Journal of Informetrics 10 (2016) 552–566

Contents lists available at ScienceDirect

Journal of Informetrics

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

A theoretical evaluation of Hirsch-type bibliometric indicators confronted with extreme self-citation

Gabriel-Alexandru Vîiu

Department of Political Science, National School of Political and Administrative Studies, Bucharest, Romania

ARTICLE INFO

Article history:
Received 29 February 2016
Received in revised form 12 April 2016
Accepted 14 April 2016
Available online 3 May 2016

Keywords:
h-Index variants
Bibliometric indicators
Self-citation
Manipulation
Research evaluation

ABSTRACT

The paper investigates the theoretical response of *h*-type bibliometric indicators developed over the past decade when faced with the problem of manipulation through self-citation practices. An extreme self-citation scenario is used to test the theoretical resistance of the research performance metrics to strategic manipulation and to determine the magnitude of the impact that self-citations may induce on the indicators. The original *h*-index, eighteen selected variants, as well as traditional bibliometric indicators are considered. The results of the theoretical study indicate that while all indicators are vulnerable to manipulation, some of the *h*-index variants are more susceptible to the influence of strategic behavior than others: elite set indicators prove more resilient than the original *h* while other variants, including most of those directly derived from the *h*-index, are shown to be less robust. Variants that take into account time constraints prove to be especially useful for detecting potential manipulation. As a practical tool which may aid further studies, the article offers a collection of functions to compute the *h*-index and several of its variants in the *R* language and environment for statistical computing.

1. Introduction

In the decade that has passed since its impromptu inception into the research evaluation landscape Hirsch's *h*-index (Hirsch, 2005) has become an outstandingly popular bibliometric indicator, albeit one generally relied on by those who are not bibliometric specialists (Bornmann, 2014). Despite its popular success – largely attributable to the simplicity of its calculation (Costas & Bordons, 2008) – the notable flaws associated with this indicator have sparked what seems to have become a virtually unstoppable autocatalytic process in which new *h*-index variants are constantly being proposed in order to correct the perceived defects of the previous ones. Only five years since Hirsch's original proposal comprehensive comparative analyses were already reviewing tens of *h*-index variants (e.g.: Schreiber, 2010; Bornmann, Mutz, Hug, & Daniel, 2011) and recent additions continue to increase an ever growing list of spin-off proposals (e.g.: Zhai, Yan, & Zhu, 2014 propose an h_l-index; Crispo, 2015 proposes an AP-index; Yaminfirooz & Gholinia, 2015 propose a multiple *h*-index; Bar-Ilan & Levene, 2015 suggest the h_w-rank, a variant for ranking web pages).

While the availability of these new research metrics has steadily risen throughout the past decade it should be borne in mind that the utility of any specific performance measure is proportional to its resistance to manipulation (Heinrich & Marschke, 2010). If a performance measure – including a bibliometric indicator – can easily be inflated by some ad hoc process than it is less suitable for the purpose of evaluation and certainly inadequate as a single evaluation instrument.

E-mail address: gabriel.viiu@yahoo.com

http://dx.doi.org/10.1016/j.joi.2016.04.010
1751-1577/

Performance and accountability measures – of which bibliometric indicators may be regarded as a particular class – create strong incentives for strategic behavior and the proliferation of dishonest coping mechanisms at the level of individuals as well as institutions (Trow, 1994; Bruijn, 2007). A related problem is of course the intuitive fact that simpler performance measures are also easier to manipulate than more complex ones. Despite these shortcomings, for those less knowledgeable of their flaws and limitations – as well as for those willing to easily dismiss them – simple measures such as the *h*-index may hold the (questionable) promise of a facile research evaluation panacea. Since its creation the *h*-index has rapidly expanded from a purely academic object known to only a few scholars to a practical and fully functional online impact assessment instrument incorporated into the major online databases of Thompson-Reuters, Elsevier and Google or available through third-party software such as *Publish or Perish* (Harzing, 2007) or *Scholarometer* (Kaur et al., 2012). Decision makers seeking cost effective evaluation instruments are far from unlikely to resort to these instruments despite their potential pitfalls.

As the use of the *h*-index and its variants expands, so too will efforts to manipulate them in the hope of securing competitive advantages such as research funding. The experience of performance-based systems indicates that highly competitive evaluation encourages "publication inflation", "game playing" and indulging in a deleterious tradeoff of quality for quantity (Butler, 2003; Geuna & Martin, 2003), all of which become rational strategies to increase chances of success. It is for these reasons that potential strategizing of the *h*-index and of the *h*-type variants remains an important topic of critical discussion.

The present paper is specifically concerned with the ability of the *h*-index and of some of its variants to cope with willful manipulation and aims to provide an analysis of the theoretical response of the original *h*-index and of eighteen of its variants when confronted with a simple but pervasive manipulation strategy involving self-citation practices.

This manipulation strategy – previously also described in Vinkler (2013) where it yields a "successively built-up indicator" – consists in an author perpetually self-citing *all* of his previous publications[1] such that at the level of each paper N he references all of his previous publications up to and including paper $N-1$. Such a strategy based on an extreme self-citation behavior obviously increases the number of total citations received by the author and, as shown by Vinkler, has the consequence that an author's *h*-index becomes readily definable as $N/2$ in the cases where N is even and $(N/2)-0.5$ in the cases where N is odd. The moral of Vinkler's counterfactual device is that self-citation can lead to significant increases in the *h*-index but the size and nature of this effect has so far only been explored in-depth for this particular index.

Since the limited number of studies explicitly devoted until now to the manipulation of the new, *h*-type bibliometric indicators have mainly focused on the original *h*-index (e.g.: Schreiber, 2007; Zhivotovsky & Krutovsky, 2008; Huang & Lin, 2011; Bartneck & Kokkelmans, 2011) the potential effects of manipulation on the *h*-index variants have received substantially less attention. The present paper aims to address this issue by presenting a detailed comparative analysis of the effects of manipulation not only on the original *h*-index but also on many of its more prominent variants (see Section 2.2 for details). This entails confronting the *h*-index and its selected variants with the sequential self-citation strategy in order to test their response and behavior.

It should be mentioned from the outset that while most studies concerned with the *h*-index and its variants usually compare a limited number of indicators based on samples of scientists and evaluate these indicators based on *empirical* data, the approach followed in the subsequent pages is markedly different: instead of selectively collecting factual, empirical data in the more usual vein of bibliometric case studies, the data that inform the present study are entirely *theoretically derived* under the assumption that the manipulation scenario of self-citation is in operation. The paper therefore relies only on hypothetical data derived under the sequential self-citation strategy (see Section 2.3) and its main objective is to study the individual response of each *h*-index variant as well as determine the change that this specific manipulative practice may induce on the indicators.

2. Materials and methods: the *h*-index, its variants and extreme self-citation

2.1. The h-index

The original statement of the *h*-index is that "a scientist has index *h* if *h* of his or her N_p papers have at least *h* citations each and the other (N_p-h) papers have $\leq h$ citations each" (Hirsch, 2005; p. 16569). The *h*-index is therefore a simple measure which has an inherent appeal because it combines in an intuitive manner the two fundamental building blocks of bibliometrics (papers and citations) in such a way that it accounts not only for an aspect of quantity (captured by the number of published items), but also for one of quality (captured by the number of citations). Some of the generally acknowledged advantages of the *h*-index include the following: it is mathematically simple and a better index than its constituent parts (publications and citations) taken separately (Rousseau, 2008; Norris & Oppenheim, 2010), it is a cumulative indicator which rewards consistent performance (Glänzel, 2006), it is robust with respect to errors in citation records and avoids false precision (Vanclay, 2007), it may be applied not only at the level of individual authors but also for higher-order collective entities such as research groups, institutions or journals (Braun et al., 2006; Prathap, 2006; Schubert, 2007). The *h*-index

[1] This theoretical possibility had been suggested in much earlier work on self-citation: noting a conceptual distinction between synchronous (i.e. *given*) and diachronous (i.e. *received*) self-citations Lawani (1982) mentioned that while improbable, a 100% diachronous self-citation rate is conceivable and would be the hallmark of egotism.

also seems to be reasonably correlated with peer review assessment and with alternative bibliometric indicators (Van Raan, 2006; Bornmann & Daniel, 2007).

