Should citations be weighted to assess the influence of an academic article?

Damien Besancenot, LIRAES and Université Paris 5, Sorbonne Paris Cité

Abdelghani Maddi, CEPN and Université Paris 13, Sorbonne Paris Cité , OST-HCERES.

Abstract: Citations are by nature heterogeneous. A citation worth may dramatically vary according to the influence of the citing article or to the journal's reputation from which it is issued. Therefore, while assessing the influence of an academic article, how should we weight citations to take into account their real influence? In order to answer this question, this article suggests various methods of weighting citations in the building of articles quality indexes. These indexes are then used to measure the influence of the articles published in the top five economic journals over the 2000-2010 period and analyses the sensibility of these indexes to the choice of the weighting schemes. Our main result is that whatever the weighting scheme, information carried by the different indexes is not significantly different. From Occam's razor principle, the number of citations provides an efficient and sufficient tool to measure research quality.

Keyword: Citations, Articles' ranking, weighting functions, Pagerank, Eigenfactor.

JEL codes : A14, C18, C43

Section 1. Introduction

In the assessment of research, citation counts are ubiquitous. Literature classifies researchers by the impact of their work captured through citations in academic journals (Medoff, 1989, 1996, Coupé 2003). Citations influence academic careers and researchers' rewards (e.g. Hamermesh et al., 1982, Diamond Jr. 1989 Moore *et al*, 1998, Bratsberg *et al*, 2010), contribute to journals reputations (Pinski, G. et al, 1976 , Palacio-Huerta and Volgi, 2004, Bollen et al, 2006 , Ritzberger, 2008 , Gonzalez-Pereira et al., 2009) and allow the assessment of research departments, universities and countries (Dusansky and Vernon 1998, Braun et al., 1996, Bordons et al., 2002). More recently, a strand of literature measures the influence of academic articles according to the number of citations they received. For instance, Van Noorden, Maher, Nuzzo, et al. (2014), provides a list of the most highly cited articles (independently of the academic field) and, in economics, Kim, Morse, and Zingales (2006) built the list of the academic articles cited more than 500 times. In this strand of research, a high number of citations appears as a signal of the intrinsic quality of an article.

However a citation doesn't exhibit the same worth depending on the citing article or on the journal where this citing article is published. In economics, being referenced by an article published in *The Amercican Economic Review* has not the same significance than being cited in a second or a third tier journal. Influent citations contribute to a higher degree to the visibility of the cited article and, logically, bibliometric indexes should take into account the heterogeneity of the pool of citations[1].

The relative weight of each citation may be taken into account in different way. Recently, the notion of PageRank appeared in bibliometric analysis by replication of the methodology developed for the search engine Google (Brin and page, 1998, Altman and Tennenholtz, 2005, 2008). In computer science, the PageRank method attributes a numerical index - the "rank" of a Web page - by quantifying the number of hyperlinks pointing on each page. In bibliometrics, this

[1] Studying the most cited articles published in Brazilian Journal of Economics, Faria (2010) reports that an article by Tan and Werlang (1992) published in the *Revista Brasileira de Economia* was often misquoted as being published in the *Journal of Economic Theory*. He found out that the first author who made the mistake was Roger Guesnerie who cited the article *in the American Economic Review, Econometrica* and *Quarterly Journal of Economics*. The repetition by other authors of the same mistake indicates that these authors cited the article because of Guesnerie's first reference.

approach leads to a definition of a citation worth as a function of the number of articles citing the citing article (see for instance Bollen, Rodriguez and Van de Sompel, 2006).[2]

The measure of a citation worth should also take into account the quality of the publishing medium in which citing articles are published. The benchmark method here is close to the Eigenfactor approach (see Bergstrom, 2007, Bergstrom and West 2008, Bergstrom, West and Wiseman 2008) which measures a journal influence according to the number of times its articles have been cited but also considers which journals have contributed these citations. When journal's quality is measured by its Eigenfactor, highly cited journals are considered as more influent than lesser cited journals.

In these setting, this article has two main objectives: first, to suggest various way of introducing the value of a citation in the assessment of articles influence and second, to measure the sensibility of the resulting indexes of academic influence to the different ways in which the weight of citations is taken into account.

Our work therefore considers the 3142 articles published in the top five economic journals *(The American Economic Review, Econometrica, Journal of Political Economy, Reviews of Economic Studies et Quarterly Journal of Economics*) over the period January 2000 - December 2010 (our benchmark articles). We identified every article citing these articles and listed in the *Web of Science* (first degree citing articles) and each article citing the first degree citing articles (second degree citing articles). With this database, we defined three families of indicators: the first assesses the quality of an article by the raw number of citations. The second adopts a PageRank approach and considers that a citation presents a higher value if the citing article (first degree) is itself frequently cited (second degree). The third approach pays attention to the journal where the citing articles (first degree) where published.

Bibliometric literature proposes various indexes measuring both the quantity and the quality of citations. In most of the cases, indicators are built as the sum of citation scores where citation scores are computed according various methodologies (see for instance Waltman and Van Eck, 2010). One of the limits to this approach is that it considers all citations in a given period of time. An article assessed according to this approach may record a high index either because it has been cited by influent articles or because it received a lot of citations from second tier articles. In order to focus only on significant citations, our indicators are built by analogy to the h and g indexes suggested by Hirsh (2005) and Egghe (2006). These indicators increase with the number

2 The Pagerank approach is close to the invariant method developed by Pinski and Narin (1976) to rank journals in chemistry and physics. In economics this seminal approach led to a series of important articles, see for instance Liebowitz and Palmer (1984), Laband and Piette (1994) or Palacio-Huerta and Volgi (2004).

of influent citations but neglect citations from poorly read articles or coming from publications in journals with low audience.

