Master Thesis
Data Science and Marketing Analytics
FEM11155-22

# Disruption in Marketing academia

*Author:*
B.H.P. van Zutphen (482262bz)

*Supervisor*:
Dr. N.M. Almeida Camacho

*Second assessor*:
Prof. Dr. D. Fok

*Abstract*

This paper investigates long-term trends in the disruptiveness of publications in the field of Marketing by employing a variation of the CD index of Funk and Owen-Smith (2017), which is selected based on a comprehensive review of disruption indicators. *Disruptive* is used to qualify science that breaks with existing trends and shifts the direction of future scientific research. Results show the least disruptive period in Marketing was around 1990-1995, after which the average disruption score rather consistently increases up until and including the final sample years. On the journal level, Marketing Science appears to produce the most disruptive papers, while the Journal of Marketing seems to produce the least disruptive studies on average. Results are verified by means of controlling regressions that correct for changes in publication, citation, and authorship practices over time.

October 20, 2023

# Contents

# 1 Introduction

Over multiple decades, despite greatly increasing research efforts, disruptive innovations in science have been declining. According to a recent study by Park et al. (2023), the indicator of disruptiveness (as defined by Funk and Owen-Smith (2017)) has decreased by over 90% for research papers from 1945 to 2010 and by over 75% for patents issued between 1980 and 2010.

Additionally, research productivity has been declining strongly, at an average of 5 percent per year (Bloom et al., 2020). This implies that the number of researchers has to double every 13 years to maintain any level of economic growth. A great illustration of this is Moore's Law, showing that 18 times more scientists are needed today to achieve the doubling in transistors compared to the 1970s (Bloom et al., 2020). In short, technological progress appears to be slowing down, negatively impacting disruptive science.

Furthermore, it's concerning that the disruptiveness of research papers and patents is decreasing, given that disruptiveness precedes innovation, which, in turn, drives productivity and economic growth (Segerstrom, 1991). An OECD (2021) report echoes these concerns, warning that the failure to support unconventional ideas could jeopardize a country's capacity to "compete economically, harness science for solving national and global challenges, and contribute to the progress of science as a whole".

The decline in disruptiveness found by Park et al. (2023) is based on the Consolidation-or-Destabilization (henceforth, CD) index, whose values range from 1 for the most disruptive, or destabilizing, to -1 for the most consolidating, and was introduced by Funk and Owen-Smith (2017). It is a citation-based metric that is built on the assumption that disruptive innovations will eventually take citation precedence over the papers those innovations are built on, while subsequent work of consolidating innovations will keep citing the papers supporting it. Hence, the metric is able to capture the direction and magnitude to which a new innovation consolidates or destabilizes the existing streams of knowledge.

Park et al. (2023) include articles from all scientific fields in their analysis to support their discovered pattern of declining disruptiveness. Other studies using data spanning all fields of science find comparable evidence, such as Uzzi et al. (2013), who investigate the related concept of novelty in science. Specifically, papers from the hard sciences (e.g., Bornmann et al., 2020 (biology & medicine); Lu et al., 2019 (biology)) appear to accumulate consistent data around this topic. Thus, few studies provide substantial quantitative evidence for the social sciences in this context, much less Marketing in particular. Similarly, and despite the tremendous recent growth[1], there exists no analyses on the disruptiveness of Artificial Intelligence (henceforth, AI).

[1] https://hai.stanford.edu/news/ai-spring-four-takeaways-major-releases-foundation-models

In this paper, I attempt to deploy the CD index of Funk and Owen-Smith (2017) to study the disruptiveness of specific fields. In light of recent studies that suggest Marketing Academia is becoming less impactful (MacInnis et al., 2020), less conceptual (Yadav, 2010) and has misaligned interests (Stremersch et al., 2021), I first investigate the top four journals in the field of Marketing: the Journal of Marketing (JM), the Journal of Marketing Research (JMR), the Journal of Consumer Research (JCR), and Marketing Science (MKS). These are widely regarded as the top four 'A' journals in the field of Marketing (Lehmann et al., 2011; Yadav, 2010).

Secondly, and principally as a benchmark for Marketing, I review four top-ranking journals in the field of AI: IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAML), the International Journal of Computer Vision (IJCV), AI[2] (AI), and the Artificial Intelligence Review (AIR). These journals have the highest impact factor in the Journal Citation Reports[3] of 2021 in the category 'Computer Science, Artificial Intelligence'. There exist a few journals that have higher impact factors, but these journals started publicizing too recently to be included in this study.

