Research Policy 48 (2019) 1513–1522

Contents lists available at ScienceDirect

Research Policy

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

The influence of editorial decisions and the academic network on self-citations and journal impact factors

Allen Wilhite[a], Eric A. Fong[b,*], Seth Wilhite[c]

[a] Department of Economics, University of Alabama in Huntsville, 301 Sparkman Dr., Huntsville, AL, 35899, United States
[b] Department of Management, University of Alabama in Huntsville, 301 Sparkman Dr., Huntsville, AL, 35899, United States
[c] College of Business Technology Support, University of Alabama in Huntsville, 301 Sparkman Dr., Huntsville, AL, 35899, United States

ARTICLE INFO

Keywords:
JIF manipulation
Networks
Academic misconduct
Editorial manipulation
Research integrity

ABSTRACT

There are many means by which editors can inappropriately manipulate journal impact factors, but questions remain as to whether these potentially inappropriate behaviors actually influence these scores to an empirically meaningful degree, and which academic disciplines are most culpable. In this manuscript, we propose a game-theoretic/information-asymmetry model that suggests manipulation is reinforced by a feedback loop that creates incentives for manipulation to spread and for disciplines to specialize in the type of manipulation used. We empirically investigate these hypotheses for four different manipulation strategies; coercive citation, self-serving review articles, editorials, and online queuing. Results show that all four of these techniques are effective, they inflate JIF scores and the h-index, and a significant part of that effect is due to inflated self-citations. We also find journals within disciplines tend to specialize in which technique they most frequently employ. Moreover, we show that disciplines are also interconnected, tied together by a journal cross-discipline content network and that disciplines that share more content also tend to rely more heavily on the same JIF influencing behaviors. Effective policy needs to change the editorial decision calculation by removing the benefits of manipulation; removing self-citations from journal metric calculations drastically reduces those benefits.

1. Introduction

Although their use can be controversial, journal impact factors have become the primary measure of journal quality and most journal rankings are now based on these calculations (Magnus, 2013; Chorus and Waltman, 2016). There are sound reasons to use citations as a quality measure, but there are also problems and one of these is that citation counts can be artificially inflated by a variety of editorial decisions that can increase self-citations or artificially extend the period during which citations are counted (Martin, 2016). In general, scholars disapprove of such manipulation, it is considered to be unethical, and some publishing organizations have adopted guidelines and/or ethics statements trying to stem the practice.[1] Unfortunately, it is unclear whether these opinions and guidelines are having much of an effect. In this study, we investigate four different editorial strategies that have been identified in the literature as some of the more commonly used manipulation techniques. They are coercive citation, review articles, editorials, and online queuing.

Investigating editorial manipulation is confounded by the fact that several of the most commonly identified manipulation strategies can also be legitimate editorial decisions that have nothing to do with manipulation. Editors may decide to increase the number of review articles or publish editorials in their journal with no intention of inflating their self-citations; they want to better serve their subscribers. However, those same mechanisms can be purposively designed to create superfluous self-citations whose only function is to inflate the journal's performance metrics. Consequentially, manipulators can claim to be doing nothing unusual while non-manipulative editorial decisions may be erroneously viewed as inappropriate. In our opinion, this blurring of actions and intent increases the need for these issues to be studied systematically and for policy to be designed to reduce the incentive to cheat.

Following a review of these citation manipulation techniques, this manuscript suggests these editorial decisions can be thought of as a game in which editorial decisions that affect the journal impact factor (JIF) of one journal influences editorial decisions of other journals. This

* Corresponding author.
E-mail addresses: wilhitea@uah.edu (A. Wilhite), eric.fong@uah.edu (E.A. Fong), wilhits@uah.edu (S. Wilhite).

[1] See Editor Ethics 2.0, https://editorethics.uncc.edu/editor-ethics-2-0-code/ (accessed on 22 June 2018) and Editors Joint Policy Statement Regarding "Coercive citations." https://depts.washington.edu/jfqa/submissions/ (accessed on 22 June 2018).

https://doi.org/10.1016/j.respol.2019.03.003
Received 5 July 2018; Received in revised form 4 March 2019; Accepted 6 March 2019
Available online 21 March 2019
0048-7333/

cross-journal influence is expected to be strongest among closely-related journals and, consequently, there should be a recognizable pattern of use, some disciplines more heavily engaged in particular types of manipulation than other disciplines. Eventually, as the use of manipulation becomes more common, it can spill over into other disciplines, particularly disciplines that publish similar content. We investigate both of these possibilities empirically. Regarding within-discipline adoption, the evidence shows that the level of manipulation differs across disciplines and that disciplines seem to specialize; some relying more heavily on one strategy while other disciplines more frequently use another.

We then study the cross-discipline effects by constructing an academic network that maps the level of cross-discipline content in the journals of eighteen different academic disciplines: chemistry, mathematics, physics, computer science, engineering, biology, ecology, medicine, nursing, sociology, psychology, political science, economics, management, accounting, finance, information systems, and marketing. We present evidence to argue that the distribution of editorial decisions across disciplines is not random either; instead the pattern of editorial manipulation decisions suggests that disciplines react to other disciplines following cross-discipline content links identified in the journal-content network.

2. Literature review

Walker et al. (2010) find that number of publications and the JIF of those publications are the two most important elements of an academic's performance review, and that increases the pressure to publish in high-ranking journals (Lawrence, 2003). Journals have responded in kind, for example, Monastersky (2005) identifies JIFs as a mark of prestige among journals, highlighting the pressure felt by editors to raise their scores to help their journals stay relevant.

While there are minor variations in JIF calculations, the basic formula is simple; the number of citations to articles appearing in a journal over some time period (commonly two years) is divided by the number of articles published in that journal (Garfield and Sher, 1963). The resulting quotient is intended to be an objective measure of a particular journal's impact; a measure of how it influences the discipline. Objective or not, the JIF has become the most common measure of a journal's relative importance in its field (Sharma et al., 2014). The impact factor has gained importance among academics and academic institutions for other applications as well (Martin, 2016; Moed and Van Leeuwen, 1995). For example, impact factors are being used to "measure and compare the scientific output of individuals or institutions" and are increasingly used to help make decisions on tenure, promotion, and merit raises (Walker et al., 2010; Kiesslich et al., 2016). Similar statements apply to another journal metric, the h-index, as its calculation also includes self-citations (in the appendix, we empirically investigate the malleability of the h-index).

A critical flaw in the use of impact factors is that they can be manipulated and there are several known methods of JIF inflation (Falagas and Alexiou, 2008). One is the use of editorials, comments, and letters in the journal. By a quirk of the JIF formula, these non-refereed entries are not included in the denominator, but citations to them are added to the numerator. A second approach to manipulation is to publish review articles or retrospectives on topics that are largely covered by the editors' journals. Such reviews tend to have large numbers of citations to the home journal, pumping up the numerator in the JIF calculation. For example, Opthof (2013) found that the *Journal of the American College of Cardiology* (*JACC*) published a review article with 292 references, 272 of which came from articles published in *JACC* within the last two years. A third strategy is to coerce citations (Wilhite and Fong, 2012). Fong and Wilhite (2017) find coercive citation spans the academic spectrum with scholars reporting coercive experiences in science, engineering, medicine, business and the social sciences. A fourth, emerging manipulation strategy, labeled "the 'online queue' stratagem" by Martin (2016), involves the posting of accepted manuscripts online and delaying their printed issue date (Bjork and Solomon, 2013; Tort et al., 2012). This allows a manuscript to accumulate citations before the JIF time-frame begins. Hopp and Hoover (2017) document its use in management journals. To be clear, editorials, review articles, and putting manuscripts online prior to print publication have legitimate academic value; for example, putting manuscripts online allows for faster dissemination of information and both editorials and reviews help further academic discussions and summarize the current state of understanding. However, each has also been identified as a means of JIF manipulation (Martin, 2016).