Finally, we also computed these indexes considering only citations published in the 600 journals ranked by Combes and Linnemer (2010). This allows us to take into account the influence of the set of journals in which citations are recorded and to build and compare 10 alternative quality indexes for our benchmark articles.

The main result of our study is that the information conveyed by the different indexes is not significantly different whatever the way citation influence is taken into account. When we rank our benchmark articles according to the various indexes, rankings do not appear as statistically different (Pearson's rank correlation coefficients are high whatever the rankings considered). More formally, a principal component analysis (PCA) allows us to show that there are only minor differences between the three families of indicators. According to Occam's razor principle, the use of the raw number of citations as a measure of articles influence appears efficient – at least for articles published in the top five journals in economics.

Note that the aim of this article is not to suggest new indicators to assess articles' quality. Even if our indexes propose an original way of measuring the influence of a scientific contribution, they may be easily criticized. For instance, they ignore the influence of the research topic (Ellison, 2013), gender (Rossiter, 1993, Maliniak et al. 2013) or coauthorship (Levitt, 2015) on the propensity to be cited. Our indexes are therefore imperfect as measures of articles influence. However, as these biases affect our indicators in the same way for each article, they don't bring any distortion between indexes while considering a given article and these indexes may be statistically compared.

The article is organized as follows: section 2 describes the database; section 3 provides the results of our statistical analysis and discusses the results; section 4 concludes the article.

Section 2: Data

Our study considers the set of articles published by the five top journals in economics (*The American Economic Review, Econometrica, Journal of Political Economy, Reviews of Economic Studies* and *Quarterly Journal of Economics*) over the period 2000-2010. In March 2015, we used the *Web of Science (WOS)* to retrieve first the citations received by these benchmark articles (first degree citations) and then the citations received by the citing articles (second degree citations). The Dataset includes 3142 benchmark articles, 57.244 first degree citing articles and 191.000 second degree citing articles. In order to normalize the time period in which citations were recorded, we restricted these citations to the four years window following the publication of the cited articles.

We then built 10 indicators aiming at reflecting the scientific influence of the articles published by the top five journals. Basically, these indicators rely on the number of citations, however they differ in the weight accorded to each citation.

The first index (T_C index) only records the number of citations:

- *T_C* index*:* this index indicates for each article of the top five journals the raw number of citations recorded in WOS.

The two following indexes (P_h *and* P_g) develop a PageRank approach to assess articles' influence (indexes P for PageRank). In these indexes, the worth of a citation is linked to the number of articles citing the citing article. Indexes are built by analogy to the h and g indexes proposed by Hirsch (2005) and Egghe (2006) to measure researchers' influence. Consequently, a citation is considered only if the citing article itself is influential.

- P_h *Index :* this index applies the Hirsch (2005) methodology. An article has therefore a P_h *Index* equal to x if there is at least x citing articles that are cited at least x time.
- P_g *Index :* this index measures an article's influence by reference to the g index (Egghe 2006). A P_g *Index* is then equal to z if the z more influential citing articles received together at least z^2 citations.

For instance, let us consider an article which is referred to in four articles. Article 1 is itself cited five times, article 2 is cited three times, article 3 only once and article 4 received no citation. Given these citations, our benchmark article has a P_h *Index* equal to 2, only two of the citing articles are cited more than twice, and a P_g *Index* of three as the sum of the citations received by the three more influential citing articles *is equal to 9 (i.e. the squared value of the rank* 3^2*=9).*

By construction, the P_g index puts a specific emphasize on the influence of the most important citing articles. If a benchmark article is cited by only a few influent articles, it may have a high P_g index and a low P_h. The two indexes don't capture the same effects.

The next indexes consider that a citation worth is linked to the quality of the journal in which is published the citing article (indexes J for Journal). In order to measure journals' quality, we use the quality indexes proposed by Combes and Linnemer (2010) in their journals' ranking. In their article, these authors give two score indexes (CLm and CLh) for each of the 1202 Econlit Journals. These measures are the result of the same estimation and therefore imply the same ordinal ranking of journals, however, they differ in the relative weights they give to the various journals. Once ranked from the most prestigious to the lowest one, the plot of the score indexes takes the shape of two decreasing functions valued in the range [0, 100] with a convexity indicating a more or less high selectivity of the index. The CLm index (where m stands for medium convexity) thus exhibits a relatively low difference between the weight assigned to the journals of the top tier and a journal in the middle of the ranking while this difference is higher with the CLh index (where h stands for high convexity). With these two quality measures, we defined two alternative sets of indexes.

First, two indexes are computed considering citing articles published in one of the 1202 Econlit Journals with an explicit reference to the h index.

- J_h^{CLh} Index: this index presents a value x if the cited article has received at least x citations from articles published in journals with a weight CL_h exceeding or equal to x. For instance consider an article cited by five journals' articles. If only three of these journals present a CL_h index exceeding or equal to 3, the article's J_h^{CLh} Index will be equal to 3.
- J_h^{CLm} Index: this index replicates the previous index but uses the CL_m weight system. An article presents a J_h^{CLm} Index with value x if the cited article has received at least x citations from articles published in journals with a weight CL_m exceeding or equal to x.

