Beyond surface correlations: Reference behavior mediates the disruptiveness-citation relationship

Alex J. Yang, Fanming Wang, Yujie Shi, Yiqin Zhang, Hao Wang, Sanhong Deng†

School of Information Management, Nanjing University, Nanjing, 210093, China

Citation: Yang, A. J., Wang, F. M., Shi, Y. J., Zhang, Y. Q., Wang, H., & Deng, S. H. (2025). Beyond surface correlations: Reference behavior mediates the disruptiveness-citation relationship. *Journal of Data and Information Science*, 10(3), 7–31. https://doi.org/10.2478/jdis-2025-0029

DOI: 10.2478/jdis-2025-0029 **CSTR:** 32295.14.jdis-2025-0029

Received: Feb. 9, 2025; Revised: Mar. 26, 2025; Apr. 21, 2025; Accepted: Apr. 24, 2025

ABSTRACT

Purpose: This study examines why papers with high CD indices (measuring research disruptiveness) increasingly show reduced citation impact and investigates whether this represents genuine impact reduction or methodological artifacts.

Design/methodology/approach: We analyzed 29 million papers (1950-2016) using Poisson regression to examine relationships between the CD index and citation count, with controls for fields, team size, and reference count.

Findings: Papers with high CD indices showed reduced citation impact over time. However, when controlling for increasing reference counts in papers, this relationship reversed, revealing a positive association. Papers with more references exhibit lower CD indices owing to the index's sensitivity to the reference count, while achieving higher citation counts. Alternative innovation metrics consistently show positive correlations with citation impact.

Research limitations: The approach may not adequately capture the reduced citation impact of highly disruptive papers with fewer references. The analysis is limited to journal articles and shows correlation rather than causality.

† Corresponding author: Sanhong Deng (Email: sanhong@nju.edu.cn).

Practical implications: The apparent undervaluation of disruptive research stems from methodological artifacts in the CD index calculation driven by evolving reference patterns. Researchers should control for the reference count when using this metric.

Originality/value: This study reveals that bias against disruptive research is specific to the CD index's calculation methodology, identifying reference behavior as the key factor affecting the relationship between disruption metrics and citation impact.

Keywords Science of science; Innovation; Disruption; CD index; Citation impact

1 Introduction

In scientific research, the evaluation of scholarly impact and innovation has long been a subject of profound interest and scrutiny (Fortunato et al., 2018). Traditionally, citation counts have served as a primary metric for assessing the significance and influence of academic papers (Waltman, 2016). However, as the landscape of scientific inquiry continues to evolve, there is a need for more nuanced and sophisticated methodologies to capture the disruptive potential of research contributions. One such endeavor in this pursuit is the introduction of the CD index, a metric proposed by Funk and Owen-Smith (2017). This index, grounded in the analysis of structural features within the deep citation network, offers a systematic approach for evaluating the disruptive nature of academic papers. By quantifying the alignment of a paper with established knowledge versus its potential to induce paradigm shifts, the CD index presents a novel lens through which to gauge scholarly innovation (Bornmann et al., 2020b).

As shown in Figure 1, the CD index is defined as:

$$CD\,index = p_D - p_C = \frac{n_i - n_j}{n_i + n_j + n_k} \tag{1}$$

where n_i represents the number of future papers citing the focal paper (FP) that do not reference its cited sources, n_j denotes the number of future papers citing FP and its references, and n_k represents the number of papers citing FP's references without citing FP itself. A lower CD index indicates alignment with established knowledge, whereas a higher CD index signifies papers with the potential to induce paradigm shifts. Following prior research methodologies (Park et al., 2023), we employed a 5-year window to compute the CD index and citation count. This approach allows for the analysis of the relationship between the CD index and citation impact across a comprehensive dataset spanning the past seven decades.

Leibel and Bornmann (2024) provided a comprehensive appraisal of the disruption index, deliberating on its conceptual underpinnings, capabilities, extensions, and constraints. Although it has some flaws and limitations, the CD index has been widely used in bibliometrics, science of science, and many other fields of study. Notably, at least three papers published in the esteemed journal *Nature* utilized the CD index as the main variable (Lin et al., 2023; Park et al., 2023; Wu et al., 2019).

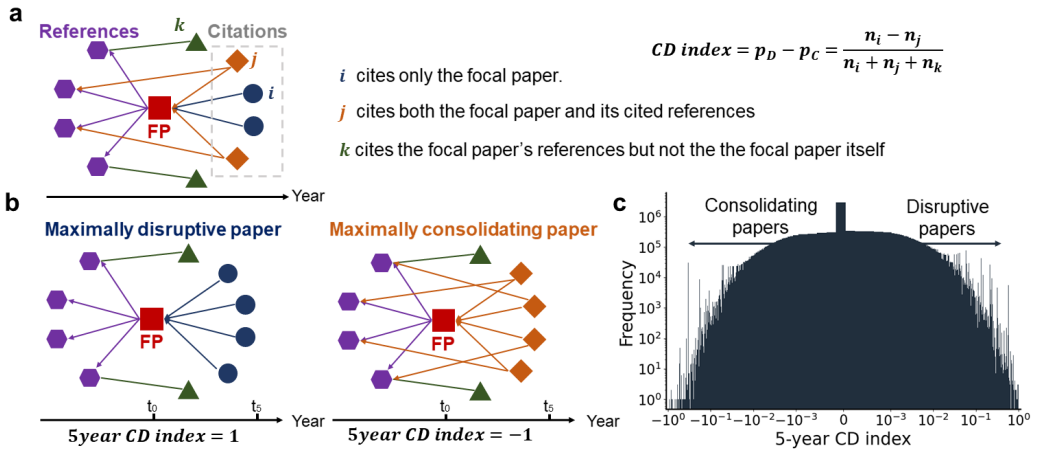


Figure 1. Quantifying the 5-year CD index.

While it has been assumed that papers and patents with high CD indices often herald paradigm shifts and predict significant scientific milestones (e.g. Nobel-winning papers) (Wu et al., 2019), recent research suggests a departure from this expectation (Zeng et al., 2023). Despite the acknowledged importance of disruptive research in precipitating breakthroughs, both this paper and the findings of Zeng et al. (2023) reveal an inconsistent, and at times negative, correlation between elevated CD indices and heightened scientific impact. This incongruity necessitates a closer examination of the potential bias against disruptive research within the scientific community. Although the CD index was designed to capture the degrees of consolidation and destabilization in papers, a concept distinct from citation impact, many researchers believe that both the CD index and citation count should reflect the quality of papers (Lin et al., 2023). Therefore, it is crucial to understand the reasons for this negative association. Is it a systemic phenomenon or is it merely due to the reference patterns in different papers?

The primary objective of this study is to scrutinize the underlying causes of this bias, with a particular focus on determining whether disruptive papers genuinely experience diminished impact or whether the issue stems from methodological nuances inherent in the calculation of the CD index. Our findings suggest that the latter proposition holds greater plausibility. Specifically, we contend that reference behavior significantly influences the observed bias against the CD index in the assessment of scientific impact.

2 Background

Several key pieces of literature motivate or relate to this study. First, Funk and Owen-Smith (2017) proposed the CD index and argued that it can (i) capture the extent to which future inventions that build on a technology also rely on that technology's predecessors, (ii) account for

variation in the extent to which an invention alters the use of its predecessors over time, (iii) capture the degrees of consolidation and destabilization, ranging from large-scale transformations to incremental shifts, and (iv) distinguish between consolidating and destabilizing technologies that may have similar impacts. Further studies suggest that different collaborative patterns exhibit varying degrees of disruptive potential (Lin et al., 2023; Wu et al., 2019). Many researchers have noted that the CD index can be biased by the reference number of the papers (Leibel & Bornmann, 2024; Ruan et al., 2021; Yang, Gong, et al., 2024). Bornmann et al. (2020a) observed that the convergent validity of the CD index is lower than its variants that use reference co-citing thresholds, which may address some of the CD index's limitations. Despite its limitations and the existence of various CD index variants (Leibel & Bornmann, 2024), it remains a widely studied metric. However, the relationship between the CD index and citation count has not yet been fully elucidated. This relationship holds significant importance, as both metrics are widely used to proxy a work's value.

Second, a cover article in *Nature* by Park et al. (2023) found that the CD index of papers decreased over time. Simultaneously, the number of papers, average citations, and average number of references have increased rapidly over the past century (Fortunato et al., 2018). Therefore, an understanding of how these patterns are linked is of significant interest. However, the findings of Park et al. (2023) have been challenged by some researchers. A recent study by Macher et al. (2024) argued that the declining trend of the CD index can be attributed to the truncation of all backward patent citations before 1976. Macher et al. (2024) constructed a larger dataset of patents, which compensated for the pre-1976 data, resulting in patents having more references and, consequently, generating a different trend. This also illustrates how the reference behavior can largely influence the distribution of the CD index. Further, Petersen et al. (2024) suggest that the decline in the CD index is driven by "citation inflation"—a systematic time-dependent bias inherent in real citation networks. One key driver of citation inflation is the increasing length of reference lists over time, which in turn increases the density of citation networks and causes the disruption index to converge toward zero. These studies underscore the complexity of measuring disruptions and innovation using citation-based metrics.

