201a

Biagioli, M., & Lippman, A. (2020). Introduction: Metrics and the new ecologies of academic misconduct. In M. Biagioli & A. Lippman (Eds.), *Gaming the metrics: Misconduct and manipulation in academic research*. MIT Press.

Bollen, J., van de Sompel, H., Hagberg, A., & Chute, R. (2009). A principal component analysis of 39 scientific impact measures. *PlosOne*, *4*(6), e6022. https://doi.org/10.1371/journal.pone.0006022

Bradshaw, C. J. A., & Brook, B. W. (2016). How to rank journals. *PlosOne*, *11*, e0149852. https://doi.org/10.1371/journal.pone.0149852

Burke, P. F., Dowling, G., & Wei, E. (2018). The relative impact of corporate reputation on consumer choice: Beyond a halo effect. *Journal of Marketing Management*, *34*(13–14), 1227–1257. https://doi.org/10.1080/0267257X.2018.1546765

Clemens, E. S., Powell, W. W., McIlwaine, K., & Okamoto, D. (1995). Career in print: Books, journals, and scholarly reputations. *American Journal of Sociology*, *101*, 433–494.

Davis, P. M. (2009). Reward or persuasion? The battle to define the meaning of a citation. *Learned Publishing*, *21*, 5–11. https://doi.org/10.1087/095315108X378712

Egert, T., & Scheufen, M. (2018). *The economics of open access: On the future of academic publishing*. Edward Elgar. https://doi.org/10.4337/9781785365768

Ewing, J. (2011). Mathematical intimidation: Driven by the data. *Notice of the AMS*, *58*, 667–673.

Garfield, E. (1955). Citation indexes for science: A new dimension in documentation through association of ideas. *Science*, *122*, 108–111.

Garfield, E. (2006). The history and meaning of journal impact factor. *Jama*, *295*, 90–93. https://doi.org/10.1001/jama.295.1.90

Griliches, Z. (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature*, *28*, 1661–1707.

Hall, B. H., Jaffe A. B., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics*, *36*: 16–38.

Hamermesh, D. S. (2018). Citations in economics: Measurement, uses, and impacts. *Journal of Economic Literature*, *56*(1), 115–156. https://doi.org/10.1257/jel.20161326

Harhoff, D., Narin, F., Scherer, F. M., & Vopel, K. (1999). Citation frequency and the value of patented inventions. *The Review of Economics and Statistics*, *81*, 511–515. https://doi.org/10.1162/003465399558265

Heckman, J. J., & Mokten, S. (2020). Publishing and promotion in economics: The tyranny of the top five. *Journal of Economic Literature*, *58*(2), 419–470. https://doi.org/10.1257/jel.20191574

Ippoliti, R., Ramello, G. B., & Scherer, F. M. (2021). Partnership and innovation in the pharmaceutical industry: The case of clinical research. *Economics of Innovation and New Technology*, *30*(3), 317–334. https://doi.org/10.1080/10438599.2019.1701782

Kahneman, D. (2010). *Thinking, fast and slow*. Random House.

Karpik, L. (2011). What is the price of a scientific paper? In J. Beckert & P. Aspers (Eds.), *The worth of goods* (pp. 62–85). Oxford University Press.

Larivière, V., & Sugimoto, C. R. (2019). The journal impact factor: a brief history, critique and discussion of adverse effects. In W. Glänzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators*. Springer Verlag. https://doi.org/10.1007/978-3-030-02511-3_1

Marciano, A., Nicita, A., & Ramello, G. B. (2020). Big data and big techs: Understanding the value of information in platform capitalism. *European Journal of Law and Economics*, *50*, 345–357. https://doi.org/10.1007/s10657-020-09675-1

Merton, R. K. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *Isis*, *79*, 606–623.

Migheli, M., & Ramello, G. B. (2018). The market of academic attention. *Scientometrics*, *114*, 113–133. https://doi.org/10.1007/s11192-017-2564-y

Migheli, M., & Ramello, G. B. (2013). Open access, social norms & publication choice. *European Journal of Law and Economics*, *35*, 149–167.

Migheli, M., & Zotti, R. (2020). The strange case of the Matthew effect and beauty contests: Research evaluation and specialisation in Italian universities. *Socio-Economic Planning Sciences*, *71*, 100749. https://doi.org/10.1016/j.seps.2019.100749

Moser, P., Ohmstedt, J., & Rhode, P. W. (2015). "Patent citations and the size of patented inventions. Evidence from hybrid corn.", NBER Working Papers 21443

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology.*, *35*(4), 250–256. https://doi.org/10.1037/0022-3514.35.4.250

Origgi, G. (2018). Reputation. *What it Is and why it matters*. Princeton: Princeton University Press.

Origgi, G., & Ramello, G. B. (2015). Current dynamics of scholarly publishing. *Evaluation Review*, *39*, 3–18. https://doi.org/10.1177/0193841X15572017

Osterloh, M., & Frey, B. S. (2020). How to avoid borrowed plumes in academia. *Research Policy*, *49*(1), 103831. https://doi.org/10.1016/j.respol.2019.103831

Pendlebury, D. A., & Adams, J. (2012). Comments on a critique of the Thomson Reuters journal impact factor. *Scientometrics*, *92*, 395–401. https://doi.org/10.1007/s11192-012-0689-6

Pepall, L. M., & Richards, D. J. (2002). The simple economics of brand stretching. *Journal of Business*, *75*, 535–552. https://doi.org/10.1086/339888

Ramello, G. B. (2006). What's in a sign? Trademark law and economic theory. *Journal of Economic Surveys*, *20*(4), 547–565. https://doi.org/10.1111/j.1467-6419.2006.00255.x

Rousseau, S., & Rousseau, R. (2021). Bibliometric techniques and their use in business and economics research. *Journal of Economic Surveys*, in press. https://doi.org/10.1111/joes.12415

Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *BMJ*, *314*, 497–502. https://doi.org/10.1136/bmj.314.7079.497

Stephan, P. (2010). The economics of science. In B. H. Hall & N. Rosenberg (Eds.), *Handbook of the economics of innovation*. North Holland. https://doi.org/10.1016/S0169-7218(10)01005-1

Trajtenberg, M. (1990). A penny for your quotes: Patent citations and the value of innovations. *RAND Journal of Economics*, *21*, 172–187. https://doi.org/10.2307/2555502

Van Dalen, H. P. (2021). How the publish-or-perish principle divides a science: The case of economists. *Scientometrics*, *126*, 1675–1694. https://doi.org/10.1007/s11192-020-03786-x

Waltman, L., & Traag, V. A. (2021). Use of the journal impact factor for assessing individual articles; statistically flawed or not? *F1000 Research*, *9*, 366. https://doi.org/10.12688/f1000research.23418.2

Woolston, C. (2015). Build a reputation. *Nature*, *521*, 113–115. https://doi.org/10.1038/nj7550-113a