Herein, we investigate four editorial decisions to measure whether or not such activities noticeably affect a journal's impact factor score and whether the extent of use of these decisions varies across disciplines. We find that when journals within a specific discipline employ manipulation techniques, they tend to use the same techniques as other, closely related journals. In short, the results suggest that disciplines specialize, some rely on one type of decision while other disciplines coalesce around a different decision. In addition, we suggest that the use of these editorial decisions can spill over from one discipline into another when journals that publish cross-discipline content adopt a given behavior. These cross-discipline journals act as a vector that transmits the incentive to inflate citations to other disciplines. These patterns are consistent with the idea that these decisions have a feedback mechanism; as journals' JIFs increase, closely related journals are more likely to adopt decisions that increase their own JIF, which reinforces the incentives of the original journals and raises the incentives for even more journals to manipulate.

In the end, editors have multiple mechanisms through which they can inflate their impact factor scores if they decide, or are pushed, to do so. In the following section, we model the pressure facing editors who are evaluated, in part, by reputation (i.e., the ranking of their journals). Their situation places them in a zero-sum game of rankings that rewards editors who inflate their scores and disadvantages those who play it straight. In any game of strategy, information plays a central role and, in this application, information on editorial decisions is asymmetric; journals are more likely to be aware of editorial decisions that influence the JIF adopted by closely related journals. As a consequence, manipulation encourages the spread of such decisions to a journal's closest competitors, similarly ranked journals in the same discipline. As a strategy spreads, it can eventually cross discipline boundaries as more journals that publish interdisciplinary content adopt manipulative strategies. To test the validity of this manipulation model, we empirically investigate four of the most commonly identified editorial strategies: coercive citation, review articles, editorials and letters, and online queuing (i.e., the number of articles pre-posted on the web).

2.1. Conceptual framework

Journal impact factors are used to establish journal rankings. Rankings, by their very nature, are relative measures; if one journal rises in the rankings, others necessarily fall. Consequently, the incentive to manipulate or to not manipulate is co-dependent; the action of one journal influences other journals. This sort of interdependence can be modeled as a zero-sum game. Consider an editor, or the editorial decision-maker of journal a, as a member of the set of journals in discipline A. Ignoring cross-discipline effects for the moment, label the net returns, r, to manipulation by this journal as $r(m_a)$. These returns depend on the costs and benefits of manipulation accruing to journal a and the manipulation decisions made by all of the other journals in discipline A, labeled m_A. Formally, the net returns to manipulation for journal a are

$$r(m_a) = m_A(b_a, c_a), a \in A \tag{1}$$

The game has two components: (i) the benefit/cost tradeoff (b_a, c_a) and (ii) the strategies adopted by other journals (m_A). The benefits of manipulating citations, b_a, come from the increased impact factor score, the expected rise in rankings, the notoriety, additional subscriptions, and bragging rights that might accompany higher rankings. The costs of such behavior, c_a, includes the unsavory reputation a journal can acquire as a manipulator as they pump up their citations through review articles or self-serving editorials and the potential backlash from authors who dislike being coerced. However, because journal rankings are relational, the net returns to journal a are affected by the manipulation decision of other journals, m_A. Every gain in ranking by one journal comes at a loss in ranking by other journals in this game. Thus, if one editor decides to play it straight and does not succumb to the temptation to manipulate, then his journal may fall in the rankings as other editors inflate their self-citation counts. This increases the pressure on non-manipulating editors to manipulate, or to see their journal's performance metrics decline. Of course, the costs and benefits vary from journal to journal because editors and publishers have different levels of concern about rankings and different levels of aversion to manipulation; still manipulation by a journal raises the net returns for other journals and at some point additional journals will also decide to manipulate.

Information is critical in strategic games because journals only respond to another journal's decisions if they know, or think they know, the other journal's decisions. We argue that this knowledge is likely to be asymmetric across the academic universe because a journal's editorial team is more likely to be aware of the editorial decisions of closely related journals. In some cases, journals may even share editorial members, again a situation that is more likely to involve closely related journals. As more journals adopt a particular behavior, the stigma of that editorial practice declines as it becomes common place. This reduces the cost of manipulation in that discipline and increases net returns, $r(m_a)$. This potential evolution of discipline norms also suggests that different disciplines might specialize and use different types of manipulation. For example, suppose a journal in discipline A decides to add review articles or retrospectives that include many self-citations to inflate their citation count. As a competing journal in the same discipline sees that journal rise in the rankings, they can respond by using any manipulation strategy. However, if self-serving editorials are already being used by closely related journals, that may establish a norm that reduces the stigma of self-citing editorials. Thus, information asymmetry and editorial norms encourage journals to respond in kind; to use review articles to respond to the use of review articles, coercion to counter others' coercing, etc.

Consequently, we do not expect the practice of manipulation to be randomly spread across the academy, nor do we expect the use of any particular type of strategy to be randomly distributed across journals that manipulate. Instead, we expect some disciplines to manipulate more than other disciplines and we also expect journals in the same discipline to gravitate towards the same types of editorial decisions.

We also suggest that there is a predictable fashion by which discipline share a common use of manipulation and to frame the potential cross-discipline similarity, consider the returns to manipulation for a second journal in a different discipline, say discipline B. As above, the net return to journal b depends on the benefits, costs, and the manipulation strategy taken by other journals in discipline B. Eq. (1) becomes $r(m_b) = m_B(b_b, c_b)$, $b \in B$. We are particularly interested in the intersection of the journals in discipline A and the journals in discipline B. This intersection, $A \cap B$, contains the journals carrying cross-discipline content that links the two disciplines. We contend that these links are the vector by which the incentive to manipulate is transmitted between disciplines.

Suppose citation manipulation is commonly practiced by many journals in discipline A, but is not common in B. If disciplines A and B share journals and journals in discipline A are successfully inflating their impact factors, then those journals move up in the rankings of discipline A and in discipline B. This increases the strategic pressure for the editors in both disciplines to engage in manipulation. The larger the intersection, $A \cap B$, more pressure passes through to discipline B. Information on manipulation, editorial overlap, and the convergence of evolving social norms are also more likely to follow this intersection of journals, $A \cap B$. This leads us to posit that manipulation can spread across disciplines and that this spread is also likely to favor a particular type of editorial decision.

3. Methods and results

3.1. The impact of manipulation on journal metrics

The methodology used to evaluate these game-driven editorial decisions is shaped by data availability. The data used here come from three sources. The primary source is Scopus (2018). Scopus collects data on journals including citations, the number of review articles, editorials, the journals primary disciplines, etc. To these data, we add additional measures from SCImajo, (2018), including information on a journal's total number of references per document, the average number of references per article for each journal, the type of publisher and home country, self-citations, and so forth. The third source is Fong and Wilhite (2017) and from their data we collect information on whether or not a particular journal has been identified as one that coerces authors for citations. A complete description of the data appears in Table A.1 of the appendix.

Using these data, we investigate the use, impact, and spread of four potential manipulation strategies: 1) coercive citation (editors urging authors to add superfluous citations to the editor's journal), the use of editorials and letters (which often contain many self-citations), the use of review articles (again potentially containing many self-citations), and online queuing (i.e., the strategic posting of articles on a website prior to their official publication date). These data impose some constraints on the empirical study. For example, the coercive citation data are limited to a single year (2012). Similarly, we limit the use of the online queuing data to a single year (2017) because these data are ephemeral. Once a physical version of an article is published the website data disappear so comparing across years would be misleading. However, data on the use of review articles and editorials have been collected in both 2012 and 2017, which allows us to compare the change in editorial decisions that influence the JIF across two time periods.