Third, another body of literature relevant to this study addresses the bias against novelty (Chai & Menon, 2019; Wang et al., 2017). Novelty serves as an ex-ante measure of scientific innovation, with numerous methodologies for its assessment (Foster et al., 2015). These studies predominantly utilized the ex-ante novelty of papers to predict their future citation impact. It is noteworthy that some measures of novelty exhibit a positive correlation with citation impact (e.g. novelty as the atypicality by Uzzi et al. (2013)), while others demonstrate a negative correlation (e.g. novelty as the first combination by Wang et al. (2017)). However, disruption fundamentally differs from novelty. Disruption measures are ex-post, similar to citation counts, and the metric of interest in this study, the CD index, is also citation-based. This offers a distinct advantage in our study: not only can we determine the direction of the relationship (whether positive or negative), but we can also gain deeper insights and explanations, given that both disruption and citation impacts are

derived from the same system.

Finally, the work of Zeng et al. (2023) is particularly relevant to this study. They argued that disruptive papers are losing impact, supported by the robust and consistent negative correlations between citation counts and the CD index. Our initial findings align with those of Zeng et al. (2023); however, we propose that the negative relationship between citation count and the CD index can be fully accounted for by reference behaviors. Specifically, when controlling for author features or reference-level variables, the negative effect dissipates. Thus, the observed negative correlations between citation counts and the CD index have explicit reasons. Consequently, we cannot conclusively state that disruptive papers are losing impact; instead, the apparent bias against disruptive papers may be attributable to reference behaviors or the CD index itself.

3 Data and method

This study utilized a combined dataset from Microsoft Academic Graph (MAG) and OpenAlex, both recognized for their comprehensive coverage of scholarly literature. The MAG dataset, spanning 1800-2021 and containing over 200 million documents (including journal articles, conference proceedings, and preprints), was integrated with OpenAlex, which encompasses 246,880,876 records across various publication types. Focusing on journal articles published between 1950 and 2016, and ensuring a minimum 5-year citation window (Wang, 2013). Papers or patents with at least one citation count and reference count were included to compute the CD index (Wu et al., 2019). Subsequently, the focus was refined solely to journal articles, acknowledging the distinctive citing behavior inherent across different scientific document types. Thus, the dataset comprises a robust cohort of 29,009,690 papers.

Given that citation counts typically conform to a log-normal distribution (Zeng et al., 2017), the primary analysis employed Poisson regression models to scrutinize the relationship between the CD index and citation count in each publication year from 1950 to 2016. The 5-year CD index was utilized as the dependent variable, while the citation count served as the independent variable. Split-sample regression was employed to analyze yearly associations.

In Section 4.1, to examine the association between the 5-year CD index and the 5-year citation count without introducing any controls, we conducted a yearly split-sample regression analysis using the following model:

$$ln(E[Citation_i])=\alpha+\beta_1(CD\ index_i)+\varepsilon_i \quad (2)$$

where i represents the article, and ε_i denotes the error term.

In Section 4.2, to provide a more nuanced analysis, we incorporated fixed effects for fields, team size, and reference count into the model while still conducting yearly split-sample regression analysis. The revised model is specified as follows:

$$ln(E[Citation_i])=\alpha+\beta_1\ (CD\ index_i)+\delta_f\ FieldFE_f+\theta_c\ Team_c+\varphi_r\ Ref_r+\varepsilon_i \quad (3)$$

In Section 4.3, to evaluate the role of a set of potential explanatory variables, we focus on papers published between 1990 and 2010. This analysis controls for year and field fixed effects without employing a yearly split-sample regression. The specification is as follows:

$$ln\left(E\left[Citation_i\right]\right)=\alpha+\beta_1\left(CD\ index_i\right)+\sum_{k=2}^{K}\beta_k\left(V_{k,i}\right)+\gamma_t YearFE_t+\delta_f FieldFE_f+\varepsilon_i \quad (4)$$

where i represents the article, t the year of publication, f the field, $V_{k,i}$ represents the potential explanation variables, $YearFE_t$ denotes year fixed effects, $FieldFE_f$ captures second-level field fixed effects, and ε_i denotes the error term. Count variables undergo natural log-transformation within the models to ensure appropriate scaling.

4 Main results

4.1 Increasing bias against the CD index

We begin by employing the basic Poisson regression model to examine the association between the CD index and citation count without introducing any controls. Specifically, we conducted split-sample regressions based on the publication year of the papers from 1950 to 2016. This approach provides initial insights into the evolving relationship between the CD index and citation impact over time.

The findings illustrated in Figure 2 reveal a noteworthy pattern. Prior to 1962, there was a statistically significant positive association between the CD index and citation count, indicating that papers with a higher CD index experienced a higher citation impact during this period. However, subsequent to this time point, the coefficient turns negative, with the exception of the years 1974 and 1975, where the coefficient is statistically insignificant. This trend persists until the present, suggesting that since then, papers with a higher CD index have been linked to a lower citation impact. Notably, the coefficient exhibits a declining trajectory (except for the latest two years, which may be due to the incomplete data in its 5-year citation windows), indicating that over time, papers with a high CD index are increasingly associated with diminished citation impact.

This result aligns with the findings of Zeng et al. (2023) and remains robust across various models (Appendix A), irrespective of whether 5-year windows are employed (Appendix A). Recognizing the skewed distribution of the CD index (Yang, Deng, et al., 2023), we transformed the index into percentiles, calculated within each publication year and field (using the 19 level-0 fields of MAG). This normalization process did not change the overall conclusions (Appendix A).

It is important to note that Funk and Owen-Smith (2017) contended that the CD index was designed to capture a concept distinct from citation impact. They also introduced the mCD index, which is the product of the CD index and citation count. Although the mCD index demonstrated a consistent positive relationship with citation count, the coefficients exhibited a declining trend over time (Appendix A). This further reinforces our argument regarding the increasing bias against the CD index.

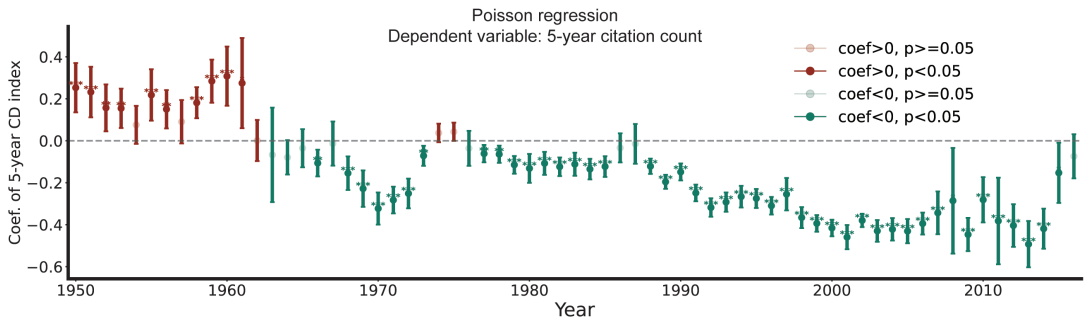


Figure 2. Basic yearly split-sample Poisson regression coefficients of the CD index on citation counts from 1950 to 2016. Error bars depict the upper and lower bounds of the 95% confidence intervals based on robust standard errors. Dark-red coloring indicates significant positive coefficients (p <0.05), and light-red coloring indicates non-significant positive coefficients. Dark-green coloring indicates significant negative coefficients (p <0.05), light-green coloring indicates non-significant negative coefficients. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

4.2 Reference count as a contributing factor

In this subsection, we augment the baseline Poisson regression models with fixed effects to elucidate the causes behind the negative association observed between the CD index and citation impact. The rationale for employing fixed effects instead of control variables is their capacity to address heterogeneity more comprehensively, thereby facilitating a deeper understanding of the underlying mechanisms.

Initially, we hypothesized that the observed negative association between the CD index and citation count may stem from heterogeneity across different fields of study. Variations in citation norms among distinct fields, such as the propensity for Medicine and Biology papers to garner more citations compared to social science papers of similar quality (Radicchi et al., 2008), could contribute to this phenomenon. Additionally, prior research by Chu and Evans (2021) indicated that larger fields tend to exhibit lower average CD index values. To explore this hypothesis, we introduced fixed effects representing 292 second-level fields into the regression model. The results depicted in Figure 3a suggest that the heterogeneity across different fields of study does not suffice to explain the escalating bias against the CD index.

