Postprint: A Study on the Stable Time Window of the Disruptive Index

Authors: Liu Xiaohui, Shen Zhesi, Liao Yu, Zhu Manman, Yang Liying

Date: 2023-04-01T16:02:55+00:00

Abstract

[Purpose/Significance] To explore the stable time window of the disruptive index, analyze its disciplinary differences, reveal the relationship between the stability of the disruptive index and the citation half-life as well as the stability of citation frequency, and provide a time window reference for the rational application of this index across various disciplines. [Method/Process] We respectively calculate the stability of the disruptive index for 22 disciplines under different citation time windows and obtain the time window required for the disruptive index in each discipline to reach a stability of 0.8; we further analyze the relationship between the disruptive index and the citation half-life, as well as between the disruptive index and the stable time window of citation frequency. [Results/Conclusion] The time window required for the disruptive index to reach a stability of 0.8 exhibits substantial disciplinary differences; at the citation half-life, the stability of the disruptive index across all disciplines reaches 0.8, suggesting that the citation half-life can be used as a reference time window for calculating the disruptive index; both disciplinary characteristics and indicator algorithms are important factors influencing the stability of the disruptive index, and due to algorithmic effects, the disruptive index generally reaches high stability later than citation frequency.

Full Text

Preamble

Research on the Stable Time Window of the Disruption Index

Liu Xiaohui[1],[2], Shen Zhesi[1], Liao Yu[1],[2], Zhu Manman[1],[2], Yang Liying[1]

[1] National Science Library, Chinese Academy of Sciences, Beijing 100190
[2] Department of Library, Information and Archives Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190

Abstract:
[Purpose/Significance] This study explores the stable time window of the disruption index, analyzes its disciplinary differences, and reveals the relationship between the stability of the disruption index and both citation half-life and citation frequency stability, providing temporal reference for the appropriate application of this index across disciplines. [Method/Process] We calculated the stability of the disruption index for 22 disciplines under different citation time windows to determine the time window required for each discipline's disruption index to reach a stability level of 0.8. We then analyzed the relationship between the disruption index and both citation half-life and citation frequency stability. [Results/Conclusion] The time window required for the disruption index to achieve 0.8 stability varies significantly across disciplines. When reaching citation half-life, the disruption index stability in all disciplines exceeds 0.8, suggesting that citation half-life can serve as a reference time window for calculating the disruption index. Both disciplinary characteristics and indicator algorithms are important factors affecting the stability of the disruption index. Due to algorithmic influences, the disruption index generally achieves high stability later than citation frequency.

Keywords: Disruption index; Citation time window; Citation indicator; Stability; Disciplinary differences
Classification Number: G250
DOI: 10.13266/j.issn.0252-3116.2021.18.006

Research evaluation serves as a fundamental tool for managing and assessing scientific activities, playing crucial roles in measuring performance, guiding development, supporting decision-making, and providing diagnostic consultation within the research management system. As noted by Price Medalist A. Raan, citation-based evaluation holds substantial application value. However, practical experience reveals that citation indicators are significantly influenced by citation time windows (hereinafter referred to as "time windows"). For stability, longer time windows are preferable, whereas for timeliness, shorter windows are desired. Therefore, selecting an appropriate time window represents a critical research question in the application of citation indicators.

The disruption index is a relatively new metric for measuring originality, assessing the disruptive nature of scientific achievements by evaluating their replacement of existing research content or paradigms from the perspective of local citation structures. As a citation indicator, the disruption index is also affected by citation time windows. L. Bornmann was among the first to raise this issue, but only examined four papers as case studies, lacking systematic investigation into the time window problem and its disciplinary variations.

This study conducts exploratory research on the time window issue of the disruption index, analyzing its stability under different time windows and examining its relationship with other citation time characteristics to better understand the stability and temporal features of this metric and provide reference for time window selection.

1. Introduction to the Disruption Index

Since its proposal, the disruption index has attracted widespread attention from scientists. It measures the originality of research papers by quantifying the degree to which scientific outputs disrupt existing research content or paradigms. The conceptual framework of the disruption index is illustrated in Figure 1 [Figure 1: see original paper].