Because the CD index is a relatively novel measure, a new way to quantify disruptiveness, it has not yet been used to study these fields individually. The key stakeholders, i.e., individuals who can learn from this study, are science policymakers, journal editors, and everyone interested in and/or can influence disruptive innovation in science. Consequently, this paper is relevant to managers because it can support any decisions on the science policies that influence the disruptiveness of science.

However, the objective of this paper is not to make any policy recommendations to boost the share of disruptive innovations in any field. On the other hand, it is the goal to quantitatively and rigorously study the fields of Marketing and AI more in-depth and identify trends in disruptiveness, both on the sub-field level (journal vs. journal) and the field vs. field level. For the latter, AI should serve as a compelling benchmark to compare Marketing to since AI is a promptly advancing field, especially with the recent progress in generative AI[1]. With the help of (the most recently available) citation data, I investigate the disruptiveness of innovations in these fields, a purpose these data haven't served to my knowledge. This way, my research contributes to the existing research by filling a gap in the study of Marketing and AI literature. The resulting insights could be used in subsequent studies that, for example, aim to research science policies and trends.

Existing literature that studies trends in the progression of Marketing science includes papers by Yadav (2010) and Mela et al. (2013) for example. Yadav (2010) studies the number of

[2] https://www.sciencedirect.com/journal/artificial-intelligence
[3] https://jcr-clarivate-com.eur.idm.oclc.org/jcr/home

conceptual articles in major Marketing journals based on subjective qualifications while Mela et al. (2013) use the history of keywords in MKS to analyze the rate of innovation. This study, thereby, clearly differs from these papers as it uses an extensive dataset covering all top Marketing journals, including a relative comparison with the field of AI. Moreover, it employs an objective, citation-based metric computed on the most recent data to quantify the disruptiveness of each individual paper. Therefore, it should give a comprehensive view of the ways the field of Marketing forms new knowledge.

In conclusion, the research question of this paper is stated as follows: which trends in the disruptiveness of papers exist in the field of Marketing (compared to the field of AI)?

In the continuation of this paper, I first present a Literature Review: What is Disruptive Science?. Next, I give Research Background on the current trends in disruptiveness and explore possible arguments for the narrowing of science, both on the field level and the overall science level. Moreover, I offer Theoretical Background: Quantifying Scientific Innovation, which is an introduction to the techniques for quantifying scientific advancements based on citation data. Then, I discuss the necessary Methodology used in my research, focused primarily on the CD index, its shortcomings, and its variations. Furthermore, I describe the Data set in detail and present an exploratory analysis of citation and reference trends. In the next step, I provide and explain the produced Results. Finally, in the Conclusion, I interpret the Main findings, provide Managerial and academic implications and discuss the Limitations & further research. Results I do not find key to my research are shown in the Appendix.

# 2 Literature Review: What is Disruptive Science?

## 2.1 Disruptive innovation theory

In their paper *A reflective review of disruptive innovation theory*, Dan and Chieh (2008) present a timeline of the evolution of *Disruptive Innovation Theory*, which finds its origins in the innovation study of Schumpeter (1942). Based on the first theory of innovation by Marx and Engels (1848), Schumpeter (1942) developed the concept of *creative destruction*. He described innovation as the "process of industrial mutation that continuously revolutionizes the economic structure from within, incessantly destroying the old one, incessantly creating a new one". In short, Schumpeter (1942) thought of innovation as the novel synthesis of existing ideas, resources, and skills.

The American science philosopher Thomas Kuhn, in his seminal work *The structure of scientific revolutions*, was the first to challenge this *development-by-accumulation* view of scientific progress. Kuhn (1962) proposed that science follows a cyclic model that alternates between phases of *normal science* and *revolutionary science*. Long periods of normal science, characterized by incremental progress, eventually result in unresolved anomalies. These anomalies, in turn, trigger scientific revolutions that overthrow the existing streams of knowledge and thus create *paradigm shifts*. Kuhn (1962) suggested that by presenting new theories or methods, revolutionary research, although infrequent, is essential for the progress of science and society.

Interestingly, Dan and Chieh (2008) do not include the work of Kuhn (1962) in their developmental timeline of Disruptive Innovation Theory. The term *disruptive* would not be introduced until decades later by the American academic and business consultant C. M. Christensen (1997) who further developed and popularized the theory, albeit focused on the commercial aspect of technological innovation.