Subsequently, we considered the potential influence of team size on the observed negative association between the CD index and citation count. The increasing prevalence of large teams in scientific endeavors may significantly alter research patterns (Jones et al., 2008; Wuchty et al., 2007). Notably, large teams often accrue higher citation counts than papers authored by individuals. Moreover, recent studies (Lin et al., 2023; Wu et al., 2019) propose that small, locally based teams may foster more disruptive research outcomes. To investigate this hypothesis, we introduced fixed effects representing team size (the number of authors) into the regression model. However, as indicated in Figure 3c, the heterogeneity in the different types of teams does not adequately explain the escalating bias against the CD index.

Finally, we propose that referencing behavior may contribute to the observed negative correlation between the CD index and citation count. This can be attributed to the sensitivity of the CD index to the number of references within the papers (Leibel & Bornmann, 2024). For instance, a paper with fewer references could potentially generate a high CD index because future citations are less likely to reference its limited sources. Conversely, papers with a higher number of references tend to have higher citation counts (Zeng et al., 2017). To test this hypothesis, we integrated reference count fixed effects into the regression model, with the maximum set at 60 (with counts above 60 capped at 60). As depicted in Figure 3b, upon inclusion of the reference count fixed effect, the coefficients exhibit positive significance, consistently observed from 1950 to 2016, without displaying a declining trend over time. The full model, encompassing all fixed effects, yields a similar outcome to the model solely incorporating reference fixed effects, yet significantly diverges from the other models (Figure 3d). Thus, we assert that referencing behavior can substantially elucidate escalating bias against the CD index.

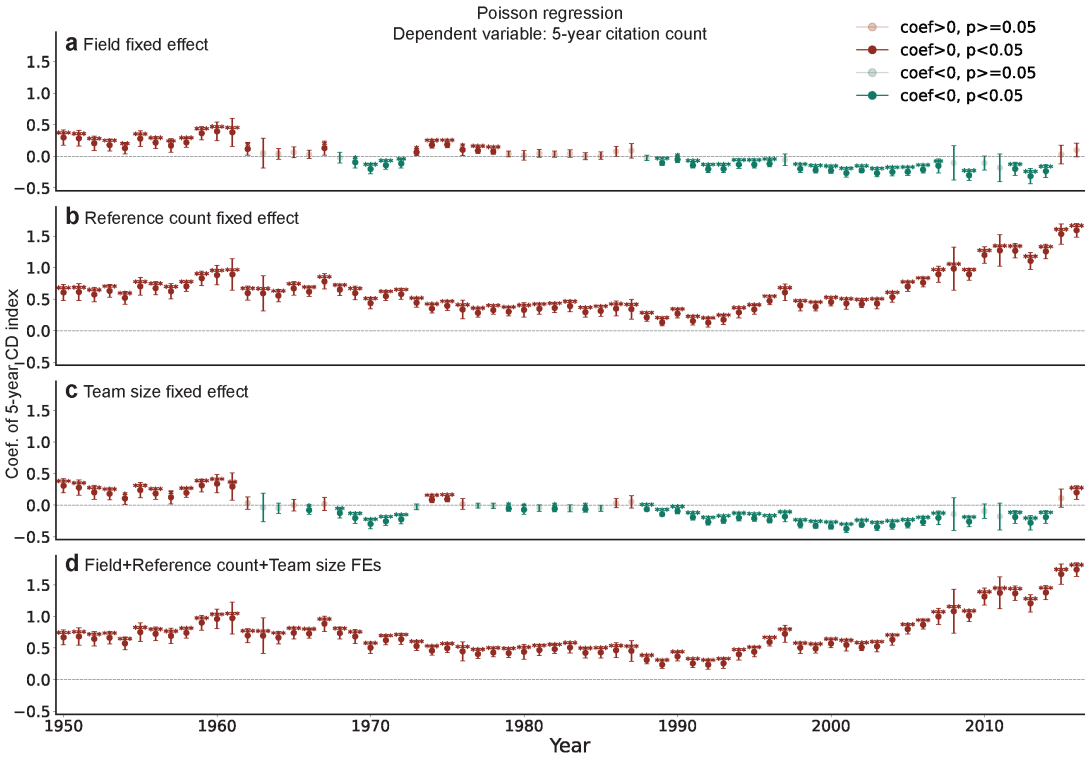


Figure 3. Yearly split-sample Poisson regression coefficients of the CD index on citation counts from 1950 to 2016, with (a) 292 fields fixed effect, (b) reference count fixed effect, (c) team size fixed effects, and (d) all fixed effects. Error bars depict the upper and lower bounds of the 95% confidence intervals based on robust standard errors. Dark-red coloring indicates significant positive coefficients (p<0.05), and light-red coloring indicates non-significant positive coefficients. Dark-green coloring indicates significant negative coefficients (p <0.05), light-green coloring indicates non-significant negative coefficients. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

To further validate our findings, we conducted an additional set of split-sample Poisson regression analyses utilizing papers' reference counts as the split rather than their publication years. This approach aimed to ascertain whether the observed negative association between the CD index and citation count could be explained by variations in referencing behavior. Our hypothesis is that if reference count adequately captures the influence of the CD index on citation count, then the coefficients of the CD index in split-sample Poisson regressions, stratified by reference count, should consistently exhibit positive effects, assuming that disruptive papers are more impactful. Essentially, by analyzing papers with equivalent reference counts, we aimed to isolate the impact of the CD index on citation counts from referencing behavior.

The results depicted in Figure 4 align with our hypotheses. When examining the relationship between the CD index and citation count within papers sharing the same reference count, ranging from 1 to 59 and above 60, all subsamples demonstrated positive effects, with the majority being statistically significant. Furthermore, upon incorporating additional fixed effects into the models, including fields, team level, author level, and year fixed effects, all subsamples exhibited significantly positive results at the 0.001 significance level. Consequently, we assert that referencing behavior substantially contributes to the elucidation of the observed bias against the CD index. However, this analysis underscores a limitation: by strictly controlling for reference count in our fixed effects regression, we may obscure the reduced citation impact of highly disruptive papers that cite fewer references, suggesting that our method may not fully capture the dynamics of disruption-driven citation patterns.

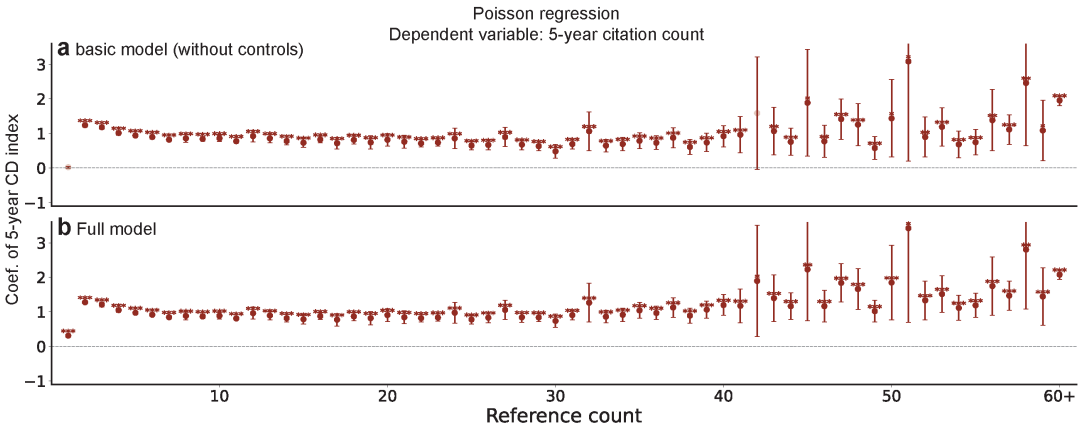


Figure 4. Split-sample Poisson regression coefficients of the CD index on citation counts based on reference counts. (a) Basic models without any fixed effects. (b) Full models including all fixed effects. Error bars depict the upper and lower bounds of the 95% confidence intervals based on robust standard errors. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

4.3 Other possible explanations

In this subsection, we analyze additional potential channels that could explain the negative relationship between the CD index and citation count. To ensure that the results are intuitive, we focused on a subset of the dataset, considering only papers published between 1990 and 2010. This

period was chosen because the negative relationship between the CD index and citation count was most significant during these years, even when controlling for the field (see Figure 3a). We utilized Poisson regression models, incorporating both year- and field-fixed effects in all analyses to obtain unbiased results.

First, we consider the potential influence of team dynamics on the observed negative association between the CD index and citation count. While in Figure 3c we only analyzed the control of team size, many other factors could also influence the disruptive potential of papers, such as cross-institutional collaboration (Yang, 2024, 2025), international collaboration (Lin et al., 2023), interdisciplinary collaboration (Liu et al., 2024), and gender diversity (Yang et al., 2022). Detailed definitions of these variables are provided in the Appendix C.