In antithesis with its advantages, the *h*-index is mired by several shortcomings. Some of these are non-specific ones shared with other, more traditional bibliometric indicators, while others are unique and pertain only to the *h*-index itself. The following issues mentioned in Egghe (2010) can affect the *h*-index but are equally likely to adversely impact other bibliometric indicators: field-dependence which prohibits meaningful comparison across different fields of study; inadequate treatment of self-citations and of multi-authored papers; strong reliance on the particular database used to retrieve the primary publication and citation information; failure to account for differences in age (especially relevant when comparing multiple authors). Among the shortcomings particular to the *h*-index itself two are considered more relevant in the context of the present paper: first, the *h*-index is defined in a completely arbitrary manner that defies any theoretical grounding (van Eck and Waltman, 2008), a feature which has undoubtedly contributed to the proliferation of *h*-index variants; second, its robustness is a double-edged sword in the sense that the index is insensitive not only to lowly cited papers, but also to very highly cited papers, a feature that unduly penalizes remarkable research (Egghe, 2006b).

2.2. Overview of selected *h*-index variants

Efforts to address both classes of deficiencies found in the *h*-index have resulted in a large number of alternative variants which either directly rely on the original concept and seek to improve it or depart from the initial method and propose entirely different working principles. From the class of *h*-dependent variants the following have been selected for analysis in the present study: the $h(2)$-index (Kosmulski, 2006), the $\hbar$-index (Miller, 2006), the *A*-index and *R*-index (Jin, Liang, Rousseau, & Egghe, 2007), the h_T-index (Anderson, Hankin, & Killworth, 2008), the h_w-index (Egghe & Rousseau, 2008), the *m*-index (Bornmann, Mutz, & Daniel, 2008), the *e*-index (Zhang, 2009), the *hg*-index (Alonso, Cabrerizo, Herrera-Viedma, & Herrera, 2010) and the q^2-index (Cabrerizo, Alonso, Herrera-Viedma, & Herrera, 2010). Three other indicators that rely on the original *h* but also take into account time constraints are the *m*-parameter (Hirsch, 2005), the *AR*-index (Jin et al., 2007) and the h^c-index (Sidiropoulos, Katsaros, & Manolopoulos, 2007). These too are included in the present study but they require additional assumptions (see Section 2.3) for the extreme self-citation scenario. From the class of *h*-independent variants the following have been selected for inclusion in the present study: the *g*-index (Egghe, 2006a), the *f*-index and *t*-index (Tol, 2009), the π-index (Vinkler, 2009), and the *w*-index (Wu, 2010). Table 1 gives a synopsis of the mathematical formulae that describe each individual indicator. Note that while the selection of the indicators is meant to reflect the diversity of the *h*-index variants and the more representative principles which have guided the construction of these alternatives, two classes of variants have been consciously omitted: (1) the class dealing with the issue of multiple authorship whose exploration would entail more elaborate assumptions than the ones needed to characterize the single author self-citation scenario employed in the present study, and (2) the class encompassing variants concerned with field normalization. Variants which differ from the *h*-index only in the sense of being mathematically calibrated to yield more discriminating results, for example the rational *h*-index (Ruane & Tol, 2008) or the real-valued *h*-index (Guns & Rousseau, 2009), have also been omitted.

Previous comparative studies which have approached the *h*-index and its variants from an empirical perspective (e.g.: Bornmann et al., 2008; Schreiber, Malesios, & Psarakis, 2012) have argued that the derived indicators may usefully be divided (although not necessarily in a mutually exclusive manner) in two broad categories: indicators that tend to underscore the *quality* (*impact*) of the scientific output being assessed and, on the other hand, indicators that tend to underscore the *quantity* (*productive dimension*) of the scientific output. In both cases the quantity and quality of core publications – i.e. the papers that contribute to the initial index, whether *h*, *g*, etc. – play a crucial role, but no clear consensus has emerged regarding the integral assignment of the indicators to one of the two categories. Note also that the generic label "*h*-type indicators" used in this paper to refer to the variants listed in Table 1 is employed mostly for convenience and should not obscure the fact that several of the indicators are not *h*-type in the pure sense of being directly derived from the initial *h*-index or in the sense of directly relying on adjacent concepts such as the *h*-core. Rather, the designation "*h*-type indicator" is only meant to highlight the fact that the variants discussed were created in response to Hirsch's original proposal and emerged as alternatives to this specific indicator.

2.3. The extreme self-citation scenario and data

To test the behavior of the *h*-index variants in response to manipulation through self-citation practices a counterfactual scenario involving sequential self-citation is used in the present study. This scenario is premised on the idea that self-citation is the simplest form of strategic behavior that may be adopted in connection to bibliometric indicators. In part, the use of this strategy may be helped by the inherent duality of self-citations: in theory they are a way of manipulating citation rates but in practice they are also a common and reasonable part of the scientific process (Garfield, 1979), often invoked for legitimate reasons such as expanding earlier work or appealing to previously published data or methods (Pichappan & Sarasvady, 2002).

Empirical studies dealing with self-citation rates have found significant variation in self-citation practices across disciplines. Aksnes (2003) found that self-citation rates in Norway varied from 17% in clinical medicine to 31% in chemistry and astrophysics while Glänzel and Thijs (2004) found that mathematics, natural sciences, engineering and life sciences have higher self-citation rates than social sciences and humanities. Both of these studies argue that self-citation is a serious

Table 1
Definition and mathematical representation of *h*-index and selected *h* variants.

Index	Formula	Details
h-Index (Hirsch, 2005)	$\max_r c_i \geq r$	The *h*-index is the highest rank *r* such that *r* papers have at least *r* citations.
m-Parameter (Hirsch, 2005)	$\frac{\max_r c_i \geq r}{a}$ a = years since first publication of author	The *m*-parameter is obtained by dividing an author's *h*-index with his academic age, defined as the number of years since his first publication.
h(2)-index (Kosmulski, 2006)	$\max_r c_i \geq r^2$	The *h*(2)-index is the highest rank *r* such that *r* papers have at least r^2 citations.
ħ-Index (Miller, 2006)	$\sqrt{\frac{1}{2}\sum_{i=1}^{r} c_i}$	The *ħ-index* is the square root of half of the total number of citations accumulated by all of the papers of an author.
g-Index (Egghe, 2006a)	$\max_r \sum_{i=1}^{r} c_i \geq r^2$	The *g*-index is the highest rank such that the first *r* papers cumulate at least r^2 citations.
A-Index (Jin et al., 2007)	$\frac{\sum_{i=1}^{\max_r c_i \geq r} c_i}{\max_r c_i \geq r}$	The *A*-index represents the average number of citations received by the publications included in the *h*-core.
R-Index (Jin et al., 2007)	$\sqrt{\sum_{i=1}^{\max_r c_i \geq r} c_i}$	The *R*-index is the square root of the number of citations corresponding to the papers included in the *h*-core.
AR-Index (Jin et al., 2007)	$\sqrt{\sum_{i=1}^{\max_r c_i \geq r} \frac{c_i}{a_i}}$ a_i = age of publication *i*	The *AR*-index implies calculating for each of the papers determined to belong to the *h*-core the ratios between the citations received and the publication's age; summing these values and taking the square root of the sum yields an index which has the formal ability to decrease (unlike the original *h* which never diminishes, even if authors become inactive).
h^c-Index (Sidiropoulos et al., 2007)	$\max_r S^c(i) \geq r$ $S^c(i) = \gamma \times (Y(\text{now}) - Y(i) + 1)^{-\delta} \times \lvert C(i) \rvert$, *Y(i)* = publication year of article *i* *C(i)* = articles citing article *i* $\gamma = 4; \delta = 1$	Each publication is assigned a particular score (S^c) based on its age, number of citations and two parameters γ and δ that express the weights attributed to publications of different ages. $\gamma = 4$ and $\delta = 1$ imply citations to an article published in the current year are four times as valuable as the ones to articles published four years ago. Once the individual scores for all of the publications are determined "a researcher has contemporary *h*-index h^c, if h^c of its N_p articles get a score of $S^c(i) \geq h^c$ each, and the rest$(N_p - h^c)$ articles get a score of $S^c(i) \leq h^c$" (Sidiropoulos et al., 2007; p. 258).