## 2.2 Definition(s)

Christensen coined the term *disruptive technologies* (Bower & Christensen, 1995), which he later (C. Christensen & Raynor, 2003) changed to *disruptive innovation*, highlighting the importance of the business model that focuses on a specific technology rather than the technology itself according to his concept of disruptive. In a more recent paper, C. Christensen et al. (2015) call attention to the common misuse of the idea of disruptive innovation (when referring to a product or service at a particular moment in time) because it is intended to refer to a process. Accordingly, Si and Chen (2020) rightly point out that disruptive innovation can stand for various ideas based on dissimilar perspectives, such as *disruptive technology innovation* (Bower and Christensen, 1995; C. M. Christensen, 1997) or *disruptive product innovation* (C. Christensen &

Raynor, 2003).

The above-mentioned interpretation of disruptive, however, is set in an evident Marketing context. In order to translate *disruptive technology* or *disruptive innovation* from a product and business perspective (according to C. M. Christensen (1997)) to a science context, it aids to be familiar with the categorization of scientific advances into dichotomous types in science of science research (Chen et al., 2021; Lin et al., 2022; L. Wu et al., 2022). One general class captures revolutionary, discontinuous, or disruptive discoveries, while the other class encompasses incremental, continuous, or evolutionary technologies (Dan & Chieh, 2008).

According to the theory of Kuhn (1962), revolutionary science is distinct from normal science in two key dimensions: it changes the direction of future research, and it has a high scientific impact on research, i.e., the recognition of the new paradigm. Revolutionary science frequently comes out of normal, non-revolutionary science, and it can therefore be complicated to differentiate between them (Casadevall & Fang, 2016). Moreover, revolutionary discoveries can be *technological* or *theoretical* in nature and can not only have a great long-term influence on the field from which they arose but also impact other fields or even spark the invention of new fields (Casadevall & Fang, 2016).

However, it is valuable to accentuate that according to the (Marketing) definitions of C. M. Christensen (1997), revolutionary innovation does not equal disruptive innovation; revolutionary innovation does not affect existing markets, while disruptive innovation either enters an existing market at the bottom or creates an entirely new market. Although widely adopted and applied, this definition of disruptive innovation by C. M. Christensen (1997) has triggered a debate in recent years (Si & Chen, 2020). Contrary to C. Christensen et al. (2015), Muller (2020) argues that Uber is a disruptive innovation because of its great lasting influence on the major stakeholders in the industry (i.e. competing producers, consumers, and service providers) and will (eventually) supplant the incumbent technology (or product or service).

Only recently have researchers initiated the use of the term *disruptive* to qualify scientific discoveries (L. Wu et al., 2019; Park et al., 2023). Interestingly, the paper of Funk and Owen-Smith (2017), the foundation of *Papers and patents are becoming less disruptive over time* by Park et al. (2023), does not once mention *disruptive*. Funk and Owen-Smith (2017), however, distinguish between *destabilizing* and *consolidating*. They, in turn, build on theories that differentiate between *competency-enhancing* and *competency-destroying* (Abernathy and Clark, 1985; Tushman and Anderson, 1986; C. M. Christensen, 1997). This famous classification , nonetheless, was also established from the company perspective (Dan & Chieh, 2008). Furthermore, L. Wu et al. (2019), based on the ideas of March (1991), suggest that the disruption of science entails the *solving* of scientific problems, contrary to the development of science by the *suggestion* of problems.

When one looks at the nomenclature of the word *disruptive*[4], they find that it means "the action of preventing something, especially a system, process, or event, from continuing as usual or as expected". This clearly integrates with the characterization of revolutionary science by Kuhn (1962) and, in particular, his concept of *paradigm shifts*, as repeatedly referred to in the context of disruptiveness (Funk and Owen-Smith, 2017; Park et al., 2023).

To summarize, the term *disruptive* is used to qualify science that breaks with existing trends and shifts the direction of future scientific research. Nonetheless, this does not necessarily mean a disruptive discovery receives a lot of followers and thus has a large impact (in terms of citations) (Funk & Owen-Smith, 2017).

## 2.3 Importance

With disruptive science precisely defined, it is meaningful to look at reasons why academic researchers should be taking disruptiveness into account. Davis (1971) defines *interesting* science as those theories that challenge, or even renounce, the audience's commonly held beliefs and are, therefore, the most impactful studies. This idea is reiterated by Shugan (2003), former editor-in-chief of MKS, who finds that research is interesting through its impact on the external audience. Moreover, he states that researchers should focus on the fundamental problems in the field since those problems have the ability to easily attract readers. This idea of destabilizing existing streams of knowledge is at the heart of disruptive research.