The results in Table 1 reveal that none of the team-level control variables can account for the negative relationship between the CD index and citation count. The coefficient of the CD index on citation count remains negatively significant in Models (1–7) after we control for these team-level variables.

Table 1. The effect of the CD index on citation count with team-level controls.

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	DV: 5-year citation count (Poisson regression)						
5-year CD index	-0.1993*** (0.0133)	-0.1541*** (0.0137)	-0.1639*** (0.0137)	-0.1789*** (0.0135)	-0.2018*** (0.0135)	-0.1618*** (0.0147)	-0.1234*** (0.0151)
ln(Team size)		0.3405*** (0.0016)					0.2533*** (0.0018)
ln(Institution count)			0.3960*** (0.0020)				0.2159*** (0.0023)
ln(Country count)				0.4382*** (0.0029)			0.0358*** (0.0034)
ln(Home field count)					0.1950*** (0.0018)		-0.0498*** (0.0018)
Gender diversity						0.2868*** (0.0021)	0.0579*** (0.0021)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13,180,603	13,180,603	13,180,603	13,180,603	13,180,603	12,262,497	12,262,497
Pseudo R2	0.05826	0.08234	0.0755	0.06733	0.06115	0.05924	0.08252

Note: Robust standard errors are reported in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Second, we hypothesize that the experience and academic status of authors may contribute to the observed negative association between the CD index and citation count. It is plausible that advancing age and increased academic stature may lead authors to produce less disruptive research due to the burden of knowledge (Bloom et al., 2020; Jones, 2009; Jones & Weinberg, 2011; Li, Tessone, et al., 2024), while also potentially enhancing the citation counts of their papers (Yang, Xu, et al., 2024). To test this hypothesis, we introduced a set of author career-level variables,

including the average career age of authors, average and maximum career productivity of authors, and average and maximum career citations of authors. These variables are dynamic and are calculated only up to the publication year of the focal paper. Detailed definitions of these variables are provided in Appendix C.

As shown in Table 2, author career age cannot fully account for the observed negative association between the CD index and citation count. The coefficient in Model (2) remained negatively significant, although the negative effect size was much smaller. However, both author career productivity and citations can account for this negative association. The coefficients in Models (3-6) are all positively significant, and the effect size is larger when controlling for author career citations and maximum author features. When combining all author career-level controls in Model (7), the coefficient of the CD index is positively significant and exhibits a larger effect size.

While these results suggest that author career level can partially explain the observed negative association between the CD index and citation count, it remains unclear whether this is due to differences in reference behavior or other factors. A deeper exploration of the mechanisms at the author level is beyond the scope of this study.

Table 2. The effect of the CD index on citation count with author career-level controls.

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	DV: 5-year citation count (Poisson regression)						
5-year CD index	-0.1993*** (0.0133)	-0.1012*** (0.0137)	0.0429** (0.0146)	0.0590*** (0.0146)	0.2263*** (0.0157)	0.2307*** (0.0157)	0.3443*** (0.0166)
ln (Avg career age+1)		0.3368*** (0.0011)					-0.3891*** (0.0019)
ln (Avg career productivity+1)			0.2601*** (0.0007)				-0.2974*** (0.0058)
ln (Max career productivity +1)				0.2445*** (0.0005)			-0.0278*** (0.0048)
ln (Avg career citations +1)					0.2198*** (0.0005)		0.4242*** (0.0044)
ln (Max career citations+1)						0.2065*** (0.0004)	0.0573*** (0.0040)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13,180,603	13,180,603	13,180,603	13,180,603	13,180,603	13,180,603	13,180,603
Pseudo R2	0.05826	0.08442	0.11291	0.12075	0.16595	0.1694	0.20231

Note: Robust standard errors are reported in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Finally, we further investigated additional reference-level factors that may influence the association between the CD index and citation count. The literature indicates that reference count, reference age, and citations of references (i.e. reference popularity) can all affect the CD index (Leibel & Bornmann, 2024; Li, Lin, et al., 2024). Unlike author-level and team-level factors, these reference-level influences may stem from the definition and calculation of the CD index itself.

The results presented in Table 3 reveal that the reference count significantly accounts for the observed negative association between the CD index and citation count. When controlling for reference count or reference age, the effect of the CD index on citation count transitions from negatively significant to positively significant. When accounting for the citations of references, the coefficient transitions from being negatively significant to insignificant at the 0.05 level. When combining all reference-level controls in Model (6), the coefficient increases to 1.92, which is significantly larger than the original estimate.

Although these coefficients should be interpreted with caution, the model suggests notable changes in the association. In Model (1), without controlling for reference-level variables, an increase of one standard deviation in the CD index (0.14, see Appendix C) is associated with a 2.8% decrease in citations (exp(-0.20*0.14) - 1). However, after accounting for all reference-level variables in Model (6), an increase of one standard deviation in the CD index is associated with a 30.1% increase in citations (exp(1.92*0.14) - 1).

Table 3. The effect of the CD index on citation count with reference-level controls.

Models	(1)	(2)	(3)	(4)	(5)	(6)
	DV: 5-year citation count (Poisson regression)					
5-year CD index	-0.1993*** (0.0133)	0.6228*** (0.0204)	0.2267*** (0.0123)	-0.0006	0.0310 (0.0170)	1.920*** (0.0212)
ln(Ref count)		0.5939*** (0.0013)				0.8642*** (0.0016)
ln(Ref age+1)			-0.4257*** (0.0011)			-0.7780*** (0.0019)
ln(Avg ref cit+1)				0.1910*** (0.0005)		0.6947*** (0.0021)
ln(Max ref cit+1)					0.1478*** (0.0003)	-0.4027*** (0.0014)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13,180,603	13,180,603	13,180,603	13,180,603	13,180,603	13,180,603
Pseudo R2	0.05826	0.16705	0.08522	0.09294	0.09803	0.27983

Note: Robust standard errors are reported in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

4.4 Mechanisms behind the influence of reference behavior

Based on these findings, we assert that the observed increasing bias against the CD index (a negative association between the CD index and citation count) can be attributed to changing referencing behaviors in scholarly papers over time. As depicted in Figure 5a, the average number of references in papers has steadily increased over the past seven decades, from 10 to 37. This upward trend in referencing behavior suggests a temporal shift in scholarly reference practices. The average 5-year citation count has also increased, climbing from 5 to 19 over the same period (Figure 5b), and the average 5-year CD index has declined from 0.09 to 0.03 (Figure 5c). Our earlier findings suggest that when controlling for the reference count, the association between the

CD index and citation count remains consistently positive. Therefore, we redirect our focus to examining the relationship between the variables of interest and reference count.

The results depicted in Figure 5d demonstrate a linear association between higher reference counts and higher 5-year citation counts. Conversely, the findings in Figure 5e indicate that a higher reference count is linearly associated with a lower 5-year CD index. Thus, as the average number of references in papers increases over time, the prevalence of papers with a high number of references contributes to the negative association between the CD index and citation count.