Table 1 (*Continued*)

Index	Formula	Details
h_T-Index (Anderson et al., 2008)	$\sum_{i=1}^{r} h_{T(r)}$ $\begin{cases} h_{T(r)} = \frac{c_i}{2r-1}, c_i \le r \\ h_{T(r)} = \frac{r}{2r-1} + \sum_{i=r+1}^{c_i} \frac{1}{2i-1}, c_i > r \end{cases}$	The tapered h-index takes into account all the citations received by an author's papers but weighs them differentially according to whether or not they correspond to papers that make up the h-core.
h_w-Index (Egghe and Rousseau, 2008)	$\sqrt{\sum_{i=1}^{r_0} c_i}, r_0 = \max_r \frac{\sum_{i=1}^{r} c_i}{\max_r c_i \ge r} \le c_i$	The h_w-index is defined as the square root of the total citations received by a subset (r_0) of papers in the h-core; r_0 is determined with the aid of weighted ranks defined as the successive divisions of the cumulative citations of the papers by the h-index: the highest rank where a paper's individual citations are greater than (or at least equal to) its weighted rank becomes the threshold value that determines r_0.
m-Index (Bornmann et al., 2008)	$\text{median}\left[\bigcup_{i=1}^{\max_r c_i \ge r} c_i\right]$	The m-index is a variation on the A-index that emphasizes the median number of citations received by papers in the h-core.
e-Index (Zhang, 2009)	$\sqrt{\sum_{i=1}^{\max c_i \ge r} c_i - (\max_r c_i \ge \mathrm{r})^2}$	Emphasizes the contribution of the excess citations of the papers in the h-core and is meant to complement the h-index by capturing the information lost by h in all the cases in which at least one of the papers that constitute the h-core has at least one additional citation not reflected by the value of h.
f-Index (Tol, 2009)	$\max_r \frac{1}{\frac{1}{r}\sum_{i=1}^{r} \frac{1}{c_i}} \ge r$	The f-index is the largest r for which the harmonic average of the number of citations is at least r.
t-Index (Tol, 2009)	$\max_r \prod_{i=1}^{r} c_i^{1/r} \ge r$	The t-index is the largest r for which the geometric average of the number of citations is at least r.
π-Index (Vinkler, 2009)	$\frac{\sum_{i=1}^{r_\pi} c_i}{100}, r_\pi = \text{round}\left(\sqrt{N}\right)$	π represents the one hundredth part of the citations cumulated over the elite set of publications; the elite set is the (rounded value of the) square root of the number of papers.
hg-Index (Alonso et al., 2010)	$\sqrt{\max_r c_i \ge r \times \max_r \sum_{i=1}^{r} c_i \ge r^2}$	The hg-index is the geometric average of Hirsch's h and Egghe's g.
q^2-Index (Cabrerizo et al., 2010)	$\sqrt{\max_r c_i \ge r \times \text{median}\left[\bigcup_{i=1}^{\max_r c_i \ge r} c_i\right]}$	The q^2-index is the geometric average of h and the m-index.
w-Index (Wu, 2010)	$\max_r c_i \ge 10r$	w is the highest number of papers r that have each received at least $10r$ or more citations.

Note: As a prerequisite for calculating the indicators the papers of an author must be sorted in decreasing order of their number of citations (c_i); r represents the rank of an author's papers determined by this sorting process; in practice r is a sequence that always starts at 1 and progresses incrementally up to the level of the last (i.e. least cited) paper of the author.

concern which should be addressed when conducting evaluation at the individual level of authors while for large collective entities such as countries self-citations need not be excluded since they are unlikely to produce biased results.

The formal assumptions that define the extreme self-citation scenario used in the current paper – which focuses on the individual level of authors, not on collective units of assessment such as institutions or countries – are the following:

A1) Within each new publication P_n an author cites all of his previous publications P_{n-1}, P_{n-2}, etc.;

A2) All the publications are single-authored papers (multiple authorship is excluded for convenience but see Section 3.4 for further remarks);

A3) The self-citing strategy is employed consistently starting with the earliest papers and continuing up to the level of the most recent one;

A4) None of the papers receive any foreign citations;

A5) The author's productivity rate – φ – is assumed to be a positive constant;

A6) The time elapsed between each successive pair of publications – θ – is also assumed to be constant.

The first four assumptions of the self-citation scenario suffice to test the response of the selected h-index variants that do not incorporate aspects related to time. The 5th and 6th assumptions of the scenario are required specifically in order to render estimable the three variants that do account for time-related aspects.

To allow the computation of the h-index and its variants a specific scenario involving a finite set of 100 papers is used. This is an admittedly arbitrary choice since 50 or 200 papers could have also been selected but opting for 100 papers has the advantages of providing a modicum of plausibility in relation to real life publication output and of also making statistical treatment of the data possible. A fictitious author conforming to the self-citation scenario just described therefore starts his career with paper P_1 and ends it with paper P_{100}. For each paper P_n of his 100 published papers we may determine an *active citation contribution*, i.e. a number of papers that P_n can cite, and a *passive citation contribution*, i.e. the number of papers that can cite P_n. The active citation contribution of a paper P_n is $n-1$ which simply means that a paper can cite all other papers preceding it (but not itself). The passive citation contribution of any paper P_n is $N-P_n$ (N being the total number of papers), meaning a paper can only be cited in the publications that succeed it. It is of course obvious that older papers have a greater passive contribution (P_1 for example will contribute only passively), while more recent ones have a greater active contribution (P_{100} for example will only contribute in this manner).[2]

At the level of each paper P_n in the self-citing sequence, the list of publications sorted in decreasing order of the number of citations will have the form $\{P_{n-1}, P_{n-2}, \ldots, P_{n-n}\}$. For example, at the level of P_5 the set of publications sorted by their citations will be {4, 3, 2, 1, 0}. This set can then be used to compute the author's bibliometric indicators. At the level of any individual paper in the finite set of 100 the total citation count – TC(P_n) – accumulated by our hypothetical self-citing author will be given by Eq. (1):

$$\mathrm{TC}(P_n) = \sum_{i=1}^{P_{n-1}} i \tag{1}$$

which simplifies to:

$$\mathrm{TC}(P_n) = \frac{n(n-1)}{2} \tag{2}$$

When paper P_1 is published it necessarily has 0 citations; when paper P_2 is published it contributes an active citation towards P_1 and itself becomes eligible to be cited in all future papers; the third published paper contributes a citation to P_2 and one citation to P_1 (therefore raising TC(P_3) to 3); as the subsequent papers are published up to the level of the one hundredth paper P_1 will reach 99 citations, while P_2 will reach 98 citations, P_3 97 citations, etc. TC(P_{100}) reaches an impressive value of 4950. The essential aspect to note is that at the level of each new published paper the self-citing author's portfolio of publications is reflected by *changing values of the bibliometric indicators used to assess his performance*. Both traditional bibliometric measures – number of published items, total number of citations, average citations per paper – and the h-index and its variants have the potential to change with each new publication. By recording the changing values and analyzing their evolution we may uncover the overall trend and the extent to which each indicator is responsive to the manipulative practice of self-citation. Note also that the impact of self-citation on the h-index and its variants may be usefully interpreted by taking into account axiomatic properties of these indicators. In particular, the sequential self-citation test proposed above may be viewed as a logical extension of Woeginger's (2008, p. 365) axiom E which states that "by adding a strong new publication and consistently improving the citations to one's old publications, one also should raise one's index". Woeginger argues that the h and g-index respect this axiom but it is precisely this otherwise desirable and intuitive property that allows the extreme self-citation scenario to unfold.

[2] Note however a further intriguing possibility: an author publishing two papers (say x and y) in separate journals at roughly the same time may be able to reference the preprint version of one paper (say y) in the other (x), thus allowing both the active contribution of one paper (in this case x) and the passive contribution of the other paper (in this case y) to rise above the theoretical thresholds mentioned above. This situation may be contrasted with one in which publication progresses completely linearly: first x is published and then y; here the thresholds of active and passive citations contributed by each paper would obey the limits outlined above.

Table 2
Descriptive statistics of time-independent h variants, average citations per paper (cit/p) and total citations per paper—TC(P_n).

	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
π	0.27	0.86	1.56	2.23	3.26	4.47	5.26	6.77	7.74	9.45
w	0.90	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00
$h(2)$	2.00	4.00	5.00	5.00	6.00	7.00	7.00	8.00	9.00	9.00
e	3.80	7.27	10.74	14.21	17.67	21.14	24.60	28.07	31.53	35.00
$\hbar$	5.19	10.15	15.10	20.05	25.00	29.95	34.90	39.85	44.80	49.75
h	5.00	10.00	15.00	20.00	25.00	30.00	35.00	40.00	45.00	50.00
h_w	5.80	10.84	16.48	22.11	27.74	33.36	38.60	44.06	49.68	55.31
hg	5.87	11.40	17.19	23.06	28.72	34.38	40.25	46.04	51.57	57.45
r	6.28	12.37	18.45	24.53	30.62	36.70	42.78	48.87	54.95	61.03
q^2	6.28	12.37	18.45	24.53	30.62	36.70	42.78	48.87	54.95	61.03
f	6.00	12.00	18.70	24.60	31.00	37.40	43.30	50.00	56.10	62.00
t	6.00	12.80	19.00	25.60	32.00	38.40	45.00	51.20	58.00	64.00
g	6.90	13.00	19.70	26.60	33.00	39.40	46.30	53.00	59.10	66.00
h_T	7.20	14.06	20.93	27.79	34.65	41.52	48.38	55.24	62.10	68.97
a	7.90	15.30	22.70	30.10	37.50	44.90	52.30	59.70	67.10	74.50
m	7.90	15.30	22.70	30.10	37.50	44.90	52.30	59.70	67.10	74.50
cit/p	4.95	9.90	14.85	19.80	24.75	29.70	34.65	39.60	44.55	49.50
TC(P_n)	54	206	456	804	1250	1794	2436	3176	4014	4950

To generate the data that reflect the changes in index values that occur with the addition of each new paper under the sequential self-citation scenario specific functions to calculate each of the selected h-index variants were written in the *R* language and environment for statistical computing (R Core Team, 2015). These functions are provided in Supplementary material 1 and may be used (bearing in mind the technical notes provided) to calculate the h-index and its variants for any alternative dataset.[3] Since the author is unaware of any similar resource that has been made publicly available, this collection of functions may itself prove to be a novel and useful tool to aid further bibliometric studies.