The two following indexes are computed considering citing articles published in Econlit journals with an explicit reference to the g index.

- J_g^{CLh} Index: for a given article, this index takes value x if the sum of the CLh score of the best x journals (ranked according to Combes and Linnemer (2010) CLh classification) publishing an article citing the reference article is higher or equal to x^2. For a given article, a J_g^{CLh} Index equal to 5 means that the sum of the five best CLh journals' score exceeds 25 (i.e. 5^2).
- J_g^{CLm} Index : for a given article, a J_g^{CLm} Index of x means that the sum of the CLm score of the x best journals (ranked according the CLm score) publishing an article citing article P is higher or equal to x^2.

Finally, the last three indexes replicate the previous methodology but restrict the set of citations to citing articles published in the most influential journals. We consider here that a journal is influential if it is included in the 600 journals list considered in Combes and Linnemer [2010]. These indexes therefore give a more elitist measure of the articles' influence by considering a restricted number of potentially more influent citations.

- P^R_h *and* P^R_g *Indexes:* these indexes use the same methodology that the P_h and the P_g indexes but with the above mentioned restricted set of citations.
- T^R_C: gives the raw number of citations coming from articles published in one of the 600 journals listed by Combes and Linnemer (2010).

a) Descriptive statistics :

Table 1 gives the number of articles published by each of the top five journals over the 2000-2010 period. The *AER* clearly appears as a leading journal in the number of publications. With a mean number of 99 published articles by year, it publishes 62% more articles than Econometrica and more than twice the number of publications of the three other top journals.

Journal	Nb articles	Articles / Year	%
AER	1084	99	35%
Eco	668	61	21%
QJE	457	42	15%
RES	480	44	15%
JPE	453	41	14%

Table 1: Number of articles published by each journal over the 2000 – 2010 period.

Table 2 presents the mean number of citations by article in the four year window. If we consider this number as representative of the influence of each journal, the *Quarterly journal of economics* appears as the most influent (with a mean number of 27 citations per article), followed by *The American Economic Review* (18) and the *Journal of Political Economy* (17). For the five journals, the mean number of citation is 18 per article in the four year window.

Journal	Articles	Citations	Citations / article
AER	1084	19814	18
Eco	668	10992	16
QJE	457	12382	27
RES	480	6461	13
JPE	453	7595	17
Total	3142	57244	18

Tableau 2: Mean number of citation in the four year window

Table 3 lists the ten articles that received the highest number of citations and their respective rank according to our 9 citation indexes. Amongst the ten most cited articles, five were published in the AER, two in the QJE, and one article in each of the three other journals (JPE, Econometrica and RES).

Author(s)	Title	Journal	Issue	Year	LP/SP	Indexes TC		Indexes J				Indexes *P*			
						T_C	T^R_C	J_h^{CLh}	J_g^{CLh}	J_h^{CLm}	J_g^{CLm}	P_h	P_g	P^R_h	P^R_g
Lawrence J. Christiano, Martin Eichenbaum, and Charles L. Evans	Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy	JPE	1	2005	LP	1	1	1	1	3	1	35	17	35	16
Marianne Bertrand, Esther Duflo, and Sendhdl Mullainathan	How Much Should We Trust Differences-In-Differences Estimates?	QJE	1	2004	LP	2	2	2	4	2	3	10	12	10	12
Jeffrey R. Kling, Jeffrey B. Liebman and Lawrence E Katz	Experimental Analysis of Neighborhood Effect	ECO	1	2007	LP	3	18	82	70	75	53	26	25	26	24
Frank Smets and Rafael Wouters	Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach	AER	3	2007	LP	4	3	9	10	4	4	67	105	67	98
Elhanan Helpman, Marc Melitz and Yona Rubinstein	Estimating Trade Flows: Trading Partners and Trading Volumes	QJE	2	2008	LP	5	5	77	107	47	80	176	314	178	281
Marc J. Melitz and Giancarlo I. P. Ottaviano	Market Size, Trade, and Productivity	RES	1	2008	LP	6	4	32	23	8	7	59	127	59	124
Lutz Kilian	Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market	AER	3	2009	SP	7	7	238	430	87	156	83	138	83	135
Gary E Bolton and Axel Ockenfels	ERC: A Theory of Equity, Reciprocity, and Competition	AER	1	2000	LP	8	6	7	9	15	4	1	1	1	1
Urs Fischbacher and Simon Gächter	Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments	AER	1	2010	SP	9	28	238	703	26	136	20	19	20	19
James E. Anderson and Eric Van Wincoop	Gravity with Gravitas: A Solution to the Border Puzzle	AER	1	2003	LP	10	8	25	50	43	36	7	3	7	3

Table 3: The ten most cited articles and their ranking according to the different citation indexes.

Table 3 suggests that the rankings of articles with respect to the J-indexes and to the raw number of citations (T_C) are closer compared to rankings resulting from the P-indexes (see for instance articles 1, 2, 4 and 6). However, in some cases, the raw number of citations leads to a ranking close to the Pagerank ranking (Cf. article 10) while for articles 7 and 9, *P* and *J* indexes lead to similar ranks. This heterogeneity illustrates the difficulties raised by the use of citations in assessment of academic influence.