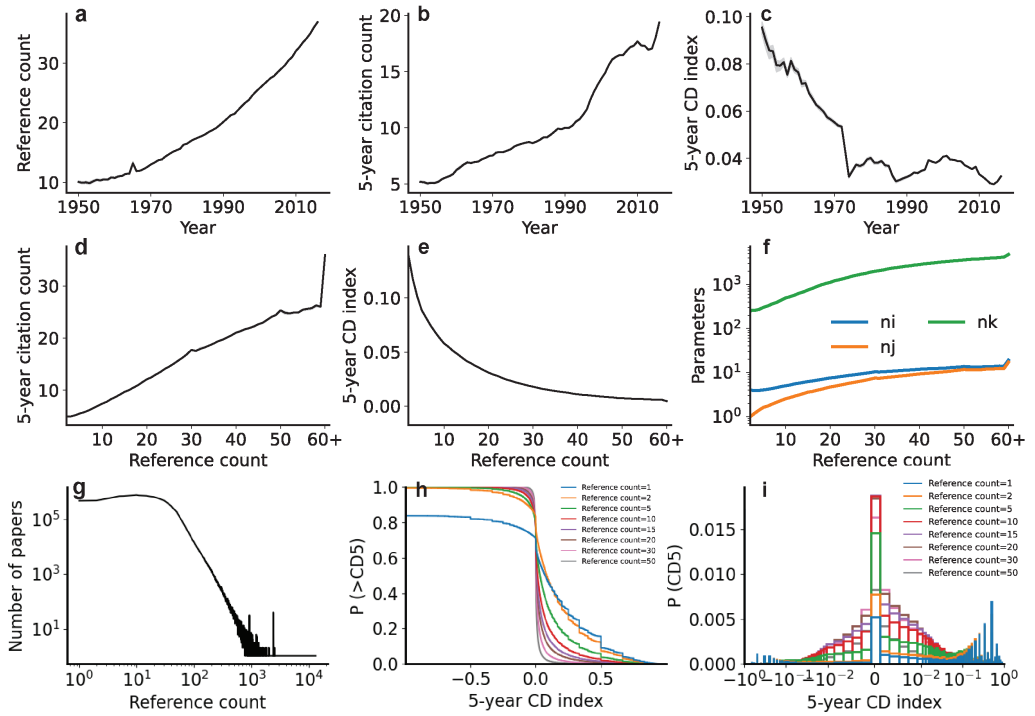


Figure 5. Reference counts as a contributing factor to the bias against CD index. (a) The average reference count of papers increases over time. (b) The average 5-year citation count of papers increases over time. (c) The average 5-year CD index of papers decreases over time. (d) Papers with higher reference count are associated with higher 5-year citation count. (e) Papers with higher reference count are associated with higher 5-year CD index. (f) Papers with higher reference count are associated with higher parameters of 5-year CD index (n_i, n_j, and n_k). (g) The number of papers with different reference count follows log-normal distribution. (h-i) The complementary cumulative distribution function (ccdf) and probability density function (pdf) of the CD index for papers with varying reference counts. Shaded areas indicate 95% confidence intervals.

Finally, we propose a potential mechanism underlying the influence of reference behavior on this association, elucidated through the parameters of the 5-year CD index (n_i, n_j, and n_k). The findings depicted in Figure 5f suggest that the impact of the reference count on distinct parameters varies. Considering n_i as indicative of disruptive citations and n_j as representative of consolidating citations, the analysis revealed that a higher reference count corresponds to a greater increase in n_j,

subsequently resulting in a decrease in the proportion of n_i. This, in turn, leads to a reduction in the CD index. Furthermore, a higher value of n_k also contributes to a lower CD index. In our dataset, where the average CD index exceeded 0, a larger denominator resulted in a diminished CD index.

The mechanism described can substantially impact the outcomes of analyses involving the CD index. As depicted in Figures 5h-i, papers with a limited number of references are likely to display extreme values on the CD index—either significantly high or low—while those with a larger number of references tend to show CD index values clustered around zero. This mechanism underscores the sensitivity of the CD index to reference count, which, combined with our fixed effects regression method's strict control for reference count, may obscure the citation dynamics of highly disruptive papers with fewer references. Consequently, while reference behavior explains the methodological bias in the CD index, our analytical approach may not fully resolve whether disruptive papers are losing impact, suggesting the need for alternative regression methods or metrics to capture this phenomenon.

5 Patterns of alternative innovation measures

In our main analysis, we found that the previously positive correlation between the CD index and citation count turned negative, with the coefficient showing a downward trend. While we attribute this shift to referencing patterns and the structural definition of the CD index, particularly the inclusion of n_k, we seek to determine whether this trend is unique to the CD index or extends to other measures of innovation. For instance, highly disruptive research may inherently cite fewer prior studies, leading to fewer references.

To explore this, we examine several alternative innovation measures, including (1) New words: The number of unique unigrams introduced by a paper for the first time. (2) New word reuse: The number of new words introduced by a paper, weighted by their future reuse. (3) New word combinations: The number of new pairwise word combinations that appear for the first time. (4) New word combination reuse: The number of new pairwise word combinations weighted by their future reuse. (5) Atypical knowledge combinations (Uzzi et al., 2013). (6) Disruptive citations (Yang & Deng, 2024; Yang, Hu, et al., 2023; Yang et al., 2025). Unlike the CD index, these measures are less likely to be influenced by reference patterns. Therefore, we expect them to exhibit different trends, providing further insight into the relationship between innovation and citation impact.

5.1 New words, new word combination and their reuse

Following the methodology of Arts et al. (2025), we processed and integrated their datasets to identify new words and their reuse. Specifically, we analyzed English, non-retracted journal publications from OpenAlex, leveraging a sophisticated text-processing pipeline. Using the spaCy model "en_core_sci_lg," the pipeline includes tokenization, POS-tagging, dependency parsing, chunking, lemmatization, cleaning, baseline removal, and vectorization. This approach enables us to track the first appearance of words and their subsequent reuse to quantify innovation. Notably,

this method is entirely text-based and independent of citation or reference information.

We begin by examining the relationship between the number of new words (log-transformed) and the 5-year citation count using a basic Poisson regression model, without introducing any controls. By conducting split-sample regressions across publication years (1950–2016), we find that the relationship between new words and citation count remains consistently positive (Figure 6a). This relationship is statistically significant at the 0.001 level for nearly all years, except 1950 and 1951, suggesting a distinct trend from the CD index. Furthermore, we assessed the impact of reference-level factors on this relationship. As shown in Table 4, controlling for reference-level variables did not alter the direction, effect size, or significance of the association. Across all models, we observe a robust, statistically significant (0.001 level) positive relationship between new words (log-transformed) and the 5-year citation count.

Using the same procedure, we then analyze the relationship between new word reuse (log-transformed) and the 5-year citation count. As shown in Figure 6b, this relationship also remained consistently positive and statistically significant at the 0.001 level. Similarly, Table 5 demonstrates that controlling for reference-level variables did not alter the direction, effect size, or significance of the association. Across all models, we find a robust, statistically significant (0.001 level) positive relationship between new word reuse (log-transformed) and the 5-year citation count.

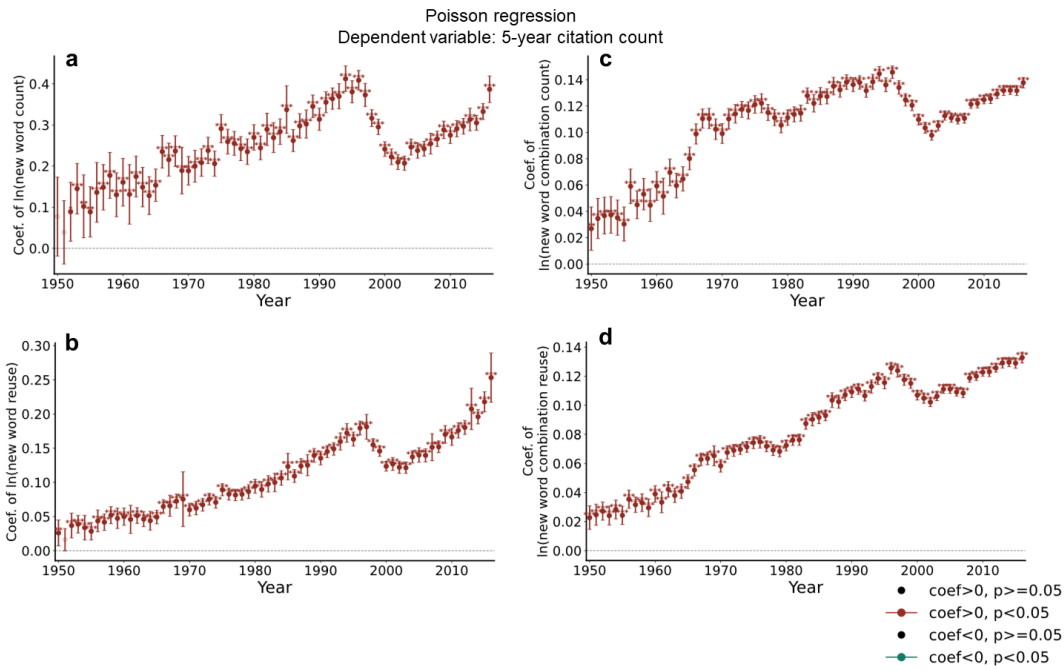


Figure 6. Yearly split-sample Poisson regression coefficients of new words, new word combinations, and their reuse on citation counts from 1950 to 2016. Error bars depict the upper and lower bounds of the 95% confidence intervals based on robust standard errors. Dark-red coloring indicates significant positive coefficients (p <0.05), and light-red coloring indicates non-significant positive coefficients. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4. The effect of the number of new words on citation count with reference-level controls.

Models	(1)	(2)	(3)	(4)	(5)	(6)
	DV: 5-year citation count (Poisson regression)					
ln(New words count+1)	0.2191*** (0.0021)	0.2195*** (0.0020)	0.2094*** (0.0021)	0.1970*** (0.0021)	0.1956*** (0.0021)	0.1771*** (0.0020)
ln(Ref count)		0.6233*** (0.0009)				0.8398*** (0.0011)
ln(Ref age+1)			-0.4137*** (0.0009)			-0.6649*** (0.0015)
ln(Avg ref cit+1)				0.1836*** (0.0004)		0.6474*** (0.0017)
ln(Max ref cit+1)					0.1439*** (0.0002)	-0.3873*** (0.0011)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25,263,987	25,263,987	25,174,032	25,262,880	25,262,880	25,174,032
Pseudo R2	0.06392	0.17867	0.08835	0.09575	0.10168	0.26690

Note: Robust standard errors are reported in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 5. The effect of new word reuse on citation count with reference-level controls.