To begin, we look at the fundamental question, are these editorial decisions effective means of influencing journal metrics? The main text focuses on the standard, 2-year journal impact factor score, but similar effects on the h-index are reported in the appendix. Since the calculation of both of these journal metrics includes self-citations, we know inflating those citations inflate the metrics. But, is this impact empirically relevant; does the manipulation of self-citations play a statistically significant role in these metrics? Measuring the impact of manipulation on journal metrics is complicated by simultaneity; attributes that might affect a journal's ranking (such as discipline and journal characteristics) also affect the likelihood that a journal's editorial team manipulates (Wilhite and Fong, 2012). To account for the simultaneous nature of these characteristics, we use an instrumental variables (IV) estimator, two-stage least squares (2sls), which allows us to parse out those differences and get a clearer picture of the extent of editorial decisions on impact factors.

The equation measuring the impact of manipulation on the two-year JIF can be represented as (2a):

$$\mathrm{JIF} = \mathrm{a_0} + \mathrm{Da_1} + \mathrm{Ja_2} + \mathrm{a_3M} \tag{2a}$$

The JIF is expected to differ across disciplines represented by D, a vector of dummy variables representing the disciplines included in the study; information systems (IS), accounting (ACC), sociology (SOC), physics (PHY), political science (PSC), psychology (PSY), medicine (MED), nursing (NUR), biology (BIO), chemistry (CHEM), computer science (CS), ecology (ECO), economics (ECN), engineering (ENG), finance (FIN), management (MGT), marketing (MKT), and mathematics (MTH). The vector J includes journal characteristics, the number of references per document, the type of publisher (such as a university, an academic association, a private company, or government sponsored entity) and the country in which the publisher resides. Finally, M is the manipulation strategy being investigated; coercive citation, the use of review articles, editorials, and online queuing. Again, the econometric complication is that manipulation has also been shown to be affected by disciplines, the variables in D, and the journal characteristics in J. To parse out that interaction, we estimate a first stage equation, (2b):

$$M = b_0 + Db_1 + Jb_2 + b_3U \tag{2b}$$

The variables in D and J match the discipline and journals characteristics in Eq. 2a as does the manipulation strategy M. To identify the system, an independent variable that affects manipulation, but not rankings (except through its impact on manipulation, M) is included in Eq. 2b. In this study, this instrument is the number of other documents (Otherdocs) published by a journal. Using definitions established by Scopus (2017), Otherdocs is calculated by subtracting the number of “citable” documents and the number of editorials and letters to the editor from the total number of documents published by a journal (see appendix). The remainder, labeled Otherdocs, includes features like case studies, news items, abstracts, and so forth. Because citations to these other types of documents are not included in the JIF calculation, they do not affect that score; however, journals that include a variety of other printed features are more likely to include reviews, editorials, and so forth. In addition, they also tend to have more complex editorial structures providing additional routes to coercion. As we see below, the significance of this variable in the first stage of estimation suggests it is a viable candidate as an instrument and post-estimation tests available in Table A.2 in the appendix supports this specification as well; this variable successfully identifies the system.

Using the estimated values of M and the residuals from Eq. 2b, the initial equation, 2a, is estimated without the simultaneous effect and the coefficient a_3 captures the impact of editorial decisions on journal impact factors. The results appear in Table 1.

As expected, academic discipline and journal characteristics affect impact factors; in general, medicine and science journals tend to have higher JIFs, the number of references per document is positive and significantly related to JIF scores, and journals with a U.S. publisher tend to be more highly rated.

Of central interest in Table 1 are the estimated coefficients on the various measures of editorial decisions appearing in the first four rows of Table 1. All four editorial decisions (coercion, the use of editorials, review articles, and online queuing) have positive coefficients, significant beyond the 0.01 level. There is little doubt that these editorial decisions are effective, each one significantly inflates the JIF score even after the confounding effects of discipline and other journal characteristics are taken into account. In the appendix, Table A.3 shows similar results for the h-index.

The results in Table 1 demonstrate that citation coercion, adding editorials, adding review articles, and online queuing significantly increase a journal’s impact factor score; however, those results do not necessarily attribute this to increased self-citations. For example, journals that frequently publish review articles or include editorials may just receive more citations from other journals. To probe more deeply into the use of these strategies and their relationship with self-citations, we estimate another set of regressions with self-citations as the dependent variable and the four editorial decisions are included as explanatory variables along with the discipline, journal and publisher characteristics included in Table 1.

Before making this calculation, we recognize that the dependent variable, self-citations, is not bounded from above and is distributed exponentially. Taking the natural log of self-citations yields a linear relationship and OLS is appropriate. We estimate Eq. (3), where the vectors D and J contain the same variables as before and M includes all four of the editorial decisions.

$$\ln(\text{self-citations}) = a_0 + Da_1 + Ja_2 + Ma_3 \tag{3}$$

Table 1
The impact of editorial decisions on impact factor scores.

	Impact factor 2012	Impact factor 2012	Impact factor 2012	Impact factor 2017	Impact factor 2017	Impact factor 2017
Coerced	34.711** (4.416)					
Reviews		0.086** (0.004)		0.106** (0.005)		
Editorials			0.046** (0.002)		0.055** (0.002)	
Online queuing						0.678** (0.076)
IS	−5.600** (1.158)	1.087** (0.290)	1.048** (0.257	1.080** (0.307)	0.932** (0.250)	−0.509 (0.780)
ACC	−4.607** (1.368)	0.376 (0.445)	0.259 (0.395)	0.138 (0.450)	−0.091 (0.366)	0.248 (1.122)
SOC	0.033 (0.458)	0.192 (0.167)	0.258 (0.148)	−0.084 (0.193)	0.039 (0.157)	−0.063 (0.482)
PHY	−0.029 (0.551)	1.331** (0.185)	1.668** (0.161)	1.209** (0.217)	1.344** (0.176)	1.065* (0.5410
PSY	−0.629 (0.541)	0.775** (0.185)	0.841** (0.164)	0.522* (0.212)	0.561** (0.173)	−0.499 (0.543)
MED	1.245** (0.451)	0.582** (0.170)	1.266** (0.145)	0.224 (0.195)	0.969** (0.154)	1.364** (0.467)
NUR	−0.605 (0.640)	0.661** (0.223)	0.791** (0.197)	0.077 (0.256)	0.496* (0.207)	0.934 (0.635)
BIO	1.509** (0.461)	1.375** (0.169)	1.929** (0.148)	0.573** (0.196)	1.263** (0.157)	1.515** (0.481)
CHEM	0.297 (0.535)	1.106** (0.187)	1.551** (0.165)	0.635** (0.214)	1.407** (0.172)	1.462** (0.526)
CS	0.444 (0.517)	1.469** (0.183)	1.475** (0.162)	1.256** (0.211)	1.268** (0.172)	−0.775 (0.574)
ECO	0.100 (0.506)	1.083** (0.178)	1.168** (0.157)	0.789* (0.205)	0.772** (0.167)	0.654 (0.512)
ECN	−1.490* (0.609)	0.678** (0.199)	0.686** (0.176)	0.461* (0.230)	0.373* (0.187)	−0.001 (0.574)
ENG	0.323 (0.486)	1.203** (0.171)	1.391** (0.151)	0.954** (0.196)	1.296** (0.159)	1.067* (0.487)
FIN	−6.186** (1.115)	0.899** (0.247)	0.837** (0.218)	0.774* (0.309)	0.603* (0.252)	0.429 (0.771)
MGT	−3.614** (0.740)	0.538* (0.190)	0.511** (0.168)	0.490* (0.217)	0.216 (0.176)	−0.314 (0.543)
MKT	−12.367** (2.077)	1.001* (0.445)	0.854* (0.395)	0.874 (0.454)	0.337 (0.369)	−0.190 (1.132)
MTH	0.256 (0.507)	1.177** (0.181)	1.185** (0.160)	0.836** (0.208)	0.891** (0.170)	0.388 (0.522)
Refs. per Doc	0.022** (0.002)	0.026** (0.001)	0.032** (0.001)	0.023** (0.001)	0.041** (0.001)	0.055** (0.003)
University	−0.242 (0.345)	0.050 (0.124)	0.125 (0.110)	−0.378* (0.155)	−0.233 (0.126)	−0.543 (0.387)
Academic	−0.615 (0.352)	0.168 (0.120)	0.257* (0.106)	−0.101 (0.149)	0.023 (0.121)	−0.230 (0.373)
Private	−0.344 (0.339)	0.254* (0.116)	0.502** (0.102)	−0.116 (0.143)	0.136 (0.116)	−1.364** (0.400)
US Publisher	−0.526** (0.181)	0.556** (0.042)	0.503** (0.037)	0.555** (0.050)	0.453** (0.041)	0.173 (0.132)
constant	−0.497 (0.559)	−1.409** (0.197)	−1.691** (0.174)	−0.537* (0.233)	−1.151** (0.188)	−1.422* (0.578)
χ^2	491.4	3,704.4	4,717.6	3,157.0	4,795.6	507.6
n	16,617			19,604		