For the purposes of the present paper the collection of functions written in the *R* software was used to compute the values of the h-index and of its variants at the level of each of the 100 papers selected to represent the finite set of publications produced by a hypothetically self-citing author. This entailed writing and executing a loop that expresses the linear expansion of papers and citations under the extreme self-citation scenario, a process also undertaken in the *R* software. The results are the index values computed starting from paper P_1 and ending with paper P_{100} which represent the theoretically-derived data that inform the present study. They are presented in their entirety in Supplementary material 2.

Before moving to the results section some further qualifications are needed with regard to the calculation of the three h-index variants that take into account time constraints. In order to facilitate the calculation of these time-dependent h-index variants – i.e. the m-parameter, *AR*-index and h^c-index – additional assumptions are necessary with regard to the productivity rate (φ) of the self-citing author and also with regard to the time elapsed between each successive pair of publications (θ). Two distinct sets of such assumptions were considered in order to permit a comparative analysis: a more prolific scenario in which φ was set to 4 papers per year and θ to 3 months and a less prolific scenario in which φ was set to only 2 papers per year and θ to 6 months.

3. Results and discussion

3.1. Effects of manipulation on time-independent indicators

The distribution of the progressive values of each time-independent h-index variant in the theoretically-derived dataset is presented in Table 2 which presents the distribution by deciles in order to allow a more granular comparison of the indicators. Given the fact that the dataset has exactly 100 cases the partitioning by deciles has the added advantage of offering not only concise, but also intuitive information regarding the progression of each index value. Note also that all of the indicators naturally start at 0 and this is the minimum value recorded in the dataset for all h variants as well as for the average citations per paper. As a complement to the table, Fig. 1 presents the distribution of the h-index variants using the more familiar notion of quartiles illustrated with the aid of boxplots. One may immediately remark two striking features: first, the elite set indicators – π, w and $h(2)$ – show much less dispersion than all other indicators included in the study; second, there are only incremental differences between the indicators and in some cases no differences at all. For example, the *A* and m indicators exhibit an identical distribution, as do *R* and q^2 and, finally, the h-index, Miller's $\hbar$ and average citations per paper.

[3] Note that this is applicable for all h-index variants with the exception of the three which take into account time (i.e. m-parameter, *AR*-index and h^c-index). These three require much more elaborate input data and therefore more comprehensive algorithmic modelling. The functions written for these three indicators under the specific self-citation scenario and under some additional assumptions about productivity and time between publications (see next two paragraphs of the main text) are nonetheless included in Supplementary material 1 for purposes of transparency and reproducibility. Note, however, that unlike the other functions they may not be used to process alternative data.

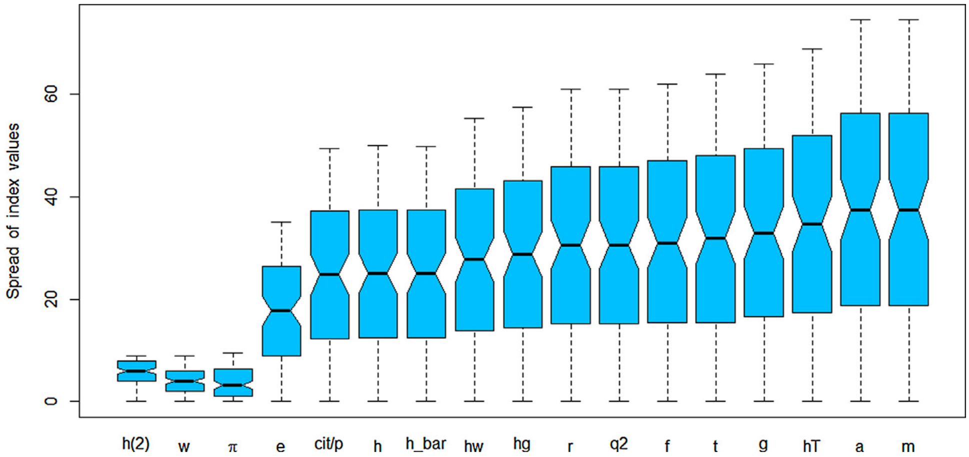

Fig. 1. Comparative distribution of time-independent h-index variants under extreme self-citation (note that to avoid confusion Miller's $\hbar$ has been labeled "h_bar").

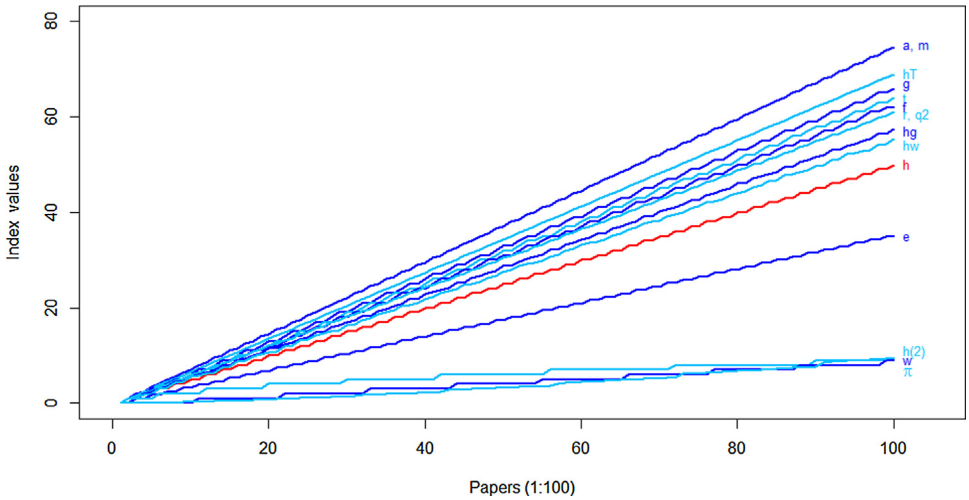

Fig. 2. Progression of h-index and h variants under extreme self-citation involving 100 papers.

A more detailed picture of the impact that the sequential self-citation strategy has on the indicators is given in Fig. 2 which presents the evolution of each indicator at the individual level of all of the 100 papers that constitute the self-citing sequence. An initial inspection of Table 2 and Figs. 1 and 2 reveal that the manipulative strategy of extreme self-citation influences the progression of all of the indicators under study, regardless of their underlying principles or mathematical distinctiveness. Despite their distinguishing features, all of the h-index variants exhibit a discouragingly unanimous receptiveness to the simple manipulation strategy of sequential self-citation. Starting from the first paper in the sequence and progressing towards the last one we may note that all of the indicators gradually increase in value, generally in a linear manner that almost perfectly reflects the linear input of papers and citations inherent in the extreme self-citation scenario.

The distinctive mathematical structure of the indicators insures however that there is noticeable variety with regard to the rate of response in the face of manipulation. Several indicators are much more robust than their alternatives, meaning that although they are not immune they are nonetheless far slower to respond to the self-citation subterfuge. In particular, all of the three elite set indicators – π, w and $h(2)$ – exhibit substantially more resilience to manipulation than the original h-index. They are joined by the e-index which is also influenced by self-citation in a milder manner, however, it should be borne in mind that e is not a stand-alone indicator, but one explicitly designed to be presented in tandem with h. This means that the formal robustness that this indicator displays when considered in isolation may be undermined by the h-index which it is in fact meant to complement. In contrast to elite set indicators we find several h-index variants that are more forcefully affected by extreme self-citation. These are indicators which place greater emphasis on citations in the h-core – for example A and m – or which were designed to place greater weight on citations in general, notably the h_T-index and the g-index. The t, f, R, q^2, hg and h_w indicators are also – to varying degrees – more heavily influenced by self-citation than the h-index.