In Figure 1, the black triangle represents the research output being evaluated, termed the “focal paper.” The circular nodes are its references, while the square nodes above represent citing papers. Based on different citation patterns toward the focal paper and its references, citing papers are classified into three categories: Type F (black, citing only the focal paper), Type B (dark gray, citing both the focal paper and its references), and Type R (light gray, citing only the focal paper’s references). These three types represent the disruptive, developmental, and continuity characteristics of the focal paper’s relationship to historical knowledge, respectively. The disruption index D is defined as the probability difference between Type F and Type B citing papers, as shown in Formula (1):

$$D = \frac{N_F - N_B}{N_F + N_B + N_R}$$

where N_F, N_B, and N_R denote the numbers of the three types of citing papers, respectively.

2. Research Status on Citation Time Windows

Previous studies have examined citation time windows for various indicators. Researchers such as D. Walters, R. Costas, and J. Levitt investigated annual citation patterns and cumulative citation frequencies over time, finding that papers with different citation levels, disciplines, and document types exhibit distinct citation patterns across time windows. L. Waltman and colleagues discovered that evaluating papers requires at least a full-year time window, while A. Nederhof found that longer time windows yield more reasonable conclusions than shorter ones. R. Costas and G. Abramo analyzed time window selection methods from individual researcher and institutional perspectives, revealing significant impacts on evaluation results. However, A. Jonathan’s study based on UK scientific data found strong correlations between early-stage (approximately one year post-publication) and long-term citation frequencies, suggesting that initial citations can predict long-term impact.

Other studies have focused more on selecting appropriate time windows. G. Abramo et al. quantified the “sufficient” length of time windows and measured errors introduced by shortening them. J. Wang calculated correlation coefficients between citation frequencies at various time windows and at 31 years, finding that field normalization cannot improve accuracy lost through short time windows. X. Wang et al. used correlation coefficients between short-term and long-term citations as weights to compare university rankings before and after weighting the CNCI indicator, finding that weighting produced more reasonable rankings. These studies precisely measured citation frequency stability across time windows and their impact on research evaluation, demonstrating that time window selection is crucial for valid assessment.

Beyond time windows, cited half-life is another commonly used indicator reflecting citation temporal characteristics. Research has shown that publication country, publication cycle, language, journal impact factor, and total citations all affect cited half-life. Studies by He Wen et al. and others have examined relationships between impact factor and cited half-life, suggesting that these relationships should be considered when constructing journal evaluation systems.

As a citation indicator, the disruption index's stability is closely related to time window length. However, since its calculation does not directly use citation frequencies, its stable time window is more complex than that of citation counts. Building on existing research, this study explores the stable time window of the disruption index across disciplines, examines disciplinary differences, and investigates its relationship with citation frequency stability and cited half-life.

3. Research Methods and Data

3.1 Research Methods

In scientometric research, post-stabilization citation frequencies are generally considered representative of a paper's long-term impact. Following existing studies, we calculated the stable time window for the disruption index in each discipline through three steps:

First, we assumed that a paper's citations stabilize by year T (sufficiently long) post-publication, treating T as the reference time window to calculate each discipline's disruption index, denoted as $D_{t=T}$.

Second, we treated 1-year, 2-year, …, (T-1)-year windows as test time windows, calculating the disruption index for each, denoted as $D_{t=1}, D_{t=2}, \ldots, D_{t=T-1}$.

Third, we calculated correlation coefficients between the reference time window and each test time window, using these as stability measures for the corresponding test windows, as shown in Table 1 .

Table 1. Principle for Calculating the Stable Time Window of the Disruption Index

Reference Time Window	Test Time Window	Correlation Coefficient (i.e., Stability)
$D_{t=T}$	$D_{t=1}$	$Corr_1$
	$D_{t=2}$	$Corr_2$
	…	…

Reference Time Window	Test Time Window	Correlation Coefficient (i.e., Stability)
	$D_{t=T-1}$	$Corr_{T-1}$

Based on these steps, we identified the shortest citation time window achieving the desired stability level as the minimum stable time window (hereinafter referred to as “stable time window”).

Since the disruption index is a citation indicator, and both cited half-life and citation frequency stable time windows are fundamental indicators reflecting citation temporal characteristics, and given that the disruption index calculation involves citation frequencies ($N_F + N_B$ represents citation count), this study further examines the relationship between the disruption index and both cited half-life and citation frequency to better understand its stability and temporal features.