PSC is the omitted reference discipline; ** indicates significance at the 0.01 level and * indicates significance at the 0.05 level.

The coefficients in Eq. (3) were estimated for 2012 and 2017 and the results appear in Table 2.

The top four rows in Table 2 clarify the source of some of this journal metric inflation. Each of these editorial decisions has a positive and significant effect on the number of self-citations and the impact of review articles and editorials is positive in both 2012 and 2017. Combining the results of these two tables we see that Table 1 implies that editorial manipulation is an effective mechanism to raise a journal's ranking metrics and Table 2 indicates that a statistically significant portion of that impact can be attributed to increased self-citations. Journals can, and do, increase their journal ranking metrics by employing actions that inflate self-citations. We do not claim that all editorials, reviews, or manuscripts posted on the web are attempts to manipulate journal metrics, or that all editors manipulate, but the evidence suggests these manipulation techniques are effective methods of citation inflation.

Table 2
Impact of editorial decisions on Self citations.

	Self-cites 2012	Self-cites 2017
Coerced	1.650** (0.064)	
Reviews	0.011** (0.001)	0.011** (0.001)
Editorials	0.014** (0.001)	0.014** (0.001)
Online queuing		0.014** (0.001)
IS	0.520** (0.195)	0.626** (0.168)
ACC	−0.128 (0.298)	−0.196 (0.249)
SOC	−0.157 (0.114)	−0.282** (0.108)
PHY	1.785** (0.125)	1.745** (0.118)
PSY	0.411** (0.126)	0.190 (0.117)
MED	0.706** (0.112)	0.567** (0.104)
NUR	0.252 (0.151)	0.286* (0.142)
BIO	1.206** (0.113)	0.933** (0.107)
CHEM	1.919** (0.127)	1.707** (0.118)
CS	0.693** (0.126)	0.686** (0.117)
ECO	1.104** (0.121)	0.975** (0.113)
ECN	0.117 (0.136)	0.021 (0.127)
ENG	1.253** (0.117)	1.304** (0.108)
FIN	0.286 (0.166)	0.395* (0.167)
MGT	0.391** (0.129)	0.421** (0.119)
MKT	0.778** (0.287)	0.939** (0.244)
MTH	0.455** (0.123)	0.321** (0.115)
Ref per Doc	0.001 (0.001)	0.003** (0.001)
University	0.121 (0.090)	0.210* (0.097)
Academic	0.430** (0.087)	0.551** (0.095)
Private	0.513** (0.084)	0.663** (0.091)
US Publisher	0.218** (0.029)	0.263** (0.028)
Constant	1.234** (0.138)	1.176** (0.137)
N	13,734	15,300
Adj. R^2	0.23	0.21

PSC is the omitted reference discipline; ** indicates significance at the 0.01 level and * indicates significance at the 0.05 level.

3.1.1. Who uses editorial decisions known to influence the JIF

Having established the manipulative capabilities of certain editorial decisions, we look to see if the use of these decisions differs across disciplines as hypothesized by the game-theoretic model. The game-theory model in Eq. (1) does not suggest any particular discipline is more likely to manipulate nor does it assign a particular type of editorial decision to specific disciplines; no discipline is assumed to be more opportunistic or less ethical than another; however, we do suggest that the use of these decisions will be concentrated such that some disciplines will more aggressively utilize some decisions than others. There are three specific mechanisms leading to this hypothesis. First, the strategic nature of the zero-sum game imposes costs (lower rankings) on journals who do not respond when another closely-related journal uses a JIF inflating technique. Second, content and editorial overlap creates asymmetric information; the knowledge that some journal editors are employing behaviors that lead to inflated citations is more likely to be known to editors in closely related journals because editors are more likely to keep up with journal practices in similar journals and because editorial overlap is more common among journals publishing similar content (Davis, 2018). Third, the stigma accompanying a particular type of editorial manipulation may decline if closely related journals use that technique because the practice is seen to be more common among peers. Thus, while we do not assume inherent ethical differences between disciplines, we do hypothesize that the use of certain editorial decisions associated with manipulation (reviews, editorials, or online queuing) and more direct manipulation (i.e., coercive citation) will differ across disciplines. This is a path dependency explanation; those who manipulate are likely to trigger greater manipulation and similar types of manipulation among closely related journals.

To empirically measure the different levels of manipulation across disciplines and across other journal characteristics, we focus on the first-stage empirical results from Eq. 2a. However, since the dependent variables used in 2a are either binary (this journal was named as a coercer or it was not) or count data (the number of review articles, editorials, and pre-posted articles for online queuing) we used logit estimation for coercion and negative binomial regression for the other manipulation techniques.[2] The results appear in Table 3.

As predicted, there are significant differences in the use of editorial decisions that influence the JIF across academic disciplines. To classify these differences we employ the Suits (1984) constrained estimation technique which sorts the disciplines into three categories; those disciplines who coerce citations (or use editorials or review articles or online queue) significantly more than the average use of these strategies, those who manipulate less than average, and those whose level of manipulation is insignificantly different from the average.[3]

Examining Table 3 more closely, consider the use of coercive citations; five disciplines fall into high-coercion category (IS, ACC, FIN, MGT, and MKT) six are in the below average group (SOC, PSC, MED, BIO, ENG, and MTH), and the rest are average coercers. If we consider the use of review articles, the grouping is different. In 2012, eight disciplines use review articles more frequently than average (SOC, PHY, MED, NUR, BIO, CHEM, ECO, and ENG), while six use them less than average (IS, ACC, CS, ECN, FIN, and MTH). Five years later (the 2017 results) these groups are largely intact although there are a few changes, for example, political science has moved from the average group into the high review group while ecology has dropped from the high group to the average group. There are six disciplines in the above average use of editorials (MED, NUR, BIO, CHEM, CS, ECO) and in 2017 that group is intact except for the movement of chemistry into the average group. In both 2012 and 2017, the same three disciplines rely on editorials statistically less than average (ECN, FIN, and MTH). Finally there are three disciplines that post more articles on the web prior to physical publication (IS, PSY, and CS) and four who post less frequently (SOC, MED, NUR, and BIO).