A somewhat peculiar (while nonetheless consistent) feature of the data generated under the extreme self-citation scenario is the fact that under the assumptions of this counterfactual device some of the indicators effectively collapse into one another and become undiscernible: because extreme self-citation causes average citations per paper to be identical to median citations per paper, the *A*-index and *m*-index that these measures support come to have identical values. Similarly, the *R* and q^2 indicators also take identical values along the entire self-citing sequence of 100 papers because under extreme self-citation the sum of citations in the *h*-core (the central element of the *R*-index) is equivalent to the product between the *h*-index and the median value of the citations in the *h*-core (a product which is central to the calculation of the q^2-index). Extreme self-citation therefore renders obsolete the nuances of the non-parametric components of the *m* and q^2 indicators and reduces them to the antecedent *A* and *R*.

With regard to the *h*-index itself we may note that its value increases with unity (beyond the initial paper P_1 which only contributes passively) for every two additional papers in the self-citing sequence. A more interesting finding is that the progression of the *h*-index under the sequential self-citation strategy is virtually indistinguishable from the progression of a more classical bibliometric index—the average citations per paper: at the level of each odd-numbered paper in the self-citing sequence the *h*-index and average citations per paper indicator have identical values while for even-numbered papers in the sequence the value of the *h*-index is always higher than the one of average citations per paper by precisely 0.5. This fact reinforces the previous claims of Bornmann, Mutz, Daniel, Wallon, & Ledin, 2009 according to which the *h*-index and its variants are not strictly necessary for research evaluation purposes in addition to standard bibliometric indicators. Miller's $\hbar$ also progresses at a rate which is nearly indistinguishable from that of the original *h*-index: for all odd-numbered papers in the self-citing sequence (except for the first one where both indicators are 0) the values of the *h*-index are (on average) approximately 0.25 points smaller than those of $\hbar$ while for the even-numbered papers the values of the *h*-index are (on average) approximately 0.25 points greater than $\hbar$. Because of the almost perfect overlap of average citations per paper and Miller's $\hbar$ with the *h*-index Fig. 2 omits the former indicators as they are essentially masked by the line corresponding to the *h*-index itself.

To determine the particular magnitude of the change elicited by manipulation on the individual indicators linear regression analysis was employed to establish the slope parameter for successive models in which each indicator was treated as the response variable and the self-citing sequence of papers was repeatedly used as the predictor variable. Table 3 presents the results for each regression model and highlights the significant differences between the indicators. In addition to the results presented in this table alternative models incorporating total citations per paper (as well as an interaction term between $TC(P_n)$ and the self-citing sequence of papers) were also considered. However, the alternative models added little to the initial analysis and therefore only the results of the more parsimonious models originally considered are presented. Note also that prior to running the regression analyses individual tests (one-sample Kolmogorov–Smirnov) were conducted to assess the normality of the underlying data. The results of these tests are presented as an Appendix A to the article. While most indicators conform to the normality assumption, in some instances – *h*(2), *m*-parameter and h^c – this assumption is infringed and the corresponding results of the regression models may therefore be less reliable. (For this reason the regression analyses involving time-dependent *h* variants are omitted altogether in Section 3.2.)

The smallest slope encountered among the successive regression models is the one that relates the paper count to the *h*(2) index: increasing the number of papers in the self-citing sequence by one only seems to increase this index by 0.079. Similarly reduced slope parameters obtain for the other two elite set indicators—*w* (slope of 0.091) and π (0.098). In the case of the *e*-index the slope more than triples in comparison to that of the elite set indicators, reaching a value of 0.354. For the *h*-index, $\hbar$, h_W and *hg* the independent papers variable yields slope parameters ranging between 0.50 and 0.58, while for *R* (and implicitly q^2), for *f*, *t*, *g* and h_T slope parameters ranging between 0.61 and 0.69 are obtained. Finally, the steepest slope parameters were obtained for the *A*-index (and implicitly for the *m*-index with which *A* shares identical values): raising the number of papers in the self-citing sequence by one has the effect of raising the *A*-index by no less than 0.75 which makes this index the most vulnerable of all the variants studied.

3.2. Effects of manipulation on variants subjected to time constraints

Table 4 presents the distribution by deciles of the progressive values of the three *h*-index variants which incorporate time constraints. As mentioned in Section 2.3, two separate scenarios were tested: one in which the hypothetical self-citing author produces 2 papers per year at regular intervals of 6 months and a further variation in which 4 papers are produced per year at regular intervals of 3 months. As a preliminary finding we may note, again, that all of these indicators also succumb to the manipulative self-citation practice. Fig. 3 illustrates the precise progression of each indicator.

Starting with the third paper in the self-citing sequence Hirsch's parameter *m* stabilizes around a value of 1 when φ is 2 and θ is 6 and around a value of 2 when φ is 4 and θ is 3. For every odd-numbered paper in the sequence (except for the very first one) the *m*-parameter is in fact exactly 1 (for $\varphi = 2$, $\theta = 6$) or 2 (for $\varphi = 4$, $\theta = 3$) while for the even-numbered papers it takes slightly higher values than 1 or 2 which incrementally decrease as the self-citing sequence progresses towards its final element. It is important to stress that these values, as remarked by Hirsch in his original paper, would indicate at the very least a successful scientist ($m = 1$) or even an outstanding one ($m = 2$) "likely to be found only at the top universities or major research laboratories" (Hirsch, 2005; p. 16571). The fact that values such as these may be obtained artificially indicates to a certain extent that in the absence of proper peer-review it is a priori unfeasible to draw clear boundaries between truly exceptional research and exceptional perseverance in strategizing.

Table 3
Intercept, slope and R^2 estimates for regression models that predict the values of time-independent *h* variants from the number of published papers under the extreme self-citation scenario.

Response variable	Model parameter	Estimate	Std. error of estimate	*t* value (for intercept and slope)/*F* value (for R^2)	*p* value of parameters
h(2)-Index	Intercept	1.825	0.102	17.920	0.000
	Slope	0.079	0.002	44.830	0.000
	R^2	0.954	0.506	2010.000	0.000
w-Index	Intercept	−0.435	0.059	−7.408	0.000
	Slope	0.091	0.001	89.829	0.000
	R^2	0.988	0.291	8069.000	0.000
π-Index	Intercept	−1.205	0.095	−12.73	0.000
	Slope	0.098	0.002	60.48	0.000
	R^2	0.974	0.470	3658.000	0.000
e-Index	Intercept	−0.217	0.037	−5.820	0.000
	Slope	0.354	0.001	552.950	0.000
	R^2	1.000	0.185	305800.000	0.000
h-Index	Intercept	−0.258	0.051	−5.062	0.000
	Slope	0.500	0.001	571.776	0.000
	R^2	1.000	0.253	326900.000	0.000
h-Bar-index	Intercept	−0.268	0.005	−54.500	0.000
	Slope	0.500	0.000	5921.300	0.000
	R^2	1.000	0.024	35060000.000	0.000
h_w-Index	Intercept	−0.419	0.038	−11.110	0.000
	Slope	0.556	0.001	855.610	0.000
	R^2	1.000	0.188	732100.000	0.000
hg-Index	Intercept	−0.439	0.038	−11.5	0.000
	Slope	0.577	0.001	878.6	0.000
	R^2	1.000	0.190	771900	0.000
R-Index & q^2-index	Intercept	−0.322	0.022	−14.940	0.000
	Slope	0.613	0.000	1655.540	0.000
	R^2	1.000	0.107	2741000.000	0.000
f-Index	Intercept	−0.767	0.059	−12.970	0.000
	Slope	0.631	0.001	621.020	0.000
	R^2	1.000	0.294	385700.000	0.000
t-Index	Intercept	−0.804	0.059	−13.680	0.000
	Slope	0.648	0.001	641.170	0.000
	R^2	1.000	0.292	411100.000	0.000
g-Index	Intercept	−0.660	0.055	−11.970	0.000
	Slope	0.667	0.001	703.420	0.000
	R^2	1.000	0.274	494800.000	0.000
h_T-Index	Intercept	−0.375	0.007	−55.260	0.000
	Slope	0.694	0.000	5951.330	0.000
	R^2	1.000	0.034	35420000.000	0.000
A-Index & *m*-index	Intercept	−0.391	0.026	−14.820	0.000
	Slope	0.750	0.000	1653.560	0.000
	R^2	1.000	0.131	2734000.000	0.000

Table 4
Descriptive statistics of *h*-index variants that account for time.