However, the ranks of the ten most cited articles don't give a correct image of the whole dataset. In order to evaluate the proximity between the ten rankings arising from our indexes, for each index, we built a ranking of the articles published in the 5 journals and computed the Spearman's coefficient for each couple of rankings. Table 4 gives the mean values of the spearman's rho in our database, the lowest coefficient is equal to 0.7 and the mean value is 0.84. According to this table, there is a strong correlation between the different rankings. Even if the 10 indexes measure the worth of a citation according to rather different criterions, the ranking they induce appears to be poorly sensitive to these criterions.

Variables	T_C	P_h	P_g	J_h^{CLh}	J_g^{CLh}	J_h^{CLm}	J_g^{CLm}	P^R_h	P^R_g	T^R_C
T_C	**1,0**	**0,7**	**0,8**	**0,8**	**0,8**	**0,8**	**0,9**	**0,7**	**0,8**	**1,0**
P_h	**0,7**	**1,0**	**0,9**	**0,7**	**0,7**	**0,7**	**0,8**	**1,0**	**0,9**	**0,7**
P_g	**0,8**	**0,9**	**1,0**	**0,7**	**0,8**	**0,8**	**0,8**	**0,9**	**1,0**	**0,8**
J_h^{CLh}	**0,8**	**0,7**	**0,7**	**1,0**	**0,9**	**0,9**	**0,9**	**0,7**	**0,7**	**0,8**
J_g^{CLh}	**0,8**	**0,7**	**0,8**	**0,9**	**1,0**	**0,9**	**1,0**	**0,7**	**0,8**	**0,8**
J_h^{CLm}	**0,8**	**0,7**	**0,8**	**0,9**	**0,9**	**1,0**	**1,0**	**0,7**	**0,8**	**0,9**
J_g^{CLm}	**0,9**	**0,8**	**0,8**	**0,9**	**1,0**	**1,0**	**1,0**	**0,8**	**0,8**	**0,9**
P^R_h	**0,7**	**1,0**	**0,9**	**0,7**	**0,7**	**0,7**	**0,8**	**1,0**	**0,9**	**0,7**
P^R_g	**0,8**	**0,9**	**1,0**	**0,7**	**0,8**	**0,8**	**0,8**	**0,9**	**1,0**	**0,8**
T^R_C	**1,0**	**0,7**	**0,8**	**0,8**	**0,8**	**0,9**	**0,9**	**0,7**	**0,8**	**1,0**

Table 4 : Pearson's Rank correlation matrix

Section 3. Statistical analysis

In order to focus on the relationship between our various indexes and the resulting measures of articles quality, we developed a Principal Components Analysis (PCA) using graphical representations in two dimensions.

One methodological issue in our study is linked to the period over which citations are collected. As our study covers articles published between 2000 and 2010, older articles naturally receive a higher mean number of citations than the more recent ones and therefore the value of our indexes is time dependant (see Figure 1).

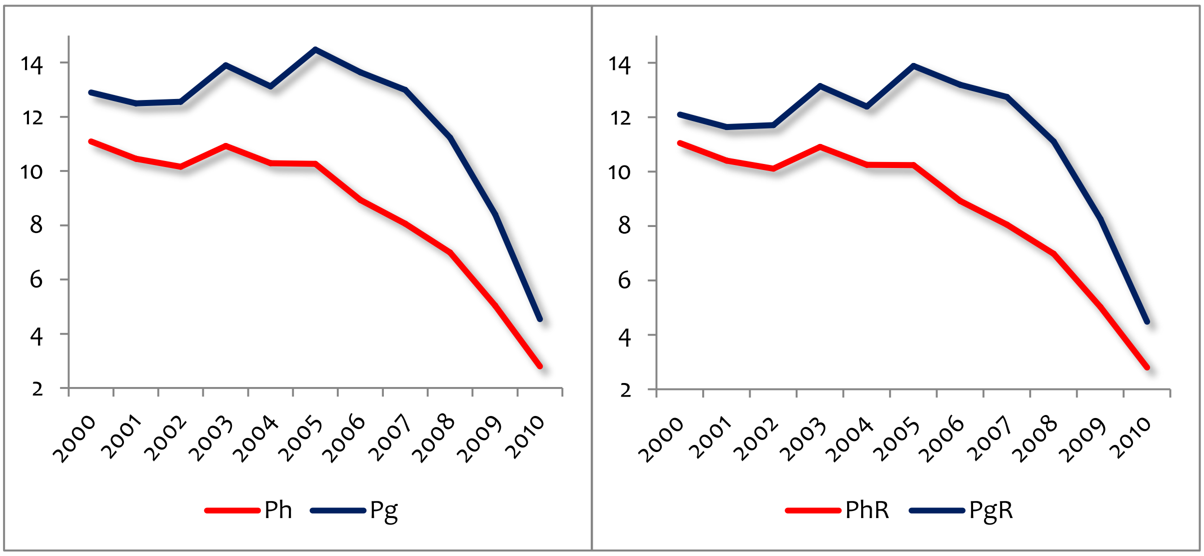


Figure 1: mean values of the Ph/g and PRh/g indexes for each year of publication

In order to avoid this issue, the dataset has been split in two subsets, a first subset considers articles published during the period 2000-2005 (i.e. 1752 articles) and the second, articles published between 2006 and 2010 (1390 articles). For each subset, we performed a specific Principal Components Analysis.

For period 2000-2005, Kaiser's criterion indicates that 97.16% of the total inertia is explained in the first two axis where the first one (Axis 1) allows explaining 94,33% of the overall information of our PCA. Table 5 shows that all variables are highly positively correlated with the first axis and weakly correlated with the second.