The statistically relevant differences across academic disciplines are consistent with our hypothesis that the use of editorial decisions will not be randomly distributed across the academic universe because the feedback on incentives accompanying the use of these decisions. As some journals utilize editorial decisions that influence their JIF, that

[2] For the 2sls results in Tables 1 and 2 the first stage was estimated using ordinary least squares. While OLS leads to biased estimated coefficients, it generates the correct covariance terms to correct for simultaneity. But, when the focus turns to who uses JIF influencing editorial decisions, we are more interested in unbiased coefficients and standard errors so the nonlinear techniques are preferred.

[3] The Suits procedure allows all disciplines to be included in the regression by constraining the sum of the coefficients to zero (Suits, 1984). The resulting estimated coefficients are then interpreted as how a specific discipline's use of manipulation differs from the average use of manipulation across all disciplines.

Table 3
Who uses certain editorial decisions known to influence the JIF.

	Coerced 2012	Reviews 2012	Editorials 2012	Reviews 2017	Editorials 2017	Online Queue 2017
IS	1.484** (0.254)	−0.385* (0.172)	−0.087 (0.170)	−0.487** (0.153)	−0.172 (0.149)	1.146** (0.334)
ACC	1.054* (0.469)	−0.719* (0.298)	−0.102 (0.290)	−0.327 (0.254)	−0.248 (0.250)	−0.726 (0.557)
SOC	−1.221** (0.171)	0.159** (0.047)	−0.071 (0.047)	0.572** (0.049)	−0.004 (0.047)	−0.214* (0.108)
PHY	0.071 (0.174)	0.608** (0.069)	0.044 (0.070)	0.289** (0.074)	0.081 (0.073)	0.120 (0.171)
PSC	−1.464** (0.553)	0.162 (0.111)	0.185 (0.114)	0.544** (0.110)	0.045 (0.109)	−0.181 (0.256)
PSY	−0.107 (0.174)	0.071 (0.071)	0.017 (0.072)	0.046 (0.070)	0.056 (0.068)	0.521** (0.160)
MED	−1.130** (0.121)	1.218** (0.039)	0.900** (0.039)	1.236** (0.038)	1.005** (0.037)	−0.441** (0.085)
NUR	−0.088 (0.279)	0.511** (0.107)	0.868* (0.101)	0.940** (0.105)	0.980** (0.099)	−0.524* (0.244)
BIO	−0.869** (0.145)	0.876** (0.045)	0.195** (0.046)	0.912** (0.046)	0.281** (0.045)	−0.422** (0.108)
CHEM	−0.215 (0.197)	0.726** (0.071)	0.177* (0.071)	0.817** (0.067)	0.061 (0.066)	−0.287 (0.160)
CS	−0.366 (0.191)	−0.307** (0.070)	0.096* (0.067)	−0.126 (0.070)	0.303** (0.066)	1.057** (0.157)
ECO	−0.245 (0.173)	0.147* (0.062)	0.145* (0.062)	0.042 (0.062)	0.123* (0.061)	−0.033 (0.144)
ECN	0.333 (0.200)	−0.275** (0.089)	−0.548** (0.096)	−0.312** (0.089)	−0.311** (0.090)	0.239 (0.198)
ENG	−0.352* (0.142)	0.153** (0.051)	−0.064 (0.051)	0.482** (0.049)	0.005 (0.048)	0.108 (0.110)
FIN	1.584** (0.199)	−0.465** (0.137)	−0.326* (0.138)	−0.730** (0.159)	−0.541** (0.158)	0.257 (0.337)
MGT	0.898** (0.140)	−0.109 (0.077)	0.109 (0.077)	−0.385** (0.077)	0.014 (0.073)	−0.010 (0.169)
MKT	2.175** (0.351)	−0.471 (0.289)	0.126 (0.284)	−1.581** (0.270)	−0.184 (0.244)	0.352 (0.551)
MTH	−0.542** (0.209)	−0.897** (0.070)	−0.666** (0.070)	−0.934** (0.071)	−0.483** (0.069)	0.039 (0 .149)
Refs per Doc	0.004** (0.001)	0.010** (0.001)	−0.009** (0.001)	0.023** (0.001)	0.003** (0.001)	−0.017** (0.001)
University	1.985* (1.017)	0.055 (0.090)	0.290** (0.094)	0.241* (0.097)	0.571** (0.105)	3.587** (0.480)
Academic	2.572* (1.007)	0.209* (0.086)	0.679** (0.090)	0.273** (0.093)	0.984** (0.101)	3.405** (0.478)
Private	2.653** (1.004)	0.377** (0.083)	0.784** (0.087)	0.327** (0.089)	1.141** (0.098)	5.009** (0.472)
US Publisher	0.808** (0.088)	−0.013 (0.030)	0.088* (0.030)	−0.087** (0.031)	0.191** (0.030)	0.473** (0.081)
Otherdocs	0.001** (0.000)	0.008** (0.000)	0.013** (0.000)	0.007** (0.000)	0.014** (0.000)	0.011** (0.001)
constant	−5.877** (1.005)	0.342** (0.088)	−0.171 (0.092)	−0.210* (0.093)	−1.129** (0.102)	−3.941** (0.475)
χ^2	768.5	2,203.0	2,535.8	4,046.1	2,395.6	1,084.7
n	16,617			19,604		

We use the Suits (1984) constrained estimation technique and thus there is no omitted reference discipline; ** indicates significance at the 0.01 level and * indicates significance at the 0.05 level.

practice alters the calculation of other, closely related journals (journals in the same discipline) making it more likely that they will also utilize editorial decisions that influence their JIF.

Disciplines specialize; for example, the business disciplines rely more heavily on coercing citations relative to other disciplines while the disciplines in the healthcare fields more frequently use reviews and editorials. From a theoretical perspective this suggests that information asymmetry, potential editor overlap, and the evolution of discipline-based editorial norms work hand-in-hand with the game driven incentives. Competition to improve, or at least maintain, a journal's position in journal rankings creates the incentive to engage in some response, and since journals mimic the practice of other journals in their field, information and evolving norms appear to be part of the decision calculus.

The feedback loop between journals also suggests that the lumpiness of editorial decisions could extend beyond specific disciplines such that disciplines that share a greater level of cross-content are more likely to participate in similar types of JIF changing editorial decisions. To test that supposition using the results in Table 3 requires two additional steps: 1) we must map out the cross-discipline content among journals and 2) we need a statistical test to see if the groups identified in Table 3 align with this mapping.

For the first step, we create an academic network that reflects the cross-discipline content in journals. In this network, disciplines are the nodes and the links connecting those nodes are based on the proportion of journals in each discipline publishing articles in another discipline. Journal content is identified using the All Science Journal Classification code (ASJC), as reported in Scopus. The ASJC code is a 4-digit number that represents fields of study published by each journal and many journals have multiple codes indicating that they publish manuscripts from different disciplines. For example, the *Journal of Finance* has codes indicating that they publish manuscripts from the fields of economics, finance, and accounting. We use these cross-discipline journal counts to calculate the weight of each link in the network. Formally, let $|A|$ be the number of journals published in discipline A (the number of journals that list this particular ASJC code) and let $|B|$ be the number of journals published in discipline B. The weight of the edge connecting disciplines A and B is, $\frac{|A \cap B|}{|A \cup B| - |A \cap B|}$, which is the proportion of journals in disciplines A and B that publish articles in both disciplines. Repeating this procedure for all eighteen disciplines in our study we get a mapping of the cross-discipline nature of these fields. In the resulting network, closely related disciplines will have more heavily weighted edges (represented by thicker lines) and more distant disciplines will have edges with less weight (thinner lines). This process yields the journal content network in Fig. 1.