	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
m-parameter ($\varphi = 2$, $\theta = 6$)	1.00	1.00	1.00	1.00	1.01	1.01	1.02	1.02	1.05	2.00
m-parameter ($\varphi = 4$, $\theta = 3$)	2.00	2.00	2.00	2.00	2.01	2.02	2.03	2.05	2.10	4.00
AR ($\varphi = 2$, $\theta = 6$)	3.16	4.47	5.48	6.32	7.07	7.75	8.37	8.94	9.49	10.00
AR ($\varphi = 4$, $\theta = 3$)	4.47	6.32	7.75	8.94	10.00	10.95	11.83	12.65	13.42	14.14
h^c ($\varphi = 2$, $\theta = 6$)	6.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00
h^c ($\varphi = 4$, $\theta = 3$)	8.00	12.00	14.00	14.00	15.00	15.00	15.00	15.00	15.00	15.00

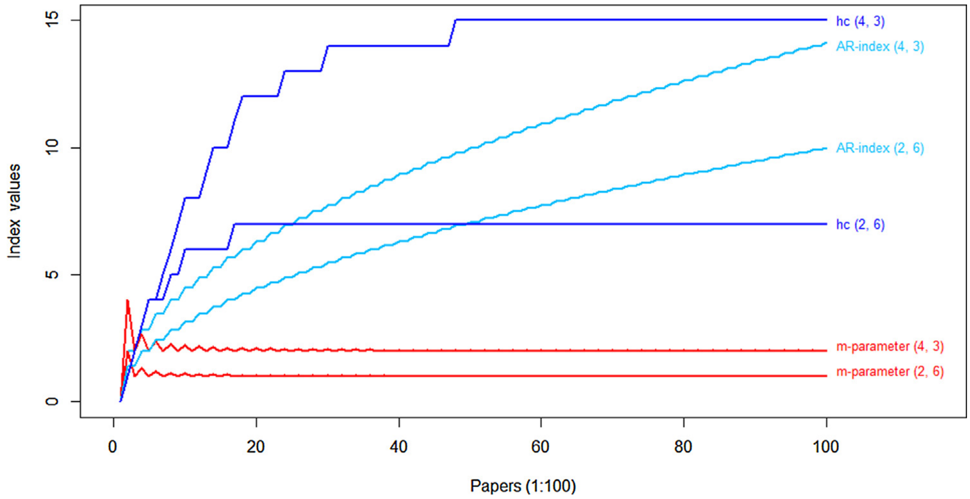

Fig. 3. Progression of *h*-index variants which account for time under the self-citation scenario (note that the parantheses next to the index names designate the productivity and time-between-publications parameters (φ, θ)).

Unlike the *m*-parameter which displays a marginal oscillation in values, both the contemporary *h*-index (h^c) and the *AR*-index turn out to be non-decreasing functions of the papers in the self-citing sequence. This fact is itself a testament to the force and extreme nature of the manipulation scenario envisaged in the present study because both of these indicators (similar to the *m*-parameter) were specifically designed to decrease with time and therefore not allow researchers to "rest on their laurels".

There is however a noticeable difference between the two variants: the *AR*-index increases in a much smoother progression while the h^c-index rises more steeply in the initial stages but eventually levels off at a peak value which can no longer be exceeded. In the less prolific scenario ($\varphi = 2$, $\theta = 6$) the h^c-index peaks at a value of 7 at the level of the 17th paper and maintains this value up to P_{100}. In the more prolific scenario ($\varphi = 4$, $\theta = 3$) the h^c-index peaks at a value of 15 at the level of the 48th paper and maintains this value up to P_{100}. It bears bringing to the attention of the reader the fact that these values are exceptionally high (especially in the latter, more prolific scenario) and that their authentic encounter in real life would indicate a very talented and successful researcher. However, high values naturally invite further scrutiny; careful peer-review initiated by excessively high values of time-dependent *h*-variants therefore has the potential to detect possible manipulative practices.

3.3. Extreme self-citation outside the realm of theory

As noted by Costas, van Leeuwen, and Bordons, 2010 the roles and effects of self-citations on research indicators are an essential topic of discussion in the field of bibliometrics. Several studies devoted to the issue of manipulating the *h*-index (e.g.: Delgado López-Cózar, Robinson-García, & Torres-Salinas, 2014; Labbé, 2010) have emphasized the technological loopholes that may be exploited to literally conjure out of nothingness world-class performance metrics: grand-scale manipulation relying on fictitious papers and equally fictitious citations augment citation counts and all derived measures including *h* and its many variants. The present study warns of an altogether different vulnerability: not instant, grand-scale technical manipulation but patient, perseverant strategizing.

A question that should not go unanswered in this context is whether or not the extreme self-citation scenario depicted in the present study is plausible in real life. Several decades ago when bibliometrics was still an emergent field such gross manipulation seemed too improbable to be of real concern. As an example, Garfield (1979, p. 362) noted in reference to manipulation through successive self-citation that "an abnormally high self-citation count would make the intent so obvious that the technique would be self-defeating". Today, however, at a time when the number of "predatory publishers" is escalating perhaps beyond measure (see for example Butler, 2013; Djuric, 2015), extreme manipulation may well be a real cause for concern.

Competitive pressures, as stated in Section 1, are notorious for galvanizing strategic responses on the part of those facing competition. As noted in the more recent literature on publication practices, "the current system of rewarding research results quantitatively through publication output creates temptations that in some research areas are extremely compelling" (Cutas & Shaw, 2015; p. 1326). The potential problem of extreme self-citation should therefore be taken quite seriously within any reasonable evaluation process. While it is of course highly implausible that the extreme scenario of self-citation would unfold in all its detail in real life, it need not manifest throughout a full hundred papers (as described in the present study) in order to yield superficially appealing performance metrics. Even 20 such papers would already artificially create a not at all negligible *h*-index of 10, supported by no less than 190 citations. In some fields this may be more than enough to secure advantages that might otherwise only be gained through substantial (honest) effort. Therefore, if quantitative, bibliometric tools are to be used by science evaluators or by decision makers engaged in research assessment processes, then the least that should be done to stave off possible manipulation is to make use of a multidimensional approach to evaluation: instead of relying on one indicator, whether it be the *h*-index or any of its variants, several indicators should be used, preferably emphasizing different aspects.

As a final note, it should be mentioned that the manipulation strategy theorized in the present study assumes self-citations are not filtered out in the calculation of the *h*-index (or its variants) and involves a single individual. However, assuming that self-citation filtering is conducted, an analogous (and arguably more egregious) manipulation strategy which can yield the exact same results may take the following form: instead of citing all of his own previous work, an author may enter a closed binary citation network whose strategy is the following: in each of his new publications author A cites all of author B's previously published papers and author B does the same for A. Assuming publication parity – i.e. a near-identical publication frequency for the members of the binary network – the *h*-index of authors A and B progresses in the same manner but formally eschews the problem of *self*-citation by turning it into a more subtle problem of *selective other*-citation. Such a strategy would render ineffective self-citation filtering and even the *h*-index variants specifically designed to mitigate the problem of self-citation such as Kosmulski's (2006) *ch* index, Schreiber's (2009) sharpened index h_s or Ferrara and Romero's (2013) discounted *h*-index *dh*. Note also that the idea of citing in a closed network may naturally be extended from the binary case to the *n*-ary case which immediately raises new issues regarding the quantitative evaluation of collective performance, for example the collective research performance of a university department or of a whole institution.

3.4. Alternative assumptions

Having analyzed the impact that the extreme self-citation scenario may induce on each individual *h*-index variant, it is worth devoting a few thoughts to the likely changes that would ensue if one or more of the assumptions that constitute the scenario would be amended:

- Altering the fundamental assumption of the scenario (A1—perpetual self-citation) would decrease all of the indicators in proportion to the decrease in the number of papers and citations;
- Allowing for multiple authorship (opposed to A2) – while significant for the *h* variants that deal with this aspect – would not alter the variants discussed in the present study;
- Inconsistent strategizing (an alternative to A3) would decrease all of the indicators but the exact magnitude of this reduction would be heavily influenced by the precise nature of the (alternative) pattern of self-citation;
- Inclusion of foreign citations would increase the indicators to varying degrees and would have a more visible impact on the variants that reflect citation counts with greater accuracy, especially the h_T-index;
- Changing the assumptions involving the time parameters φ and θ could either decrease or further increase the values of the three variants accounting for time aspects, depending on the direction of the change: a higher productivity rate and a shorter time between successive publications would further increase the indicators while an opposite shift would reduce them.

4. Summary and conclusions

The main objective of this paper has been to study the individual response of some of the more prominent *h*-index variants when confronted with a hypothetical manipulation scenario based on extreme self-citation. This scenario has been constructed using several simplifying assumptions regarding the progression of self-citations along a specific sequence of one hundred papers. Additional assumptions regarding productivity and time between successive publications were also specified to allow estimation of *h*-index variants designed to capture time constraints. Using an algorithmic implementation of the mathematical formulae that describe each *h*-index variant, the values of eighteen *h* variants were calculated in the *R* language and environment for statistical computing at the level of each individual paper that constitute the self-citing sequence of 100 publications. These functions are offered as a separate, online material that may aid further research, including studies based on empirical data.