3.2 Data Sources

Using Web of Science as the sample database, we selected SCI papers published in 1998 by the 20 most productive countries (including China, the United States, the United Kingdom, Germany, etc.) as our analysis objects, retrieved on October 17, 2020. We used 1-year, 2-year, …, 18-year windows as test time windows for cumulative citation statistics, with 19 years as the reference time window (treating total citations through 2017 as stable citations) to analyze the stable time of the disruption index across disciplines. The sample comprised 396,000 papers with 17.17 million cumulative citations. Given disciplinary differences in citation patterns, we analyzed variations in the stable time window across 22 ESI discipline categories.

4. Stable Time Windows of the Disruption Index Across Disciplines

Based on the above data and methods, we obtained the stability of the disruption index across disciplines at different time windows, as shown in Figure 2 [Figure 2: see original paper].

Figure 2. Stability of the Disruption Index Across Disciplines at Different Time Windows
Note: Legend names correspond to groups in Table 2

In correlation analysis, 0.8 is considered the threshold for strong correlation. Following this convention, we adopted 0.8 as the standard for strong stability of the disruption index. Figure 2 reveals that: (1) The time required to achieve 0.8 stability varies dramatically across disciplines. Visually, this variation warrants further analysis of its causes. (2) At the same time window, substantial differences exist between disciplines. For instance, at a 4-year window, some disciplines (e.g., Molecular Biology & Genetics) exceed 0.8 stability, while others (e.g., Agriculture) remain below 0.7. Uniform application of a 4-year window for the disruption index would thus introduce bias for certain disciplines, necessitating discipline-specific time windows.

According to the time windows required to reach 0.8 stability, disciplines can be roughly divided into seven groups (see Table 2). Molecular Biology & Genetics requires only 2 years, while Agriculture needs 8 years.

Table 2. Time Windows Required for D to Reach 0.8 Stability Across Disciplines

Group	Disciplines	Time Window (Stability = 0.8)
1	Molecular Biology & Genetics	2
2	Computer Science	3
3	Clinical Medicine, Space Science, Biology & Biochemistry, Social Sciences, Neuroscience, Economics & Business, Immunology, Microbiology	4
4	Physics, Psychiatry	5
5	Chemistry, Engineering, Mathematics, Environment/Ecology, Geosciences	6
6	Pharmacology & Toxicology, Plant & Animal Science, Materials Science	7
7	Agriculture	8

5. Analysis of Factors Influencing Disruption Index Stability Trends

Since the disruption index conceptually measures disruption through citing papers and their citation structure characteristics relative to the focal paper, this section analyzes causes for stable time window variations from two perspectives: changes in citation structure across time windows and mathematical operations.

From the citation structure perspective, the disruption index is affected by the stability trends of Type F, Type B, and Type R citations. The relative quantities of these three citation types directly impact disruption index stability. If these three categories stabilize, the disruption index will also stabilize, with the latest-stabilizing category ultimately determining the disruption index's stable time.

Figure 3 [Figure 3: see original paper] illustrates stability trends for N_F, N_B, N_R, and the disruption index D across time windows, using Engineering, Clinical Medicine, Computer Science, and Mathematics as examples. The figure shows similar stability patterns across disciplines for the three citation types, with N_F and N_B exhibiting greater variation and primarily influencing disruption index stability, while N_R remains relatively stable across disciplines with minimal impact. In summary, differences in citing literature stability trends cause disciplinary variations in disruption index stable time windows.

Figure 3. Stability Trends of Type F, B, and R Citing Papers and D

On the other hand, the disruption index's stabilization time is also influenced by its calculation formula. The disruption index D is a function of N_F, N_B, and N_R, which exert secondary influence on stability through this function, directly determining D's stability across time windows. In Formula (2), citation frequency has a linear relationship with N_F and N_B:

$$C = N_F + N_B$$

Once these stabilize, citation frequency stabilizes. However, in Formula (1), the disruption index has a non-linear relationship with N_F, N_B, and N_R, making its stable time more complex. Nevertheless, even if the three components haven't stabilized, the disruption index can still achieve high stability if their relationships remain consistent with those at the 19-year window.