	Axis 1	Axis 2
T_C	0,988	-0,046
P_h	0,956	-0,247
P_g	0,989	-0,081
J_h^{CLh}	0,920	0,324
J_g^{CLh}	0,959	0,183
J_h^{CLm}	0,979	0,102
J_g^{CLm}	0,990	0,031
P^R_h	0,958	-0,239
P^R_g	0,988	-0,066
T^R_C	0,983	0,054

Table 5: Correlations between variables and factors (2000-2005)

Furthermore Table 6 shows that our ten indexes contribute equally, 10% on average, to the formation of the main axis (axis 1).

	Axis 1	Axis 2
T_C	10,340	0,756
P_h	9,698	21,503
P_g	10,369	2,306
J_h^{CLh}	8,977	37,000
J_g^{CLh}	9,749	11,771
J_h^{CLm}	10,163	3,661
J_g^{CLm}	10,397	0,334
$P^R{}_h$	9,726	20,102
$P^R{}_g$	10,343	1,555
T^R_C	10,238	1,012

Table 6: Contributions variables (%) - (2000-2005)

While considering period 2006-2010, results appear to be very similar. Kaiser's criterion indicates that 95,31% of the information is captured by the two first axes (84,18% by axis 1 and 11,13% by axis 2) and the different indexes are highly correlated with axis 1 (see table 7). Therefore, these indexes are strongly correlated with each other.

Variables	Axis 1	Axis 2
T_C	0,945	-0,188
P_h	0,874	0,461
P_g	0,910	0,380
J_h^{CLh}	0,880	-0,333
J_g^{CLh}	0,914	-0,288
J_h^{CLm}	0,944	-0,271
J_g^{CLm}	0,958	-0,246
$P^R{}_h$	0,876	0,458
$P^R{}_g$	0,923	0,344
T^R_C	0,946	-0,255

Tableau 7 : Correlations between variables and factors (2006-2010)

Table 8 replicates table 6 for period 2006-2010 showing that all variables contribute equally to Axis 1. One difference between the two sub periods comes from Axes 2 wich seems to introduce a slight difference between the P and the J indexes. We will come back to this point later on.

Fig. 2. Below presents the various indexes on the correlation circle for the two sub periods. Each index is associated with a dot whose coordinates are given by the correlation between the factor and the index. The strength of the relationship between two variables is given by the squared cosine of the angle between the radius of the circle passing through the

representative dots. The smaller the angle, the stronger the correlation between the two variables.

Variables	Axis 1	Axis 2
T_C	10,619	3,166
P_h	9,084	19,073
P_g	9,841	12,978
J_h^{CLh}	9,206	9,983
J_g^{CLh}	9,923	7,441
J_h^{CLm}	10,575	6,616
J_g^{CLm}	10,893	5,435
$P^R{}_h$	9,119	18,832
$P^R{}_g$	10,114	10,612
T^R_C	10,625	5,861

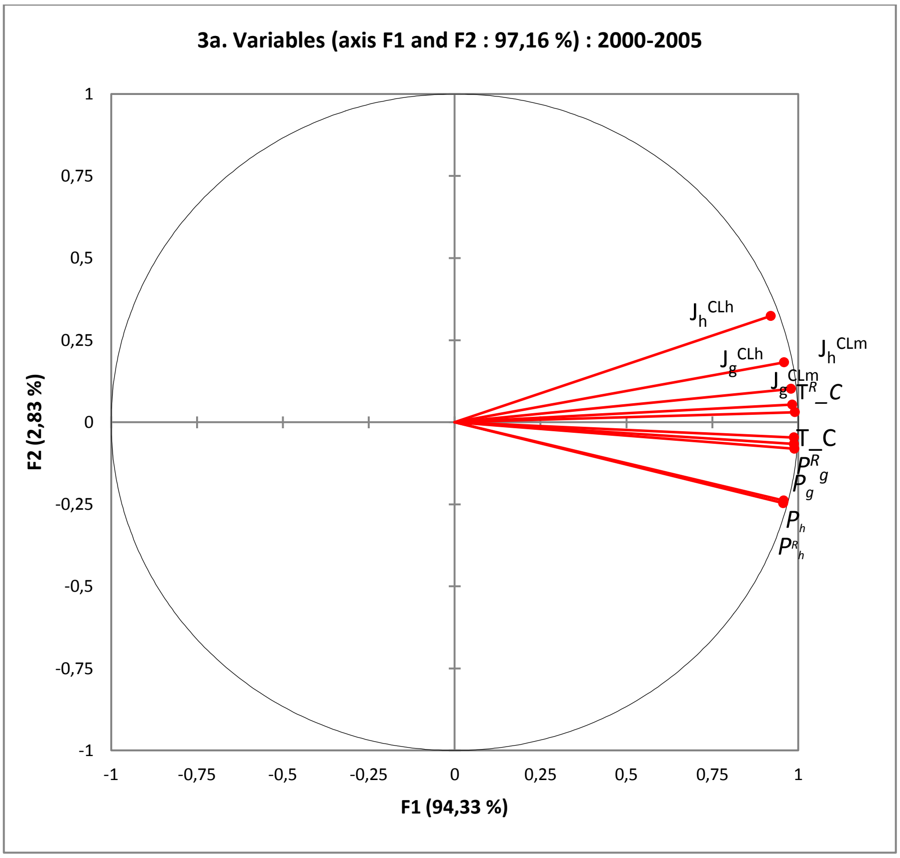
Tableau 8 : Contributions variables (%)

Note that, in both circles figures, dots locations are very close to the edge of the circle and are therefore very representative of the variables (indexes) on the plane. The proximity between the vectors – as measured by the angles they form pairwise - allows us to determine whether indexes (or evaluation methods) are similar.