Fig. 1 demonstrates that there is substantial cross-discipline content in academic journals. Because we are interested in the cross-discipline content's potential role in the spread of manipulation we look for clusters of disciplines that share greater content among themselves and share less with other clusters (i.e., content sharing sub-groups). To identify those components, we employ a tabu search procedure that maximizes the value of the weighted links within each cluster and minimizes the weight of the links between clusters. The procedure starts with an arbitrary partition and then moves nodes between clusters in an attempt to maximize the fit criterion. Tabu search reduces the chances that the algorithm will get stuck at a local maximum and repeating the procedure with different initial arbitrary partitions increases the probability of finding the global maximum (Glover, 1989). Specifying the desired number of clusters *a priori*, we can compare different partitions by initially dividing the eighteen disciplines into two, three, four, and five clusters. Those clusters unfold in a reasonable fashion as shown in Table 4.

The two-cluster case separates disciplines with the science, medicine, and engineering in one group and business and the social sciences in the other. When the network is divided into three clusters, the medicine/life science disciplines and social sciences split off from their respective clusters to form their own group. The four-cluster optimization splits the medicine/life sciences and the social sciences and results in clusters consisting of science and engineering, business, medicine, and the social sciences. Five clusters have only a small impact as the business disciplines split into two groups, information systems and marketing forming their own group. Although not shown in the table,

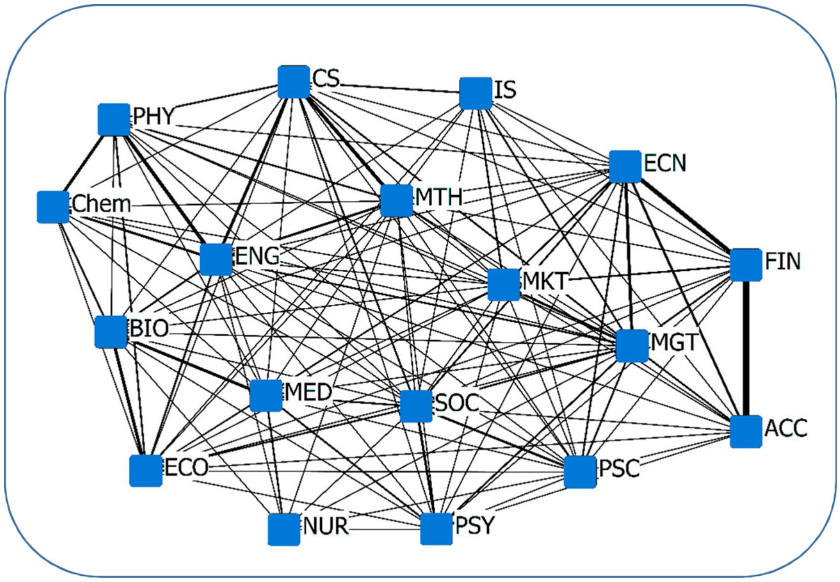

Fig. 1. The journal content network. Heavier links indicate a higher proportion of the journals in these two disciplines share cross-discipline content.

allowing for a sixth cluster simply cuts nursing out of the medical group and puts it into its own cluster and further divisions that add additional clusters continue that process; one discipline after another is split off to become its own "group" adding little additional information.

A visual presentation of the four cluster model can be constructed by removing the weaker links from the network presented in Fig. 1, keeping the stronger within-cluster links. The resulting network appears in Fig. 2.

We now have two different groupings of these academic disciplines. One consists of network clusters produced by the tabu search procedure that identifies which disciplines share the most cross-discipline content in their journals. The other grouping sorted disciplines by their use of editorial decisions that can be used to manipulate their self-citations and journal metrics. These groups are identified by the first-stage regressions results in Table 3. If the game-theoretic framework plays a role in editorial decisions made across disciplines these two groupings should resemble each other.

To rigorously test that matching we use the Freeman and Halton (1951) extension of the Fisher Exact Test. Fisher's exact test is a non-parametric statistic that calculates the probability that the observed

Table 4
Results of the tabu search procedure using two thru six clusters.

	One	Two	Three	Four	Five
Two Clusters	Biology Ecology Physics Mathematics Medicine Nursing Computer Science Engineering Chemistry	Accounting Management Information Systems Marketing Economics Finance Psychology Sociology Political Science			
Three Clusters	Chemistry Mathematics Physics Engineering Computer Science	Management Accounting Finance Economics Marketing Information Systems	Ecology Biology Nursing Psychology Sociology Medicine Political Science		
Four Clusters	Chemistry Mathematics Physics Engineering Computer Science	Management Accounting Finance Economics Marketing Information Systems	Ecology Biology Nursing Medicine	Psychology Sociology Political Science	
Five Clusters	Chemistry Mathematics Physics Engineering Computer Science	Management Accounting Finance Economics	Ecology Biology Nursing Medicine	Psychology Sociology Political Science	Marketing Information Systems

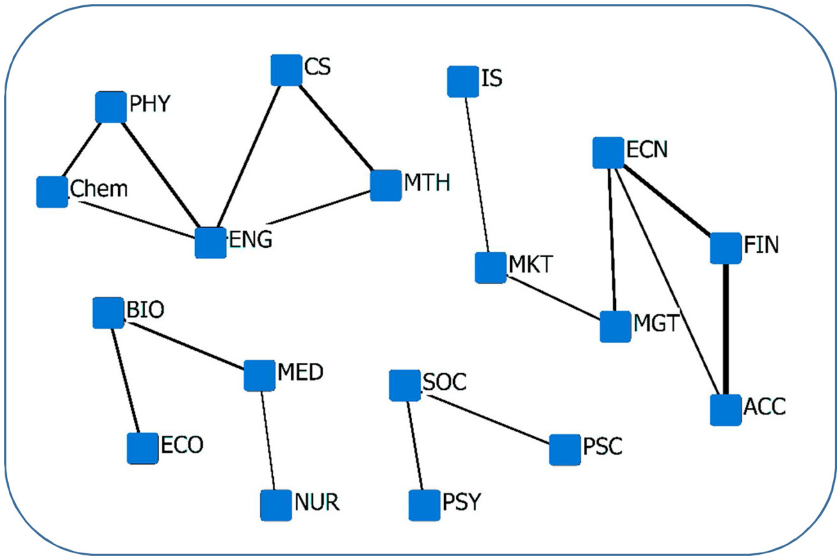

Fig. 2. The four cluster journal content network.

Table 5
Numbers of disciplines in each cluster falling into each type of coercer using Fisher's exact test on the four cluster case.

	Cluster 1 (science & engineering)	Cluster 2 (business)	Cluster 3 (medicine)	Cluster 4 (social sciences)	Totals
High coercers	0	5	0	0	5
Average coercers	3	1	2	1	7
Low coercers	2	0	2	2	6
Totals	5	6	4	3	18

distribution of manipulating disciplines aligns in a statistically significant fashion among the clusters identified by the tabu search procedure (Fisher, 1932). To demonstrate the application of Fisher's exact test we use the four-cluster case, three levels of JIF influencing editorial decisions (high, average, and low), and the coercive citation results from Table 3. We begin by constructing a 4 × 3 cross table with the coercion classification as rows and network clusters as columns. Thus, Table 5 quantifies the network displayed in Fig. 2 and overlays the coercion results from Table 3.