Therefore, if the disruption index achieves high stability before citation frequency, this represents a coincidence under the algorithm—a temporary stability. Generally, for sustained stability, N_F, N_B, and N_R must all stabilize.

Ultimately, disciplinary differences in disruption index stable time windows are determined by research paradigms and citation practices, which manifest as unique citation patterns affecting the stability trends of N_F, N_B, N_R, and consequently D. These influences primarily appear through research duration, review periods, and recognition delays. For instance, experimental disciplines in biomedicine require lengthy and complex experiments, directly affecting publication timing and thus the disruption index's stable time window. Similarly, verifying mathematical theorems may require extensive review time, while recognition of focal papers also involves discipline-specific time lags, creating disciplinary variations in stable time windows.

6. Disruption Index Stability and Cited Half-Life

As a citation indicator, the disruption index is primarily affected by citation patterns. For most papers, citations accumulate rapidly before reaching cited half-life, then slow down and enter a decline phase. This raises the question: when using the disruption index to measure long-term disruptive performance at the cited half-life point, are the results stable and reliable? What stability level can be achieved?

This study treats cumulative citations at the 19-year window as final citation counts, with the publication year as year 1. The time required for cumulative citations to reach 50% of final counts is defined as cited half-life. Table 3 illustrates this calculation method with examples.

Table 3. Illustration of Cited Half-Life Calculation

Paper	1998	1999	2000	2001	...	2017	Half-Life
Paper 1	1	8	20	30	...	40	3.33
Paper 2	1	8	...	...	...	...	0.94

For Paper 1, cumulative citations reach half of the 2017 total between years 3 and 4, yielding a cited half-life of:

$$\text{Paper 1 Half-Life} = 3 + \frac{20 \times 0.5 - 9}{20 - 9} = 3.33$$

Similarly, Paper 2's half-life is 0.94 years. Disciplinary cited half-life is the average of all papers' half-lives in that discipline. Since half-lives aren't necessarily integers, we applied fractional calculations when determining corresponding disruption index stability. For example, Space Science has a half-life of 5.77 years, with disruption index stability values of 0.86 and 0.89 at years 5 and 6, respectively, resulting in a stability of 0.88 at half-life (i.e., $0.86 + (0.89 - 0.86) \times (5.77 - 5)$).

Figure 4. Disciplinary Cited Half-Life and Corresponding Disruption Index Stability
Note: Dashed lines represent median values

Figure 4 [Figure 4: see original paper] shows that: (1) At cited half-life, all disciplines achieve disruption index stability exceeding 0.8, with some surpassing

0.9 (e.g., Economics & Business at 0.92). (2) Many disciplines have similar half-lives but vastly different disruption index stability levels. For instance, Physics and Molecular Biology & Genetics both have half-lives around 5.5 years, but the latter's disruption index stability is significantly higher. This can be explained by the differential stability of N_F, N_B, and N_R discussed in Section 5. Table 4 compares stability levels for these components in both disciplines, showing Physics has lower stability for all three categories, particularly N_F, resulting in lower D stability.

Table 4. Stability of Related Variables in Physics at 5-Year Window

Variable	Stability
N_F	0.75
N_B	0.82
N_R	0.91
D	0.78

7. Relationship Between Disruption Index and Citation Frequency Stable Time Windows

For most research papers, citation performance after publication allows estimation of long-term relative citation impact, indirectly indicating paper quality. Since both disruption index and citation frequency measure research quality but reflect different dimensions (originality vs. academic impact), two key questions arise: (1) Do these indicators require consistent time windows for assessment? If not, how large are the differences? (2) Which indicator stabilizes first?

Using the same 19-year reference window, we calculated citation frequency stability across time windows. Figure 5 [Figure 5: see original paper] shows the relationship between the two indicators' stability for four representative disciplines: Economics & Business, Mathematics, Clinical Medicine, and Neuroscience, each representing distinct patterns.

Figure 5. Stability Curves of Disruption Index and Citation Frequency Across Time Windows
Note: D represents disruption index, C represents citation frequency ($C = N_F + N_B$*)*

For question (1), Figure 5 shows significant differences in stability growth rates across disciplines. The line Y=X serves as a reference: points above indicate faster disruption index stabilization, while points below show faster citation frequency stabilization. Adapting the Gini coefficient concept, we quantified these differences by calculating the area S between the stability curve and the Y=X line (Figure 6 [Figure 6: see original paper]). Positive S values indicate faster disruption index stabilization overall.