Figure 2. therefore illustrates the high correlation between the different variables for articles published during period 2000-2005. We note a slight difference between the J- indexes (J_h^{CLh}, J_g^{CLh}, J_h^{CLm}, J_g^{CLm}) that take into account the quality of the citing journal, indexes which are above axis 1 and the P¨-indexes (P_h, $P^R{}_h$, P_g, $P^R{}_g$) *below* the same axis and focusing on the influence of the citing article. Both types of indexes seem also highly correlated with the two raw indexes: T_C and T^R_C (respectively the total number of citations received by each article and the number of citations from a journal surveyed in Combes and Linnemer, 2010). Note that these two last indexes present the highest correlation with axis 1 – respectively 0.945 and 0.946 – meaning that this axis is closely related to the raw number of citations.[3]

Rather equivalent results may be observed while considering period 2006-2010. In Fig. 4, two sets of vectors are still distributed on each side of the horizontal axis. However, things seem to be more contrasted as the two blocks of variables exhibits a V shape with a 45 degrees angle. Above axis 1, vectors are representatives of indexes focusing on citing articles (P-indexes), below this axis we find the J-indexes (built considering the influence of the citing journal). As shown in Fig. 4, indexes T_C and T^R_C are plotted below the horizontal axis, meaning that they are more correlated to the J-indexes than to the *P* ones. Like for the period 2000-2005, the indexes grounded on the quality of the citing journal are higly correlated with the raw number of citations – whatever the way these citations are recorded. Things are rather different for the P-indexes as

[3] Index J_g^{CLm} is the is the only one with a stronger correlation.



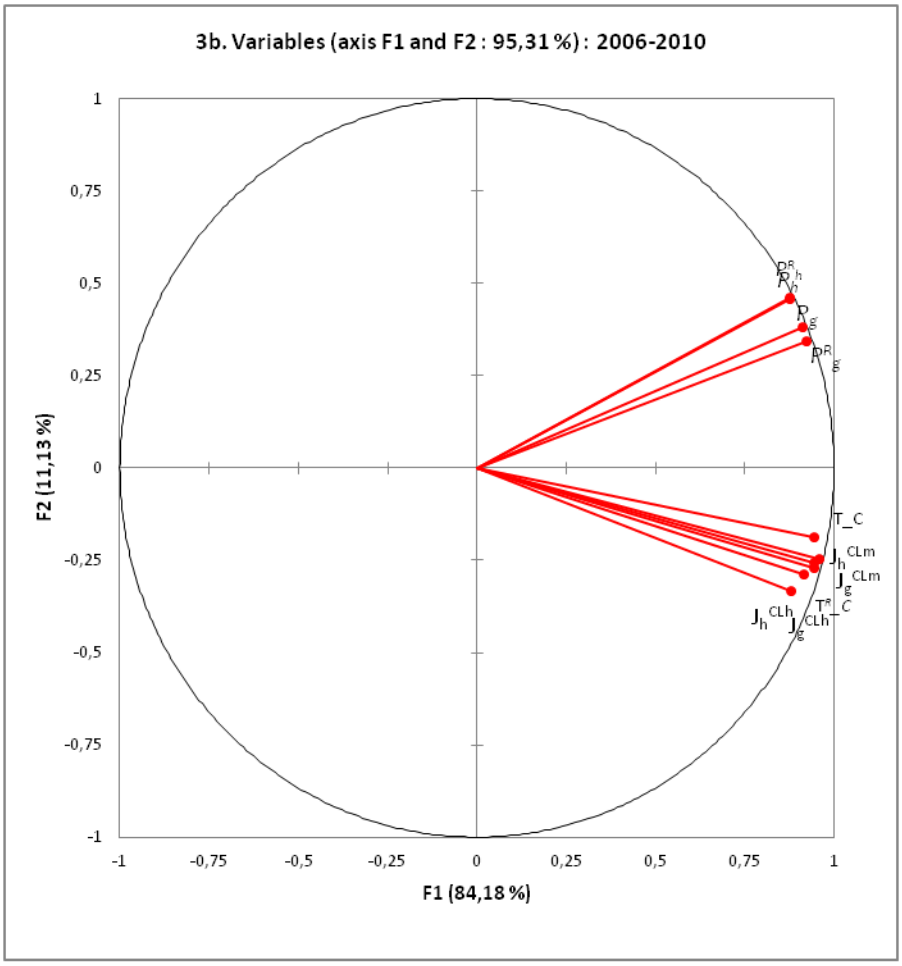


Fig 2. Figure 2 : ACP Axis

time plays a crucial role in the number of citations. Recall that the P-indexes measure the worth of a citation according to the number of citations received by the citing article. As writing and publishing articles is a time consuming process, an article may hardly record a significant P index before four or five years after its publication. For articles published in the 2006-2010 period, the P-indexes computed with citations recorded until 2015 appear to give an incomplete measure of articles influence, an imperfection which is corrected through time as reflected by the correlation between the P and the J-indexes for the 2000-2005 period.