Given the counts in Table 3, the exact test calculates the probability that this particular relationship between disciplines that coerce and the ideological clusters would happen randomly, given the number of clusters and the grouping of coercers. In this example, that probability is:

$$\frac{\binom{5}{0}\binom{5}{3}\binom{6}{5}\binom{1}{1}\binom{4}{0}\binom{4}{2}\binom{3}{0}\binom{3}{1}}{\binom{18}{5}\binom{13}{7}} = 0.000073$$

With a point probability of 0.000073, it is extremely unlikely that the observed pattern of coercive citation across the ideological landscape is random. The implication is that the cross-content network clusters and the level of coercive citation practiced in various disciplines are related, consistent with our hypothesis that the adoption of coercion is influenced by the network links (cross-over journals) that transmit the incentive to manipulate from one discipline to another. Proposing a stricter test, Freeman and Halton (1951) suggest that aggregating the sum of the point probabilities for all other potential combinations of manipulation and clusters that are even less likely to arise than the observed dispersion yields a p-value. In the above example, the p-value for the coercive citation example is 0.017, also significant using the typical 0.05 alpha level.

This test is repeated for each type of editorial decision measured in Table 3 and the results are displayed in Table 6. Focusing on the 4 cluster network, the likelihood that the observed alignment between all four types of editorial decisions and the network clusters is very small, less than 0.0001, and in five of the six cases they also pass the more strenuous p-value test at the typical 0.05 level. The single exception, online queuing, has a p-value of 0.075.

Table 6 also presents the Fisher exact tests for the two, three, and five cluster networks. Again, in most cases the alignment of editorial

Table 6
Fisher exact tests: Significance of match between disciplines that manipulate and disciplines with cross-content journals.

	Test statistic	2 clusters	3 clusters	4 clusters	5 clusters
Coercive citation (2012)	Point prob.	0.0065	0.00014	0.00007	0.000049
	p-value	0.038	0.005	0.017	0.068
Reviews (2012)	Point prob.	0.0025	0.00034	0.00005	0.000026
	p-value	0.014	0.015	0.007	0.020
Editorials (2012)	Point prob.	0.0022	0.0039	0.00011	0.000044
	p-value	0.011	0.184	0.013	0.021
Reviews (2017)	Point prob.	0.021	0.00027	0.00016	0.0001
	p-value	0.138	0.009	0.035	0.118
Editorials (2017)	Point prob.	0.0074	0.0043	0.00012	0.000049
	p-value	0.039	0.147	0.009	0.017
Online queuing (2017)	Point prob.	0.1	0.0028	0.00065	0.00022
	p-value	0.658	0.076	0.075	0.084

The point probability is the probability that the specific contingency table occurs randomly. The p-values are the sum of the probabilities of the observed dispersion and all equally extreme or more extreme dispersions (Fisher, 1932).

decisions and the network clusters are unlikely to be random, there are only five instances where the p-value exceeds the 0.10 level. The weakest fit is online queuing; it never meets the 0.05 alpha level, but it is still significant at the 0.10 level in four of the five network tabu search clusters.

Finally, because we have data on the use of review articles and editorials in two time periods, we can investigate disciplines that change classification from 2012 to 2017. For example, consider the use of review articles; in 2012, political science was in the average group (their use of review articles did not significantly differ from the average use), but in 2017 political science has moved into the above-average use category. In total, there were only seven disciplines that changed categories from 2012 to 2017, six changed their use of review articles and one changed their use of editorials. In these seven instances, we aggregate the link-weights of each specific discipline to other disciplines in their same 2012 category (above average manipulators, average manipulators, or below average) and compare that aggregated weight to that discipline's aggregated link-weights to the disciplines in the other categories. Of the seven changes observed here, five followed their network link-weight, meaning they moved to the group with whom they shared the stronger aggregate link weight. Two disciplines move to a category with a lower aggregate link-weight; they did not follow the network.

Overall the use of reviews and especially editorials has been fairly stable. Only seven disciplines moved categories at all and in every case they moved only one level, in other words, no discipline moved from the low category to the high category or visa-versa. Since reviews and editorials have been published for decades, this stability is not unexpected. In sum the exact test rejects the conclusion that the matching of the use of editorials and review articles with the groups sharing cross-discipline content is random both years, and the bulk of the changes that do arise move disciplines into categories with more highly weighted links, in other words, they follow the network's links.

4. Discussion

Research shows that competition among scientists can contribute to unsavory actions by scholars; examples include falsification of data or results, sabotage, and interference with peer review (Anderson et al., 2007). The increase in the use of JIFs creates similar pressure on editors, it is easy for journals to be compared on these scores and so journals compete for higher scores. This means editors face pressure to raise their JIF scores and because JIFs can be manipulated, some editors run self-serving editorials, coerce citations, add review articles, or use online queuing to artificially inflate their metrics (Martin, 2016). We propose that editors are locked in a zero-sum game and as some journals manipulate their citations, the pressure for others to manipulate rises. Furthermore, because of information asymmetry and the evolution of ethical norms, journals with similar content are likely to adopt similar types of JIF influencing editorial decisions. While there may be other explanations for the pattern of JIF influencing editorial decisions, these results align with the simple, game-theoretic/information asymmetry model offered here.

Before closing, we discuss some shortcomings, possible extensions, and policy implications. First, while it is difficult to defend editors who coerce authors to add superfluous citations to a manuscript before publication, the other types of JIF influencing editorial decisions (review articles, editorials, and online queuing) are reasonable editorial decisions that do not necessarily indicate an attempt to artificially inflate citations. However, numerous examples have been identified in the literature demonstrating that these techniques have been used to inappropriately manipulate (Falagas and Alexiou, 2008; Martin, 2016; Opthof, 2013; Hopp and Hoover, 2017). In this study, we look at this statistically and the strong and positive relationships between reviews, editorials, and online queuing and the improved JIF scores in association with the inflated number of self-citations suggests that they are effective manipulation tools. Nevertheless, just because there are some bad actors out there we should not unilaterally condemn the use of review articles, editorials, and online queuing. Second, in some sense, this may be a late chapter in the story of editorial manipulation. Comparing the use of review articles and editorials from 2012 to 2017 identifies only small adjustments. Since editorials and review articles have been part of the academic literature for decades, we may be near a manipulation equilibrium, at least for those strategies, although some changes are expected as editorial teams change personnel. Future work that reconstructs the evolution of reviews and editorials as artificial inflation strategies could help us understand how we arrived at our current situation.

To the extent that game-theoretic pressure funneled by asymmetric information and evolving editorial norms has led to the use of manipulation strategies, it will be difficult to reign in the manipulators. Manipulation is effective; these techniques significantly inflate a journal's impact factor score, a result that is difficult to ignore. For a single journal's editorial board to reverse an existing policy of manipulation, that board must be prepared to see their outlet fall in the rankings and that fall would be more precipitous in disciplines where manipulation is common. Furthermore, with the cross-discipline effects documented here, a journal that stops manipulating will fall relative to journals in other disciplines as well. That suggests that even if an entire discipline decided to reverse their path and encouraged all of its journals to stop manipulating, impact factors in that discipline will fall in the rankings relative to other, closely related disciplines. It would appear to the academy at large that their discipline is becoming less relevant. This could be a bitter pill to swallow.

5. Conclusions

The previous literature combined with this study paints a gloomy picture; JIF influencing editorial decisions seem to be effective and widespread and their use reinforces additional use and undermines the incentive to stop manipulation. But who cares; what harm is created by journals inflating JIF scores? First, these activities change rankings. Not all journals practice manipulation, nor do the manipulators do so to the same degree. Thus, as journal metrics become increasingly important in academic decision-making and resource allocation, those decisions are flawed. Tenure and promotion, research funding, annual evaluations, and hiring decisions frequently rely on journal metrics and those decisions are distorted. But, distorted rankings aren't the largest concern; manipulation misleads academics. Credit is given where it is not due when false citations are added to manuscripts. Incremental contributions can have inflated reputations and over shadow substantial contributions. In addition, the use of self-serving editorials and self-promoting review articles casts a shadow over all review articles and editorials. These editorial options can make valuable contributions to our understanding of the world, but when some people use these tools to cheat then it taints any contribution as a potential cheat. Finally, the ethical violation should bother us all. Falsely inflating citations strikes at the heart of scientific research, it promotes things that are not true.