Models	(1)	(2)	(3)	(4)	(5)	(6)
	DV: 5-year citation count (Poisson regression)					
ln(New words reuse+1)	0.1213*** (0.0011)	0.1180*** (0.0011)	0.1147*** (0.0011)	0.1112*** (0.0011)	0.1113*** (0.0011)	0.0936*** (0.0011)
ln(Ref count)		0.6230*** (0.0009)				0.8390*** (0.0011)
ln(Ref age+1)			-0.4117*** (0.0009)			-0.6622*** (0.0015)
ln(Avg ref cit+1)				0.1827*** (0.0004)		0.6453*** (0.0017)
ln(Max ref cit+1)					0.1434*** (0.0002)	-0.3865*** (0.0011)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25,263,987	25,263,987	25,174,032	25,262,880	25,262,880	25,174,032
Pseudo R2	0.06638	0.18096	0.09056	0.09790	0.10385	0.26838

Note: Robust standard errors are reported in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

We then analyzed the relationship between new word combinations and their reuse (both log-transformed) and 5-year citation counts. Here, new word combinations are based on all possible pairs between any of the words in a paper. As shown in Figures 2c-d, this relationship also remained consistently positive and statistically significant at the 0.001 level. Similarly, Tables 6-7 demonstrates that controlling for reference-level variables does not alter the direction, effect size,

or significance of the association. Across all models, we find a robust, statistically significant (0.001 level) positive relationship between new word combinations and their reuse (both log-transformed) and 5-year citation count.

Table 6. The effect of the number of new word combinations on citation count with reference-level controls.

Models	(1)	(2)	(3)	(4)	(5)	(6)
	DV: 5-year citation count (Poisson regression)					
ln(New words count+1)	0.0978*** (0.0004)	0.0665*** (0.0004)	0.1002*** (0.0004)	0.0894*** (0.0004)	0.0841*** (0.0004)	0.0598*** (0.0004)
ln(Ref count)		0.6144*** (0.0010)				0.8333*** (0.0011)
ln(Ref age+1)			-0.4242*** (0.0009)			-0.6654*** (0.0015)
ln(Avg ref cit+1)				0.1800*** (0.0004)		0.6491*** (0.0017)
ln(Max ref cit+1)					0.1398*** (0.0002)	-0.3900*** (0.0011)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25,263,987	25,263,987	25,174,032	25,262,880	25,262,880	25,174,032
Pseudo R2	0.07284	0.18198	0.09806	0.10322	0.10814	0.26989

Note: Robust standard errors are reported in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 7. The effect of new word combination reuse on citation count with reference-level controls.

Models	(1)	(2)	(3)	(4)	(5)	(6)
	DV: 5-year citation count (Poisson regression)					
ln(New words reuse+1)	0.0970*** (0.0003)	0.0741*** (0.0003)	0.0952*** (0.0003)	0.0888*** (0.0003)	0.0860*** (0.0003)	0.0597*** (0.0003)
ln(Ref count)		0.6078*** (0.0010)				0.8260*** (0.0011)
ln(Ref age+1)			-0.4181*** (0.0009)			-0.6545*** (0.0014)
ln(Avg ref cit+1)				0.1755*** (0.0004)		0.6402*** (0.0017)
ln(Max ref cit+1)					0.1363*** (0.0002)	-0.3860*** (0.0011)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25,263,987	25,263,987	25,174,032	25,262,880	25,262,880	25,174,032
Pseudo R2	0.08425	0.19005	0.10848	0.11290	0.11761	0.27463

Note: Robust standard errors are reported in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

5.2 Atypical combinations of knowledge

To measure atypical combinations of knowledge, we adopted the framework of Uzzi et al. (2013), analyzing the distribution of journal pairs within a paper's reference list. Each journal pairing was converted into z-scores, representing standardized measures of novelty. We quantified atypical combinations using Monte Carlo simulations, which generate reshuffled citation networks by randomly reassigning edges while preserving the temporal and distributional properties of the original network. The atypicality score for each journal pair is calculated as follows:

$$Z\ score_{i,j} = \frac{obs\left(pair_{ij}\right) - exp\left(pair_{ij}\right)}{\sigma\left(pair_{ij}\right)} \tag{5}$$

where $obs(pair_{ij})$ signifies the observed frequency of the journal pair in the real dataset, $exp(pair_{ij})$ represents the mean, and $\sigma(pair_{ij})$ denotes the standard deviation of the journal pairs obtained from 10 randomized simulations of the reshuffled network.

To summarize the distribution of z-scores, we extracted the 10th percentile value for each reference list. Given that the distribution of these values is highly skewed, we normalized atypicality scores into percentiles within each publication year and field (across 19 Level-0 fields in MAG).

Next, we analyze the relationship between atypicality (percentile) and the 5-year citation count using a basic Poisson regression model without additional controls. Split-sample regressions across publication years (1950–2016) revealed a consistently positive relationship between atypicality and citation count (Figure 7a). This association remained statistically significant at the 0.001 level for nearly all years, diverging from the trend observed for the CD index. Furthermore, as shown in Table 8, incorporating reference-level controls slightly reduces the effect size (especially the number of references) but does not alter the direction or significance of the association. Across all models, we find a robust, highly significant (0.001 level) positive relationship between atypicality and the 5-year citation count.

5.3 Disruptive citations

Beyond the CD index, which captures the relative proportion of disruptive impact, we employ disruptive citations to measure the absolute disruptive impact of papers, defined as:

$$Disruptive\ citation = n_i \tag{6}$$

Disruptive citations follow a power-law distribution (Yang et al., 2025), similar to other network-based metrics, providing a more granular perspective on a paper's disruptive impact. Unlike the CD index, this metric is independent of the focal paper's reference list, thereby eliminating potential biases introduced by the citing behavior of the paper itself (Yang & Deng, 2024; Yang, Hu, et al., 2023; Yang et al., 2025). Additionally, its broad range enables more precise differentiation between papers with varying levels of disruptive impact, making it a powerful tool for ranking scholarly contributions. Prior research has shown that disruptive citations effectively identify milestone papers and laureates (Lin et al., 2022; Yang, Hu, et al., 2023) while shedding

light on the relationship between scientific impact and disruption.

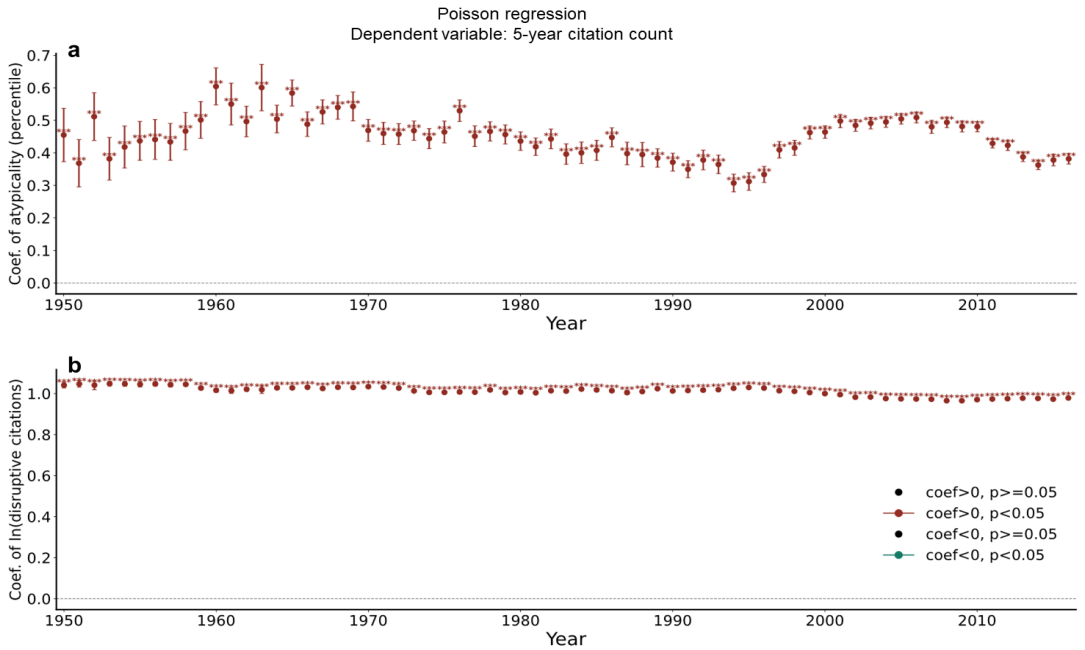


Figure 7. Yearly split-sample Poisson regression coefficients of atypicality and disruptive citations on citation count from 1950 to 2016. Error bars depict the upper and lower bounds of the 95% confidence intervals based on robust standard errors. Dark-red coloring indicates significant positive coefficients (p <0.05), and light-red coloring indicates non-significant positive coefficients. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 8. The effect of atypicality on citation counts with reference-level controls.