Finally, the Principal Components Analysis allows emphasizing the strong link between the various indexes. Statistically, the assessment of articles' influence leads to very similar results if one counts only the raw number of citations or if one considers more subtle measures taking into account the influence of the citing articles or quality of the citing journals.

This result is not surprising. First, the fact that rankings are poorly influenced by the choice of the set of citing journals reflects the specific nature of our indexes. By construction, these indexes consider the most influent citations, i.e. citations coming from highly cited articles or from publications in influent journals. Restricting the set of journals to the best ones only removes citations coming from second tier journals or authors with only little influence. It barely affects our indexes and the resulting ranking.

Moreover, even if they rely on different measures of a citation worth, our indexes are basically grounded on the same benchmark. If the best academic journals were able to attract and select only the highest quality articles, these articles should receive a high number of citations from very influential authors published in the best journals. Measuring citation quality with a specific focus on the number of citations that the citing article receives or on the journal where this citing article has been published should not make a large difference. In this case, our indexes would lead to the same ranking whatever the way a citation worth is measured.

The Matthew effect (Merton, 1968) also contributes to smooth the differences between our measures of a citation worth. While choosing their bibliographical references, researchers present an obvious bias toward recognized authors and prestigious journals. First degree citing articles have a higher probability of being cited if they are published in top tier journals and a high J-index should induce a high P-index. Moreover, a citation by an influent article increases the visibility of the cited article and fosters further citations (see footnote 1). A high J index, meaning that first degree citing articles are published in top tier journals also implies a high T_C index (citation in good journals induces additional citations) and guaranties that citing articles have a high probability of being cited (the P-indexes must also be high).

The difference revealed by the PCA between the J and the P-indexes would not exist if top tier journals were able to publish only the highest quality articles. However, articles' selection is sometime inaccurate and referees may reject good articles or accept articles with little scientific contribution.[4] If an influent journal selects articles that will be poorly cited, a bias is naturally introduced between our indicators. A first degree citing article with low audience published in such a journal may contribute both to an increase of the J indexes and a relative decrease of the P indexes.

Alternative mechanisms are at work considering the influence of research topics on the volume of citations (Ellison, 2013). If an influent journal deals with highly specialized topics that induce only few citations, the J and the P indexes will diverge. Journal specificities may also induce a difference between the two kinds of indicators. For instance, when a journal publishes literature reviews, articles may be cited by a high number of PhD students but will be ignored by senior researchers (Bollen et al. 2006). Only few citations will be found in the best journals and, once again, the two indicators will diverge.

Finally, our study tends to demonstrate that if the J and the P indexes may capture different phenomena, these differences are low and the simple citation count is highly correlated with other indexes. In a nutshell, despite its simplicity, citation count appears efficient.

Section 4. Conclusion

Over the last thirty years, the rise of research assessment procedures induces the search for efficient tools able to give correct measures of research efficiency. Most of the time, researcher performance and journal influence are evaluated through a simple count of the citations they receive. However, citations are intrinsically heterogeneous. The set of journals in which citations are collected and the impact of the citing articles lead to high discrepancies in the measure of researchers or journals' influence.

The aim of this work is to measure the sensibility of rankings to the metrics used to assess citation worth. More specifically, our work focuses on the assessment of articles influence. To do so, we considered all articles published by the top five journals of economics over the period 2000-2010 and built several indicators reflecting their influence. Each indicator was grounded on citation count but the indexes differ from each other in the way citations were valued. Our purpose was to test the sensibility of these indexes to the way we measure citation worth.

[4] The matching between articles and journal is frequently imperfect. For instance, Oswald (2007) pointed out that the most cited articles published in a second tier journal receive more citations than the 4 least cited articles of any first tier journal such as the AER. In the same way, Gans and Shepherd 1994 give multiple examples of seminal articles rejected by top tier journal. In our database, 34 articles published in the top five journals received no citation, 10 of them were published in the JPE, 9 in the AER, 9 in Econometrica, 5 in RES, and one in the QJE.

Despite their differences, our indexes appear to be highly correlated and lead to statistically similar rankings. According to Occams rule, for a given scientific field, assessing articles' influence by a simple citation count appears to be an efficient procedure.

This result is not exempt from criticism. First, our database only considers a very specific set of benchmark articles. A more general study should check these results with a more important set of articles. In this case, the problem of self-citation would have to be considered. Moreover, our approach let apart the characteristics of the citing authors. A citation coming from a Nobel price would not have the same value than a citation from a PhD student. Citation worth is certainly linked to the characteristics of its author. In turn this rises important methodological problems when a citing article results from a collaboration between authors with different level of notoriety. These points deserves in depth analysis in further research.

References

Altman A. and M. Tennenholtz (2005), "Ranking systems: the PageRank axioms", Proceedings of the ACM Conference on Electronic Commerce, p. 1-8.

Altman A. and M. Tennenholtz (2008), "Axiomatic Foundations for Ranking Systems," *Journal of Artificial Intelligence Research*, p. 473-495.

Bedeian, A.G., (2003). The Manuscript Review Process: The Proper Role of Authors, Referees and Editors, *Journal of Management Inquiry,* 12, p. 331-338.

Bergstrom C. (2007), Eigenfactor: Measuring the value and prestige of scholarly journals, C&RL News, 68, p. 314-316.

Bergstrom, C.T. and. J.D. West (2008), Assessing citations with the Eigenfactor™ metrics, Neurology, 71, p. 1850–1851.