The main findings related to the analysis of the theoretically-constructed dataset that constitutes the object of the present study highlight the fact that all of the *h*-index variants (including more traditional bibliometric indicators) are decidedly responsive to manipulative practices. Nonetheless, while this holds as a general principle, it must be stressed that not all indicators are equally receptive to the possibility of strategizing: elite set indicators which emphasize work of remarkable quality – π, w, $h(2)$ – are more resilient in the face of manipulation while metrics that place greater weight on citation counts – including A, h_T and g – are easier to artificially inflate. The h_T-index is particularly interesting because it demonstrates that technical complexity does not necessarily guarantee imperviousness to simple strategizing. The analysis of time-dependent variants indicates that these indicators may prove especially useful in detecting manipulative publication practices. Because their inherent design implies decreasing values as a function of time, detecting instances of consistently high values for any of these indicators should prompt awareness of potential strategizing and of the need for a more detailed (peer) review.

These remarks point toward one final issue that should always be borne in mind by decision makers during a research assessment process: similar to more general aggregation procedures, the *h*-index and its variants obey the basic principle of converting inputs into outputs. Therefore, the indicators in themselves are "blind" to the quality of the input data much as a voting rule is "blind" to whether or not the individual votes being aggregated into a collective decision have been tainted by fraud. Just as consistent efforts are made to ensure the fairness of votes in collective decisions, so too efforts should be made to ensure the fairness of the input data that underpin bibliometric indicators. This ultimately remains a crucial task that can only be reliably entrusted to well-informed peers.

Author contributions

Gabriel-Alexandru Vîiu: conceived and designed the analysis, collected the data, contributed data or analysis tools, performed the analysis and wrote the paper.

Acknowledgements

Financial support from the National Research Council under grant number PN-II-ID-PCE-2011-3-0746 is gratefully acknowledged. Several comments made to the author through the peer-review process conducted by the *Journal of Informetrics* are also acknowledged for helping to clarify and improve aspects of the paper.

Appendix A. Results of one-sample Kolmogorov–Smirnov tests of normality.

Indicator	Mean	SD	Test statistic D	p-value
π	3.76	2.89	0.112	0.160
w	4.14	2.64	0.111	0.171
$h(2)$	5.79	2.33	0.148	0.025
e	17.66	10.27	0.066	0.776
cit/p	24.75	14.51	0.061	0.850
$\hbar$	24.99	14.51	0.061	0.850
h	25.00	14.51	0.066	0.779
h_w	27.64	16.12	0.067	0.753
hg	28.72	16.75	0.066	0.777
r	30.62	17.77	0.063	0.828
q^2	30.62	17.77	0.063	0.828
f	31.12	18.32	0.067	0.761
t	31.92	18.80	0.069	0.732
g	33.00	19.34	0.066	0.777
h_T	34.65	20.12	0.061	0.851
a	37.49	21.77	0.063	0.827
m	37.49	21.77	0.063	0.827
m-parameter ($\varphi = 2, \theta = 6$)	1.02	0.15	0.438	0.000
m-parameter ($\varphi = 4, \theta = 3$)	2.04	0.30	0.438	0.000
AR ($\varphi = 2, \theta = 6$)	6.66	2.39	0.086	0.457
AR ($\varphi = 4, \theta = 3$)	9.42	3.37	0.086	0.457
h^c ($\varphi = 2, \theta = 6$)	6.58	1.23	0.473	0.000
h^c ($\varphi = 4, \theta = 3$)	13.03	3.49	0.320	0.000

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.joi.2016.04.010.

References

Aksnes, D. W. (2003). A macro study of self-citation. *Scientometrics*, *56*(2), 235–246. http://dx.doi.org/10.1023/A:1021919228368

Alonso, S., Cabrerizo, F. J., Herrera-Viedma, E., & Herrera, F. (2010). *hg*-Index: a new index to characterize the scientific output of researchers based on the *h*- and *g*-indices. *Scientometrics*, *82*(2), 391–400. http://dx.doi.org/10.1007/s11192-009-0047-5

Anderson, T. R., Hankin, R. K. S., & Killworth, P. D. (2008). Beyond the Durfee square: enhancing the *h*-index to score total publication output. *Scientometrics*, *76*(3), 577–588. http://dx.doi.org/10.1007/s11192-007-2071-2

Bar-Ilan, J., & Levene, M. (2015). The hw-rank: an *h*-index variant for ranking web pages. *Scientometrics*, *102*(3), 2247–2253. http://dx.doi.org/10.1007/s11192-014-1477-2

Bartneck, C., & Kokkelmans, S. (2011). Detecting *h*-index manipulation through self-citation analysis. *Scientometrics*, *87*(1), 85–98. http://dx.doi.org/10.1007/s11192-010-0306-5

Bornmann, L. (2014). *H*-Index research in scientometrics: a summary. *Journal of Informetrics*, *8*(3), 749–750. http://dx.doi.org/10.1016/j.joi.2014.07.004

Bornmann, L., & Daniel, H. D. (2007). Convergent validation of peer review decisions using the *h* index: extent of and reasons for type I and type II errors. *Journal of Informetrics*, *1*(3), 204–213. http://dx.doi.org/10.1016/j.joi.2007.01.002

Bornmann, L., Mutz, R., & Daniel, H.-D. (2008). Are there better indices for evaluation purposes than the *h* index? A comparison of nine different variants of the *h* index using data from biomedicine. *Journal of the American Society for Information Science and Technology*, *59*(5), 830–837. http://doi.org/10.1002/asi

Bornmann, L., Mutz, R., Daniel, H.-D., Wallon, G., & Ledin, A. (2009). Are there really two types of *h* index variants? A validation study by using molecular life sciences data. *Research Evaluation*, *18*(3), 185–190. http://dx.doi.org/10.3152/095820209X466883