Models	(1)	(2)	(3)	(4)	(5)	(6)
	DV: 5-year citation count (Poisson regression)					
Atypicality (percentile)	0.4457*** (0.0018)	0.0249*** (0.0019)	0.5505*** (0.0017)	0.3110*** (0.0018)	0.2475*** (0.0019)	0.0217*** (0.0018)
ln(Ref count)		0.6557*** (0.0009)				0.8536*** (0.0011)
ln(Ref age+1)			-0.5450*** (0.0010)			-0.6462*** (0.0014)
ln(Avg ref cit+1)				0.1764*** (0.0004)		0.6479*** (0.0017)
ln(Max ref cit+1)					0.1407*** (0.0002)	-0.3834*** (0.0011)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,656,587	27,656,587	27,584,165	27,656,569	27,656,569	27,584,165
Pseudo R2	0.06756	0.17913	0.10371	0.09461	0.0997	0.26379

Note: Robust standard errors are reported in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

We examined the relationship between disruptive citations and the 5-year citation count using a Poisson regression model without additional controls. Our split-sample regressions across publication years (1950–2016) revealed a consistently positive relationship between disruptive citations and citation count (Figure 7b), with statistical significance at the 0.001 level for nearly all years. Moreover, as Table 9 demonstrates, incorporating reference-level controls did not alter the direction, effect size, or significance of the association. Across all models, we find a robust, highly significant (0.001 level) positive relationship between disruptive citations and the 5-year citation count.

Table 9. The effect of disruptive citations on citation count with reference level controls.

Models	(1)	(2)	(3)	(4)	(5)	(6)
	DV: 5-year citation count (Poisson regression)					
ln(Disruptive citations+1)	0.9791*** (0.0004)	0.9435*** (0.0004)	0.9657*** (0.0004)	0.9682*** (0.0004)	0.9663*** (0.0004)	0.8993*** (0.0005)
ln(Ref count)		0.1827*** (0.0004)				0.2734*** (0.0006)
ln(Ref age+1)			-0.2951*** (0.0005)			-0.3831*** (0.0007)
ln(Avg ref cit+1)				0.0497*** (0.0002)		0.1780*** (0.0007)
ln(Max ref cit+1)					0.0435*** (0.0001)	-0.0992*** (0.0005)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	29,009,690	29,009,690	28,888,580	29,007,831	29,007,831	28,888,580
Pseudo R2	0.70948	0.71983	0.71953	0.71123	0.71217	0.73775

Note: Robust standard errors are reported in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

6 Conclusion

This study illuminates the growing bias inherent in the CD index, revealing its negative correlation with citation impact and pinpointing the significant role of reference behavior in driving this trend. Our findings emphasize the critical need to account for referencing patterns when using the CD index to evaluate disruptive innovation. Direct comparisons of CD indices across papers with varying reference behaviors should be approached with caution, as these patterns substantially affect the index without necessarily reflecting the true disruptive nature of the research (Yang, Gong, et al., 2024).

Our analysis demonstrates that the observed negative correlation between the CD index and citation count stems largely from evolving scholarly practices, particularly the exponential increase in references per paper over time. However, when we control for reference count, either by integrating reference-level variables into our models or by conducting split-sample regressions on papers with identical reference counts, the relationship between the CD index and citation impact

shifts to a consistently positive one. This suggests that disruptive papers do indeed attract greater attention, challenging recent claims that their impact is diminishing (Zeng et al., 2023). However, by strictly controlling for reference count, our fixed effects regression method may also not adequately capture the citation dynamics of highly disruptive research, as lower reference counts may be a proxy for disruption that leads to fewer citations.

In contrast, alternative innovation measures, such as the introduction of new words, novel word combinations and their subsequent reuse, atypical knowledge pairings, and disruptive citations, consistently demonstrate positive associations with citation impact. Unlike the CD index, these metrics appear less sensitive to reference behavior because they focus on the content and knowledge structure of papers rather than their citation networks. This striking difference suggests that the bias observed with the CD index may be a methodological artifact tied to its specific calculation, particularly its dependence on the reference count, rather than a broader indication that disruptive research is losing influence. These findings highlight the value of complementing the CD index with alternative measures to achieve a more robust and nuanced evaluation of scientific innovation.

Our work cautions against interpreting the CD index's negative correlation with citation count as evidence that disruptive research is undervalued. Instead, we argue that the index's sensitivity to reference inflation, driven by the growing number of references in modern papers, can obscure the disruptive potential of highly cited works. This necessitates rethinking how the CD index can be applied. Researchers should explicitly control for reference counts to avoid misleading conclusions regarding the interplay between disruption and impact. Future studies could explore the development of a refined CD index, perhaps through normalization or weighting strategies that mitigate the influence of reference behavior. Such an improved metric would provide a clearer and more reliable gauge of scientific disruption.

These insights have significant implications for science policy and research evaluation. While the CD index remains a useful tool, its interpretation must be tempered by an awareness of referencing practices to prevent the misrepresentation of innovative contributions. Policymakers and evaluators might consider integrating alternative innovation measures that are less affected by reference patterns to foster a more accurate and holistic assessment of scientific impact. Failing to account for these dynamic risks skewing evaluations and misguiding funding or policy decisions.

Despite its contributions, this study had some limitations that warrant consideration. First, while fixed-effects models were used to address confounding factors, unobserved heterogeneity may still influence our results. Second, our reliance on the MAG dataset, though comprehensive, introduces potential biases or inaccuracies that could affect the findings. Additionally, by focusing solely on journal articles, our conclusions may not fully extend to other publication types, such as books or conference proceedings, which may exhibit distinct referencing behaviors. Third, the link between the reference count and the CD index rests on observed correlations; further research is needed to establish causality and clarify the mathematical interplay between these factors. Fourth, "disruption" remains a multifaceted concept, and while the CD index offers one lens, other metrics or qualitative approaches could provide complementary perspectives. Finally, our

aggregate-level analysis may mask variations across disciplines or research communities, suggesting the need for more granular future investigations.

In conclusion, this study advances our understanding of the CD index by highlighting the pivotal role of reference behavior and integrating insights from alternative innovation measures. Together, these findings advocate for a cautious, context-aware application of the CD index and underscore the need for more resilient metrics that can adapt to the evolving landscape of scholarly communication. By refining how we measure and interpret disruptions, we can better recognize and support transformative scientific contributions.

Acknowledgments

The study was conducted independently, and the author would like to express gratitude to the anonymous reviewers for their valuable insights.

Author contributions

Alex Jie Yang (alexjieyang@outlook.com; ORCID: 0000-0002-1276-0483): Conceptualization (Equal), Data curation (Equal), Formal analysis (Equal), Funding acquisition (Equal), Investigation (Equal), Methodology (Equal), Project administration (Equal), Resources (Equal), Visualization (Equal), Writing - original draft (Equal), Writing - review & editing (Equal).

Fanming Wang (527288991@qq.com): Conceptualization (Equal), Data curation (Equal), Formal analysis (Equal), Resources (Equal), Software (Equal), Writing - original draft (Equal), Writing - review & editing (Equal).

Yujie Shi (809297016@qq.com): Conceptualization (Equal), Resources (Equal), Software (Equal), Writing - original draft (Equal), Writing - review & editing (Equal).

Yiqin Zhang (miraclezyq@foxmail.com): Conceptualization (Equal), Formal analysis (Equal), Visualization (Equal), Writing - original draft (Equal), Writing - review & editing (Equal).

Hao Wang (ywhaowang@nju.edu.cn): Supervision (Equal), Validation (Equal), Writing - review & editing (Equal).

Sanhong Deng (sanhong@nju.edu.cn): Conceptualization (Equal), Supervision (Equal), Validation (Equal), Writing - original draft (Equal), Writing - review & editing (Equal).

Funding information

This research was supported by the Basic Research Program for Young Students (Doctoral Students) of the National Natural Science Foundation of China (#724B2015).