Figure 6. Schematic of Stability Differences Between Indicators

Table 5 presents these differences. Economics & Business, Computer Science, Engineering, Social Sciences, Mathematics, Materials Science, and Environment/Ecology show faster disruption index stabilization, reflecting disciplinary characteristics that affect N_F, N_B, and N_R stability patterns.

Table 5. Stability Differences Between Disruption Index and Citation Frequency Across Time Windows

Discipline	S Value
Economics & Business	0.161
Computer Science	0.128
Mathematics	0.097
Engineering	0.042
Materials Science	0.021
Environment/Ecology	0.019
Physics	-0.004
Chemistry	-0.014
Clinical Medicine	-0.023
Neuroscience	-0.024
Molecular Biology & Genetics	-0.032

For question (2), Figure 7 [Figure 7: see original paper] shows which indicator reaches 0.8 stability first across 22 disciplines. Most disciplines show citation frequency stabilizing first, while Computer Science and Economics & Business show disruption index stabilizing first, and Engineering shows near-simultaneous stabilization. The scattered distribution reflects substantial disciplinary differences.

Figure 7. Time Windows for Disruption Index and Citation Frequency to Reach 0.8 Stability

Note: TC(0.8) = time for citation frequency to reach 0.8 stability; TD(0.8) = time for disruption index to reach 0.8 stability

Fundamentally, these differences stem from disciplinary research paradigms and nature, manifested through research duration, review periods, and recognition delays. While seven disciplines showed faster initial disruption index stabilization, only two ultimately achieved 0.8 stability first, because differential stabilization patterns of N_F and N_B altered the relative growth rates over time.

8. Summary and Discussion

Using 1998 SCI papers as samples with 19 years as the stable reference window and 1-18 years as test windows, we calculated disruption index stability across 22 disciplines. We found substantial disciplinary differences in time windows required to achieve 0.8 stability (Table 2), which has important implications for the timeliness and stability of research evaluation and provides guidance for time window selection in scientific activity studies.

We further revealed that at cited half-life, all disciplines achieve disruption index stability exceeding 0.8, suggesting cited half-life can serve as a practical reference time window when discipline-specific stable time windows are unavailable. However, the relationship between disruption index stable time windows and citation frequency stable time windows is complex. Seven disciplines showed faster initial disruption index stabilization, but only two ultimately achieved 0.8 stability first. This demonstrates that citation frequency stable time windows cannot be directly applied to the disruption index.

These disciplinary differences in stability ultimately arise from distinct research paradigms and practices, manifested through research duration, review periods, and recognition delays. While this study provides recommendations for using the disruption index, some outliers (e.g., Computer Science, Economics & Business) lack clear explanation, and whether national differences exist in disruption index stable time windows requires further investigation.

References

[1] Hu Jinxiu. International comparative study on university research evaluation systems[J]. Evaluation and Management, 2016, 14(1): 30-34.
[2] Li Xiaoxuan, Yang Guoliang. Research on the logical model of institute evaluation at Chinese Academy of Sciences[J]. Science of Science and Management of S.&T., 2009, 30(4): 5-8.
[3] Van R. Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises[J]. Scientometrics, 1996, 36: 397-420.
[4] Wu L, Wang D, Evans J. Large teams develop and small teams disrupt science and technology[J]. Management science, 2017, 63(3): 791-817.
[5] Funk R, Smith J. A dynamic network measure of technological change[J]. Nature, 2019, 566(7744): 378-382.
[6] Bornmann L. Disruption index depends on length of citation window[J]. EI profesional de la información, 2019, 2(28): 1-2.
[7] Bornmann L, Devarakonda S, Tekles A, et al. Disruptive papers published in Scientometrics: meaningful results by using an improved variant of the disruption index originally proposed by Wu, Wang, and Evans (2019)[J]. Scientometrics, 2020, 123(2): 1149-1155.
[8] Liu Xiaohui, Shen Zhesi, Liao Yu, et al. Research on improvement of disruption index and its influencing factors[J]. Library and Information Service, 2020, 64(24): 84-91.
[9] Walters D. The citation life cycle of articles published in 13 American Psychological Association journals: a 25-year longitudinal analysis[J]. Journal of the American Society for Information Science and Technology, 2011, 62(8): 1629-1636.
[10] Costas R, Leeuwen T, Raan A. The “Mendel syndrome” in science: durability of scientific literature and its effects on bibliometric analysis of

individual scientists[J]. Scientometrics, 2011, 89(1): 177-205.