Policy can help alter the editorial calculus. First, decision makers who embrace manipulation need to realize the full consequences of their behavior. Editors hold considerable sway in academia, what message do they send when they to commit acts of research misconduct? Junior faculty are in the stage of their career when mentorship most matters; where will the ethical direction of our profession go as more faculty members face these dilemmas at a young age and begin to accept them as the norm? In fact, Seeber et al. (2019) show that assistant and associate professors seem to be strategically using self-citations to increase their own, individual citation metrics. We need better guidelines and clearer ethical standards for editors. For example, Hall and Martin (2019) develop a taxonomy to help academics determine appropriate and inappropriate research behaviors to help shape good research practices; such a guide should exist for editors.

As noted in the introduction, guidelines by themselves are unlikely to be effective without a second step, which is the elimination of self-citations from the impact factor, the h-index, and other journal metric calculations. This simple act erases the benefits of coercing authors, of writing self-serving editorials, and publishing review manuscripts with dozens of citations to the home journal. It reduces the penalty faced by journals that decide not to manipulate by eliminating the feedback loop that continually increases the pressure to manipulate. Ethical editors are not penalized and manipulative editors are not advantaged. Of course, there are alternative impact factor calculations that already exist and some of those omit self-citations, but the presence of a smorgasbord of metrics compounds the problem. Opportunistic journals can pick the factor score that most compliments their preferred type of manipulation; self-citations need to be removed.

However, removing self-citations has only a small impact on the online queuing strategy. Instead, developing policy about a consistent start date would remove the incentive to utilize this as a means of manipulation. For example, if the two-year JIF window begins when manuscripts are made public (when they appear online or in print, whichever comes first) then there is no advantage to strategically delaying the physical version to collect citations. Unfortunately, this might also negatively impact an important, legitimate reason for online posting, which is to disseminate information more quickly. In essence, adopting a consistent start date may encourage some journals with longer publishing backlogs to delay online posting if they perceive that the simultaneous appearance of print and electronic versions makes the biggest impact.

Finally, removing self-citations allows editors to return to making decisions for editorial reasons, not to artificially inflate citations. It also stops us from questioning ethical editors who include reviews or editorials as part of their offerings. It is a simple fix, but it is not an easy one. Impact factor calculations are made by third-parties who may not be concerned with the ethics of the issue and currently the academy does little to sway their decisions. But it can. The academy can identify a slate of transparent citation and impact metrics that minimize the ability to fudge the numbers. We can encourage the universal adoption of those measures, which would eliminate or reduce unwanted incentives. Absent the removal of the incentive to manipulate citations, we should expect manipulation to persist and even grow. Scholars may lament the decline in publishing ethics, but that decline seems to be our destiny; unless we find the resolve to change.

Acknowledgements

This work was supported by a grant from the Office of Research Integrity (ORI) through the United States Department of Health and Human Services (HHS) [Grant number ORIIR130003]. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of ORI or HHS.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.respol.2019.03.003.

References

Anderson, M.S., Ronning, E.A., De Vries, R., Martinson, B.C., 2007. The perverse effects of competition on scientists' work and relationships. Sci. Eng. Ethics 13 (4), 437–461.
Bjork, B., Solomon, D., 2013. The publishing delay in scholarly peer-reviewed journals. J. Infometrics 7 (4), 914–923.
Chorus, C., Waltman, L., 2016. A large-scale analysis of impact factor biased journal self-citations. PLoS One 11 (8), e0161021. https://doi.org/10.1371/journal.pone.0161021.
Davis, P., 2018. Oh, what a tangled web! Citation network underscores editorial conflicts of interest. Scholarly Kitchen 12 (December), 2018.
Falagas, M.E., Alexiou, V.G., 2008. The top-ten in journal impact factor manipulation. Arch. Immunol. Ther. Exp. 56 (4), 223–226.
Fisher, R.A., 1932. Statistical Methods for Research Workers. Oliver & Boyd, Edinburgh.
Fong, E.A., Wilhite, A.W., 2017. Authorship and citation manipulation in academic research. PLoS One 12 (12), e0187394. https://doi.org/10.1371/journal.pone.0187394.
Freeman, G.H., Halton, J.H., 1951. Note on the exact treatment of contingency, goodness of fit and other problems of significance. Biometrika 38 (1/2), 141–149.
Garfield, E., Sher, I.H., 1963. New factors in evaluation of scientific literature through citation indexing. Am. Doc. 14 (8), 195–201.
Glover, F., 1989. Tabu search-part I. ORSA J. Comput. 1 (3), 190–206.
Hall, J., Martin, B.R., 2019. Towards a taxonomy of research misconduct: the case of business school research. Res. Policy 48 (2), 414–427.
Hopp, C., Hoover, G.A., 2017. How prevalent is academic misconduct in management research? J. Bus. Res. 80, 73–81.
Kiesslich, T., Weineck, S.B., Koelblinger, D., 2016. Reasons for journal impact factor changes: influence of changing source items. PLoS One 11 (4), e0154199. https://doi.org/10.1371/journal.pone.0154199.
Lawrence, P.A., 2003. The politics of publication. Nature 422 (6929), 259–261.
Magnus, D., 2013. Overthrowing the tyranny of the journal impact factor. Am. J. Bioeth. 13 (7), 1–2.
Martin, B.R., 2016. Editors' JIF-boosting stratagems – which are appropriate and which are not? Res. Policy 45 (1), 1–7.
Moed, H.F., Van Leeuwen, T.H., 1995. Improving the accuracy of institute for scientific information's journal impact factors. J. Assoc. Inf. Sci. Technol. 46 (6), 461–467.
Monastersky, R., 2005. The number that's devouring science. Chron. High. Educ. 52 (8), A12–A17.
Opthof, T., 2013. Inflation of impact factors by journal self-citation in cardiovascular science. Netherlands Heart J. 21 (4), 163–165.
SCImajo, 2018. SJR – SCImago Journal & Country Rank. http://www.scimagojr.com.
Scopus, 2017. Content Coverage Guide. Updated August, 2017. .
Scopus, 2018. Abstract and Citation Database of Peer-reviewed Literature. Elsevier. https://www.scopus.com.
Seeber, M., Cattaneo, M., Meoli, M., Malighetti, P., 2019. Self-citations as strategic response to the use of metrics for career decisions. Res. Policy 48 (2), 478–491.
Sharma, M., Sarin, A., Gupta, P., Sachdeva, S., Desai, A.V., 2014. Journal impact factor: its use, significance and limitations. World J. Nucl. Med. 13 (2), 146.
Suits, D.B., 1984. Dummy variables: mechanics v. interpretation. Rev. Econ. Stat. 66 (1), 177–180.
Tort, A.B.L., Torgino, Z.H., Amaral, O.B., 2012. Rising publication delays inflate journal impact factors. PLoS One 7 (12), e53374. https://doi.org/10.1371/journal.pone.0053374.
Walker, R.L., Sykes, L., Hemmelgam, B.R., Quan, H., 2010. Authors' opinions on publication in relation to annual performance assessment. BMC Med. Educ. 10, 1–9.
Wilhite, A.W., Fong, E.A., 2012. Coercive citation in academic publishing. Science 335 (6068), 542–543.