References

Arts, S., Melluso, N., & Veugelers, R. (2025). Beyond citations: Measuring novel scientific ideas and their impact in publication text. *The Review of Economics and Statistics*, 1-33. https://doi.org/10.1162/rest_a_01561

Bloom, N., Jones, C. I., Van Reenen, J., & Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, 110(4), 1104-1144. https://doi.org/10.1257/aer.20180338

Bornmann, L., Devarakonda, S., Tekles, A., & Chacko, G. (2020a). Are disruption index indicators convergently valid? The comparison of several indicator variants with assessments by peers. *Quantitative Science Studies*, 1(3), 1242-1259. https://doi.org/10.1162/qss_a_00068

Bornmann, L., Devarakonda, S., Tekles, A., & Chacko, G. (2020b). Disruptive papers published in scientometrics: Meaningful results by using an improved variant of the disruption index originally proposed by Wu, Wang, and Evans (2019). *Scientometrics*, 123(2), 1149-1155. https://doi.org/10.1007/s11192-020-03406-8

Chai, S., & Menon, A. (2019). Breakthrough recognition: Bias against novelty and competition for attention. *Research Policy*, 48(3), 733-747. https://doi.org/10.1016/j.respol.2018.11.006

Chu, J. S. G., & Evans, J. A. (2021). Slowed canonical progress in large fields of science. *Proceedings of the National Academy of Sciences of the United States of America*, 118(41), e2021636118. https://doi.org/10.1073/pnas.2021636118

Fortunato, S., Bergstrom, C. T., Boerner, K., Evans, J. A., Helbing, D., Milojevic, S., Petersen, A. M., Radicchi, F., Sinatra, R., Uzzi, B., Vespignani, A., Waltman, L., Wang, D., & Barabasi, A.-L. (2018). Science of science. *Science*, 359(6379), eaao0185. https://doi.org/10.1126/science.aao0185

Foster, J. G., Rzhetsky, A., & Evans, J. A. (2015). Tradition and innovation in scientists' research strategies. *American Sociological Review*, 80(5), 875-908. https://doi.org/10.1177/0003122415601618

Funk, R. J., & Owen-Smith, J. (2017). A dynamic network measure of technological change. *Management Science*, 63(3), 791-817. https://doi.org/10.1287/mnsc.2015.2366

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review of Economic Studies*, 76(1), 283-317. https://doi.org/10.1111/j.1467-937X.2008.00531.x

Jones, B. F., & Weinberg, B. A. (2011). Age dynamics in scientific creativity. *Proceedings of the National Academy of Sciences of the United States of America*, 108(47), 18910-18914. https://doi.org/10.1073/pnas.1102895108

Jones, B. F., Wuchty, S., & Uzzi, B. (2008). Multi-university research teams: Shifting impact, geography, and stratification in science. *Science*, 322(5905), 1259-1262. https://doi.org/10.1126/science.1158357

Leibel, C., & Bornmann, L. (2024). What do we know about the disruption index in scientometrics? An overview of the literature. *Scientometrics*, 129, 601–639. https://doi.org/10.1007/s11192-023-04873-5

Li, H., Tessone, C. J., & Zeng, A. (2024). Productive scientists are associated with lower disruption in scientific publishing. *Proceedings of the National Academy of Sciences of the United States of America*, 121(21), e2322462121. https://doi.org/10.1073/pnas.2322462121

Li, L., Lin, Y., & Wu, L. (2024). Displacing science. *arXiv.* https://doi.org/10.48550/arXiv.2402.16839

Lin, Y., Evans, J. A., & Wu, L. (2022). New directions in science emerge from disconnection and discord. *Journal of Informetrics*, 16(1), 101234. https://doi.org/10.1016/j.joi.2021.101234

Lin, Y., Frey, C. B., & Wu, L. (2023). Remote collaboration fuses fewer breakthrough ideas. *Nature*, 623(7989), 987-991. https://doi.org/10.1038/s41586-023-06767-1

Liu, X., Bu, Y., Li, M., & Li, J. (2024). Monodisciplinary collaboration disrupts science more than multidisciplinary collaboration. *Journal of the Association for Information Science and Technology*, 75(1), 59-78. https://doi.org/10.1002/asi.24840

Macher, J. T., Rutzer, C., & Weder, R. (2024). Is there a secular decline in disruptive patents? Correcting for measurement bias. *Research Policy*, 53(5), 104992. https://doi.org/10.1016/j.respol.2024.104992

Park, M., Leahey, E., & Funk, R. J. (2023). Papers and patents are becoming less disruptive over time. *Nature*, 613(7942), 138-144. https://doi.org/10.1038/s41586-022-05543-x

Petersen, A. M., Arroyave, F., & Pammolli, F. (2024). The disruption index is biased by citation inflation. *Quantitative Science Studies*, 5(4), 936-953. https://doi.org/10.1162/qss_a_00333

Radicchi, F., Fortunato, S., & Castellano, C. (2008). Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences of the United States of America*, 105(45), 17268-17272. https://doi.org/doi:10.1073/pnas.0806977105

Ruan, X., Lyu, D., Gong, K., Cheng, Y., & Li, J. (2021). Rethinking the disruption index as a measure of scientific and technological advances. *Technological Forecasting and Social Change*, 172, 121071. https://doi.org/10.1016/j.techfore.2021.121071

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, 342(6157), 468-472. https://doi.org/10.1126/science.1240474

Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365-391. https://doi.org/10.1016/j.joi.2016.02.007

Wang, J. (2013). Citation time window choice for research impact evaluation. *Scientometrics*, 94(3), 851-872. https://doi.org/10.1007/s11192-012-0775-9

Wang, J., Veugelers, R., & Stephan, P. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, 46(8), 1416-1436. https://doi.org/10.1016/j.respol.2017.06.006

Wu, L. F., Wang, D. S., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, 566(7744), 378–382. https://doi.org/10.1038/s41586-019-0941-9

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, 316(5827), 1036-1039. https://doi.org/10.1126/science.1136099

Yang, A. J. (2024). Unveiling the impact and dual innovation of funded research. *Journal of Informetrics*, 18(1), 101480. https://doi.org/10.1016/j.joi.2023.101480

Yang, A. J. (2025). Unraveling topic switching and innovation in science. *Information Processing & Management*, 62(4), 104171. https://doi.org/10.1016/j.ipm.2025.104171

Yang, A. J., & Deng, S. (2024). Dynamic patterns of the disruptive and consolidating knowledge flows in Nobel-winning scientific breakthroughs. *Quantitative Science Studies*, 1-17. https://doi.org/10.1162/qss_a_00323

Yang, A. J., Deng, S., Wang, H., Zhang, Y., & Yang, W. (2023). Disruptive coefficient and 2-step disruptive coefficient: Novel measures for identifying vital nodes in complex networks. *Journal of Informetrics*, 17(3), 101411. https://doi.org/10.1016/j.joi.2023.101411

Yang, A. J., Gong, H., Wang, Y., Zhang, C., & Deng, S. (2024). Rescaling the disruption index reveals the universality of disruption distributions in science. *Scientometrics*, 129(1), 561-580. https://doi.org/10.1007/s11192-023-04889-x

Yang, A. J., Hu, H., Zhao, Y., Wang, H., & Deng, S. (2023). From consolidation to disruption: A novel way to measure the impact of scientists and identify laureates. *Information Processing & Management*, 60(5), 103420. https://doi.org/10.1016/j.ipm.2023.103420

Yang, A. J., Xu, H., Ding, Y., & Liu, M. (2024). Unveiling the dynamics of team age structure and its impact on scientific innovation. *Scientometrics*, 129, 6127–6148. https://doi.org/10.1007/s11192-024-04987-4

Yang, A. J., Yan, X., Hu, H., Hu, H., Kong, J., & Deng, S. (2025). Are disruptive papers more likely to impact technology and society? *Journal of the Association for Information Science and Technology*, 76(3), 563-579. https://doi.org/10.1002/asi.24947

Yang, Y., Tian, T. Y., Woodruff, T. K., Jones, B. F., & Uzzi, B. (2022). Gender-diverse teams produce more novel and higher-impact scientific ideas. *Proceedings of the National Academy of Sciences of the United States of America*, 119(36), e2200841119. https://doi.org/10.1073/pnas.2200841119

Zeng, A., Fan, Y., Di, Z., Wang, Y., & Havlin, S. (2023). Disruptive papers in science are losing impact. *arXiv.* https://doi.org/10.48550/arXiv.2305.03589

Zeng, A., Shen, Z. S., Zhou, J. L., Wu, J. S., Fan, Y., Wang, Y. G., & Stanley, H. E. (2017). The science of science: From the perspective of complex systems. *Physics Reports*, 714, 1-73. https://doi.org/10.1016/j.physrep.2017.10.001



Copyright: © 2025 Alex J. Yang, Fanming Wang, Yujie Shi, Yiqin Zhang, Hao Wang, Sanhong Deng. Published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

Appendix material

-pdf file

The data can be accessed at: https://doi.org/10.57760/sciencedb.25215