[11] Levitt J, Thelwall M. Patterns of annual citation of highly cited articles and the prediction of their citation ranking: a comparison across subjects[J]. Scientometrics, 2008, 77(1): 41-60.

[12] Waltman L, Van Eck N. Toward a new crown indicator: an empirical analysis[J]. Scientometrics, 2011, 3(87): 467-481.

[13] Nederhof A, Van L, Clancy P. Calibration of bibliometric indicators in space exploration research: a comparison of citation impact measurement of the space and ground-based life and physical sciences[J]. Research evaluation, 2012, 21(1): 79-85.

[14] Abramo G, Cicerero T. A sensitivity analysis of researchers' productivity rankings to the time of citation observation[J]. Journal of informetrics, 2012, 6(2): 192-201.

[15] Costas R, Leeuwen T, Raan A. Effects of the durability of scientific literature at the group level: case study of chemistry research groups in the Netherlands[J]. Research policy, 2013, 42(4): 886-894.

[16] Abramo G, Cicerero T, Angelo C. A sensitivity analysis of research institutions' productivity rankings to the time of citation observation[J]. Journal of informetrics, 2012, 6(2): 298-306.

[17] Jonathan A. Early citation counts correlate with accumulated impact[J]. Scientometrics, 2005, 63(3): 567-581.

[18] Abramo G, Cicerero T, Angelo C. Assessing the varying level of impact measurement accuracy as a function of the citation window length[J]. Journal of informetrics, 2011, 5(4): 659-667.

[19] Wang J. Citation time window choice for research impact evaluation[J]. Scientometrics, 2013, 94(3): 851-872.

[20] Wang X, Zhang Z. Improving the reliability of short-term citation impact indicators by taking into account the correlation between short- and long-term citation impact[J]. Journal of informetrics, 2020, 14(2): 101019.

[21] Ren Shengli, Li Jialin, Jin Bihui, et al. Statistical analysis of references and citations of Chinese scientific journals[J]. Acta Editologica, 2001, 4(5): 261-263.

[22] Tsay M, Chen Y. Journals of general & internal medicine and surgery: an analysis and comparison of citation[J]. Scientometrics, 2005(1): 17-30.

[23] Zhou Erqiang. Study on citation half-life changes and influencing factors of library and information science journals in SSCI from 2013-2017[J]. Jiangsu Science & Technology Information, 2020, 37(12): 12-15.

[24] He Wen, Ye Jiyuan. Comparative study on the relationship between journal impact factor and cited half-life: taking library and information science journals as examples[J]. New Century Library, 2015, 4(4): 39-43.

[25] Dai Liqiong. Relationship between cited half-life and impact factor, total citations in oncology medical journals[J]. Journal of Medical Informatics, 2012, 33(6): 61-64.

[26] Zhang Huangqun. Correlation between journal cited half-life and related indicators[J]. Acta Editologica, 2011, 4(1): 114-116.

[27] Clarivate Analytics. Journal list[EB/OL].[2021-05-23]. https://help.incites.clarivate.com/incitesLiveESI/E
Group/overviewESI/esiJournalsList.html.

[28] Shan Weihua, Zhou Erqiang, Li Jianhua, et al. Analysis of publication delay of Chinese journal papers in Medline database[J]. Information Research, 2016, 4(4): 37-39.
[29] Liu Xueli. Control of publication delay and journal publication cycle in medical papers[J]. Chinese Journal of Scientific and Technical Periodicals, 2003, 14(2): 153-155.

Author Contributions

Liu Xiaohui: Designed the research structure, collected and processed data, wrote the initial draft.
Shen Zhesi: Supervised research structure, revised the paper.
Liao Yu: Processed data, created visualizations, revised the paper.
Zhu Manman: Processed data, revised the paper.
Yang Liying: Supervised research structure, revised the paper.

Note: Figure translations are in progress. See original paper for figures.