Bergstrom, C.T., J.D. West and M.A. Wiseman (2008), The Eigenfactor™ metrics, Journal of Neuroscience, 28(45), p. 11433–11434.

Bratsberg, B., Ragan, J., and J., Warren, (2010), "Does raiding explain the negative returns to faculty seniority?", Economic Inquiry, 48(3), p. 704–721.

Bollen, J., Rodriguez, A., and H. Van de Sompel, (2006). Journal Status, Scientometrics, 69(3), p. 669-687.

Bordons, M., Fernández, T. and I. Gomez, (2002), "Advantages and limitations in the use of impact factor measures for the assessment of research performance", Scientometrics, 53(2), p.195-206.

Braun, T., and A. Schubert, (1996), "Indicators or research output in the sciences from 5 central European countries, 1990-1994", scientometrics, 36(2) p. 14- 165.

Brin. B. and L. Page, (1998) "The anatomy of a large-scale hyper-textual web search engine," Computer Networks and ISDN Systems, 30, p. 107-117.

Combes, P-P. and L. Linnemer (2010), "Inferring missing citations. A quantitative multi-criteria ranking of all journals in economics". *GREQAM Working Article*.

Coupé, T., (2003), "Revealed performances: worldwide rankings of economists and economics departments, 1990–2000", Journal of the European Economic Association, 1(6), p. 1309–1345.

Diamond, A., (1986), "What is a Citation Worth?", The Journal of Human Resources , 21(2), p. 200-215.

Dusansky, R. and C. Vernon, (1998), "Rankings of U.S. Economics Departments", Journal of Economic Perspectives, 12(1), p. 157-170.

Egghe, L. (2006), "Theory and practise of the g-index", *Scientometrics*, vol.69, n° 1, p. 133.

Ellison, G. (2013). "How Does the Market Use Citation Data? The Hirsch Index in Economics.", American Economic Journal: Applied Economics, 5(3), p. 63-90.

Faria, J., (2010), "Most Cited Articles Published in Brazilian Journals of Economics: Google Scholar Rankings," Economia, ANPEC - Associação Nacional dos Centros de Pósgraduação em Economia [Brazilian Association of Graduate Programs in Economics], 11(1), p. 1_25.

González-Pereira, B., Guerrero-Bote, V. P., and F., Moya-Anegón, (2009). « The SJR indicator: A new indicator of journals' scientific prestige", Journal of Informetrics, 4(3), p. 379–391.

Gans, J., and G., Shepherd, (1994), "How Are the Mighty Fallen: Rejected Classic Articles by Leading Economists", Journal of Economic Perspectives, 8(1), p. 165-179.

Hamermesh, D., Johnson, G. and B. Weisbrod, (1982), "Scholarship, Citations and Salaries: Economic Rewards in Economics", Southern Economic Journal, 49, p. 472-481.

Hirsch, J., (2005), "An index to quantify an individual's scientific research output", *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 102(46), p. 165-169.

Kim, H., Morse, A. and L. Zingales. (2006), "What Has Mattered to Economics Since 1970", Journal of Economic Perspectives, 20(4), p. 189–202.

Laband D., and M. Piette, (1994). – « The Relative Impacts of Economics Journals: 1970- 1990 », *Journal of Economic Literature*, 32 (2), p. 640-66.

Liebowitz, S., and Palmer, P., (1984), "Assessing the Relative Impacts of Economic Journals," Journal of Economic Literature, American Economic Association, 22(1), p. 77-88, March.

Levitt M., (2015). "What is the optimal number of researchers for social science research?" Scientometrics, 102, p. 213-225.

Maliniak, D., Powers R., and B. Walter (2013), "The Gender Citation Gap in International Relations", International Organization, 67(04), p. 1 - 34

Medoff M., (1989), “The Ranking of Economists”, The Journal of Economic Education, 20(4), p. 405-415.

Medoff M., (1996), “A Citation-Based Analysis of Economists and Economics Programs”, The American Economist, 40(1), p. 46-59 .

Merton, R., (1968), « The Matthew Effect », Science, 159, p. 56-63.

Moore, W., Newman R. and G. Turnbull. (1998). “Do Academic Salaries Decline with Seniority?” Journal of Labor Economics, 16(2), p. 352-366.

Oswald, A., (2007), “An Examination of the Reliability of Prestigious Scholarly Journals: Evidence and Implications for Decision-Makers”, Economica, 74(293), p. 21–31, February 2007

Palacios-Huerta, I. and O., Volij, (2004), “The measurement of intellectual influence,” Econometrica, 72, p. 963-977.

Pinski, G. and F. Narin (1976). “Citation influence for journal aggregates of scientific publications”, *Theory, with applications to the literature of physics*. *Information Processing and Management* 12, p. 297–312.

Ritzberger, K. (2008) “A Ranking of Journals in Economics and Related,” German Economic Review 9(4), p. 402–430

Rossiter, M., (1993). “The Mathilda effect in science”, Social Studies of Science, 23, p. 325–341.

Tan T. and S. Werlang (1992) , “On Aumann’s notion of common knowledge: An alternative approach”, Revista Brasileira de Economia, 46(2), p. 151-166.

Van Noorden, R., Maher, B., and R. Nuzzo, (2014), “The top 100 articles.”, Nature, 514(7524), p.550–553.

Waltman, L., and N., Van Eck, (2010), “The relation between Eigenfactor, audience factor, and influence weight”, Journal of the American Society for Information Science and Technology, p. 1476–1486.