Bornmann, L., Mutz, R., Hug, S. E., & Daniel, H. D. (2011). A multilevel meta-analysis of studies reporting correlations between the *h* index and 37 different *h* index variants. *Journal of Informetrics*, *5*(3), 346–359. http://dx.doi.org/10.1016/j.joi.2011.01.006
Braun, T., Glänzel, W., & Schubert, A. (2006). A Hirsch-type index for journals. *Scientometrics*, *69*(1), 169–173. http://dx.doi.org/10.1007/s11192-006-0147-4
Bruijn, H. (2007). *Managing Performance in the Public Sector*. New York: Routledge.
Butler, L. (2003). Explaining Australia's increased shared of ISI publications—the effects of a funding formula based on publications counts. *Research Policy*, *32*(1), 143–155.
Butler, D. (2013). Investigating journals: the dark side of publishing. *Nature*, *495*(7442), 433–435. http://doi.org/10.1038/495433a
Cabrerizo, F. J., Alonso, S., Herrera-Viedma, E., & Herrera, F. (2010). q^2-Index: quantitative and qualitative evaluation based on the number and impact of papers in the Hirsch core. *Journal of Informetrics*, *4*(1), 23–28. http://dx.doi.org/10.1016/j.joi.2009.06.005
Costas, R., & Bordons, M. (2008). Is *g*-index better than *h*-index? An exploratory study at the individual level. *Scientometrics*, *77*(2), 267–288. http://dx.doi.org/10.1007/s11192-007-1997-0
Costas, R., van Leeuwen, T. N., & Bordons, M. (2010). Self-citations at the meso and individual levels: effects of different calculation methods. *Scientometrics*, *82*(3), 517–537. http://dx.doi.org/10.1007/s11192-010-0187-7
Crispo, E. (2015). A new index to use in conjunction with the *h*-index to account for an author's relative contribution to publications with high impact. *Journal of the Association for Information Science and Technology*, *66*(11), 2381–2383. http://doi.org/10.1002/asi.23426
Cutas, D., & Shaw, D. (2015). Writers blocked: on the wrongs of research co-authorship and some possible strategies for improvement. *Science and Engineering Ethics*, *21*(5), 1315–1329. http://dx.doi.org/10.1007/s11948-014-9606-0
Delgado López-Cózar, E., Robinson-García, N., & Torres-Salinas, D. (2014). The Google scholar experiment: how to index false papers and manipulate bibliometric indicators. *Journal of the Association for Information Science and Technology*, *65*(3), 446–454. http://doi.org/10.1002/asi.23056
Djuric, D. (2015). Penetrating the omerta of predatory publishing: the Romanian connection. *Science and Engineering Ethics*, *21*(1), 183–202. http://dx.doi.org/10.1007/s11948-014-9521-4
Egghe, L. (2006a). An improvement of the *h*-index: the *g*-index. *ISSI Newsletter*, *2*(1), 8–9.
Egghe, L. (2006b). Theory and practise of the g-index. *Scientometrics*, *69*(1), 131–152. http://dx.doi.org/10.1007/s11192-006-0144-7
Egghe, L. (2010). The Hirsch index and related impact measures. *Annual Review of Information Science and Technology*, *44*(1), 65–114. http://dx.doi.org/10.1002/aris.2010.1440440109
Egghe, L., & Rousseau, R. (2008). An *h*-index weighted by citation impact. *Information Processing and Management*, *44*(2), 770–780. http://dx.doi.org/10.1016/j.ipm.2007.05.003
Ferrara, E., & Romero, A. E. (2013). Scientific impact evaluation and the effect of self-citations: mitigating the bias by discounting the *h*-index. *Journal of the American Society for Information Science and Technology*, *64*(11), 2332–2339. http://doi.org/10.1002/asi.22976
Garfield, E. (1979). Is citation analysis a legitimate evaluation tool. *Scientometrics*, *1*(4), 359–375. http://doi.org/10.1007/BF02019306
Geuna, A., & Martin, B. (2003). University research evaluation and funding: an international comparison. *Minerva*, *41*, 277–304.
Glänzel, W. (2006). On the opportunities and limitations of the *H*-index. *Science Focus*, (May). Library of Chinese Academy of Sciences. Retrieved from http://eprints.rclis.org/9378/1/H_Index_opprtunities.pdf
Glänzel, W., & Thijs, B. (2004). The influence of author self-citations on bibliometric macro indicators. *Scientometrics*, *59*(3), 281–310. http://doi.org/10.1023/B:SCIE.0000018535.99885.e9
Guns, R., & Rousseau, R. (2009). Real and rational variants of the *h*-index and the *g*-index. *Journal of Informetrics*, *3*(1), 64–71. http://dx.doi.org/10.1016/j.joi.2008.11.004
Harzing, A. W. (2007). *Publish or Perish*.. Retrieved from. http://www.harzing.com/pop.htm
Heinrich, C. J., & Marschke, G. (2010). Incentives and their dynamics in public sector performance management systems. *Journal of Policy Analysis and Management*, *29*(1), 183–208. http://doi.org/10.1002/pam
Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(46), 16569–16572. http://dx.doi.org/10.1073/pnas.0507655102
Huang, M.-H., & Lin, W.-Y. C. (2011). Probing the effect of author self-citations on *h* index: a case study of environmental engineering. *Journal of Information Science*, *37*(5), 453–461. http://dx.doi.org/10.1177/0165551511412027
Jin, B., Liang, L., Rousseau, R., & Egghe, L. (2007). The *R*- and *AR*-indices: complementing the *h*-index. *Chinese Science Bulletin*, *52*(6), 855–863. http://dx.doi.org/10.1007/s11434-007-0145-9
Kaur, J., Hoang, D. T., Sun, X., Possamai, L., JafariAsbagh, M., Patil, S., & Menczer, F. (2012). Scholarometer: a social framework for analyzing impact across disciplines. *PLoS One*, *7*(9), 1–13. http://dx.doi.org/10.1371/journal.pone.0043235
Kosmulski, M. (2006). A new Hirsch-type index saves time and works equally well as the original *h*-index. *ISSI Newsletter*, *2*(3), 4–6. Retrieved from http://sci2s.ugr.es/hindex/pdf/kosmulski2006.pdf
Labbé, C. (2010). Ike Antkare one of the great stars in the scientific firmament. *ISSI Newsletter*, *6*(2), 48–52. Retrieved from http://hal.archives-ouvertes.fr/hal-00713564
Lawani, S. M. (1982). On the heterogeneity and classification of author self-citations. *Journal of the American Society for Information Science*, *33*(5), 281–284.
Miller, C. W. (2006). *Superiority of the*.. Retrieved from. http://arxiv.org/abs/physics/0608183
Norris, M., & Oppenheim, C. (2010). The *h*-index: a broad review of a new bibliometric indicator. *Journal of Documentation*, *66*(5), 681–705. http://dx.doi.org/10.1108/00220411011066790
Pichappan, P., & Sarasvady, S. (2002). The other side of the coin: the intricacies of author self-citations. *Scientometrics*, *54*(2), 285–290. http://dx.doi.org/10.1023/A:1016070029935
Prathap, G. (2006). Hirsch-type indices for ranking institutions' scientific research output. *Current Science*, *91*(11), 1439.
RCoreTeam. (2015). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/
Rousseau, R. (2008). Reflections on recent developments of the *h*-index and *h*-type indices. *Collnet Journal of Scientometrics and Information Management*, *2*(1), 1–8. http://dx.doi.org/10.1080/09737766.2008.10700835
Ruane, F., & Tol, R. S. J. (2008). Rational (successive) *h*-indices: an application to economics in the Republic of Ireland. *Scientometrics*, *75*(2), 395–405. http://dx.doi.org/10.1007/s11192-007-1869-7
Schreiber, M. (2007). Self-citation corrections for the Hirsch index. *European Physics Letters*, *78*(3), 1–6. http://dx.doi.org/10.1209/0295-5075/78/30002
Schreiber, M. (2010). Twenty Hirsch index variants and other indicators giving more or less preference to highly cited papers. *Annalen Der Physik (Berlin)*, *522*(8), 536–554. http://dx.doi.org/10.1002/andp.201000046
Schreiber, M. (2009). The influence of self-citation corrections and the fractionalised counting of multi-authored manuscripts on the Hirsch index. *Annalen Der Physik (Berlin)*, *18*(9), 607–621. http://dx.doi.org/10.1002/andp.200910360
Schreiber, M., Malesios, C. C., & Psarakis, S. (2012). Exploratory factor analysis for the Hirsch index, 17 *h*-type variants, and some traditional bibliometric indicators. *Journal of Informetrics*, *6*(3), 347–358. http://doi.org/10.1016/j.joi.2012.02.001
Schubert, A. (2007). Successive *h*-indices. *Scientometrics*, *70*(1), 201–205. http://dx.doi.org/10.1007/s11192-007-0112-x
Sidiropoulos, A., Katsaros, D., & Manolopoulos, Y. (2007). Generalized Hirsch *h*-index for disclosing latent facts in citation networks. *Scientometrics*, *72*(2), 253–280. http://dx.doi.org/10.1007/s11192-007-1722-z
Tol, R. S. J. (2009). The *h*-index and its alternatives: an application to the 100 most prolific economists. *Scientometrics*, *80*(2), 317–324. http://dx.doi.org/10.1007/s11192-008-2079-7

Trow, M. (1994). Managerialism and the academic profession: the case of England. *Higher Education Policy*, *7*(2), 11–18. http://doi.org/10.1057/hep.1994.13
van Eck, N. J., & Waltman, L. (2008). Generalizing the *h*- and *g*-indices. *Journal of Informetrics*, *2*(4), 263–271. http://dx.doi.org/10.1016/j.joi.2008.09.004
Van Raan, A. F. J. (2006). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics*, *67*(3), 491–502. http://dx.doi.org/10.1556/Scient.67.2006.3.10
Vanclay, J. K. (2007). On the robustness of the *h*-index. *Journal of the American Society for Information Science and Technology*, *58*(10), 1547–1550. http://doi.org/10.1002/asi.2061
Vinkler, P. (2009). The π-index: a new indicator for assessing scientific impact. *Journal of Information Science*, *35*(5), 602–612. http://dx.doi.org/10.1177/0165551509103601
Vinkler, P. (2013). Would it be possible to increase the Hirsch-index, π-index or CDS-index by increasing the number of publications or citations only by unity? *Journal of Informetrics*, *7*(1), 72–83. http://dx.doi.org/10.1016/j.joi.2012.08.001
Woeginger, G. J. (2008). An axiomatic characterization of the Hirsch-index. *Mathematical Social Sciences*, *56*(2), 224–232. http://dx.doi.org/10.1016/j.mathsocsci.2008.03.001
Wu, Q. (2010). The *w*-index: a measure to assess scientific impact by focusing on widely cited papers. *Journal of the Association for Information Science and Technology*, *61*(3), 609–614. http://doi.org/10.1002/asi.21276
Yaminfirooz, M., & Gholinia, H. (2015). Multiple *h*-index: a new scientometric indicator. *Journal of Managerial Psychology*, *33*(3), 547–556. http://dx.doi.org/10.1108/EL-07-2013-0137
Zhai, L., Yan, X., & Zhu, B. (2014). The Hl-index: improvement of *H*-index based on quality of citing papers. *Scientometrics*, *98*(2), 1021–1031. http://dx.doi.org/10.1007/s11192-013-1039-z
Zhang, C. T. (2009). The *e*-index, complementing the *h*-index for excess citations. *PLoS One*, *4*(5), 1–4. http://dx.doi.org/10.1371/journal.pone.0005429
Zhivotovsky, L. A., & Krutovsky, K. V. (2008). Self-citation can inflate *h*-index. *Scientometrics*, *77*(2), 373–375. http://dx.doi.org/10.1007/s11192-006-1716-2