Kohli and Haenlein (2021) highlight the often misunderstood difference between *importance* and *relevance*. Namely, not all problems relevant to a field are important, without the vice versa being true. They recommend researchers to operationalize the importance of a topic by assessing the change in the behavior of related stakeholders, including the magnitude. This "status quo altering" property of science is another characteristic of disruptive works.

The upcoming section Trend in Marketing further emphasizes the importance and demand of disruptive research that challenges the status quo. Additionally, it elaborates on the Marketing-specific issues and possible solutions for creating important research.

## 2.4 Related concepts

Various related concepts to disruption exist to describe scientific innovation, which can easily be ambiguous. Therefore, this section provides context to the three most prevailing concepts.

[4]https://dictionary.cambridge.org/dictionary/english/disruption

A summary of these and previously mentioned concepts can be found in Table 1 below. In addition, the forthcoming section Theoretical Background: Quantifying Scientific Innovation highlights the importance of appropriate metrics to overcome these ambiguities.

Table 1: This table reports a summary of the definitions of science qualifications, i.e., constructs.

| **Construct** | **Definition** | **Main source(s)** |
|---|---|---|
| Disruptive | Breaks with existing trends and shifts the direction of future scientific research, without necessarily generating a large impact. | Schumpeter (1942); Funk and Owen-Smith (2017) |
| Revolutionary | Changes the direction of future research, and has a high scientific impact on research through recognition of new paradigms. | Kuhn (1962); Casadevall and Fang (2016) |
| Transformative | Causes radical shifts in the understanding of important concepts and/or induces the creation of new paradigms[5]. | Staudt et al. (2018) |
| Novel | Introduces new combinations between existing and often unrelated elements. | Uzzi et al. (2013); J. Wang et al. (2017) |
| Interesting | Challenges the audience's commonly held beliefs. | Davis (1971) |
| Important | Changes behavior of (many) related stakeholders | Kohli and Haenlein (2021) |
| Relevant | Involves field-related problems. | Kohli and Haenlein (2021) |

Firstly, the previously mentioned concept of *revolutionary* science has two characteristics (that set it aside from normal science) of changing the direction of future research and imposing the recognition of a new paradigm to the current streams of knowledge and thereby having a high impact (Kuhn, 1962; Casadevall and Fang, 2016).

*Transformative* research signifies the radical shift in the understanding of an important concept in science and/or induces the creation of new paradigms or fields of science[5]. A study needs to be radical, i.e. breaking from an existing paradigm, and impactful in order to be transformative, according to Staudt et al. (2018), who classify scientific work on the axes of *incremental* versus *radical* and low versus high *impact*. Staudt et al. (2018) claim the distinction between *incremental* and *radical* is in parallel to the distinction between normal and revolutionary science. However, this claim generates a contradiction because, according to Kuhn (1962), a revolutionary work needs to have a high impact by definition. Moreover, Staudt et al. (2018) present seven aspects of transformative research (each with a proposed metric): Radical-Generative, Radical-Destructive, Risky, Multidisciplinary, Wide Impact, Growing Impact, and Impact (overall).

Finally, and perhaps most importantly because of its broad adoption, the concept of *novel* research. Novelty describes the introduction of new combinations between existing and often

[5]https://new.nsf.gov/funding/learn/research-types/transformative-research

unrelated elements, such as scientific concepts or journals (Hofstra et al., 2020; Uzzi et al., 2013; J. Wang et al., 2017). This combinatorial perspective of novelty links back to the innovation theory of Schumpeter (1942) and is widely supported throughout science (Weitzman, 1998). Novelty is a fundamental ingredient for generating creative ideas since it is regarded as one of its two principal components, in addition to impact (Lee et al., 2015; Uzzi et al., 2013). Accordingly, Hofstra et al. (2020) defines *impactful novelty* to capture the uptake of a new conceptual link and refers to novelty as *conceptual novelty.* Nonetheless, both disruption and novelty are classifications for science that capture a shift in the formation of knowledge.

## 3 Research Background

### 3.1 Trend in science (overall)

#### 3.1.1 Citation-based evidence

Park et al. (2023) provide evidence, based on the $CD_5$ index by Funk and Owen-Smith (2017), of a decline in disruptiveness of papers across all fields of science, categorizing them into *life sciences and biomedicine*, *physical sciences*, *social sciences*, and *technology*, see Figure 1. Since 1945, the decline has been the lowest for the *life sciences and biomedicine*, with the *physical sciences* following thereafter. Starting around 1995 (and until 2010), the average disruptiveness of a paper has essentially been constant for these two categories. Contrarily, it has been declining consistently for the final 15 years of the time series regarding the *social sciences* and *technology*, the two categories that have experienced the largest drop since 1945.

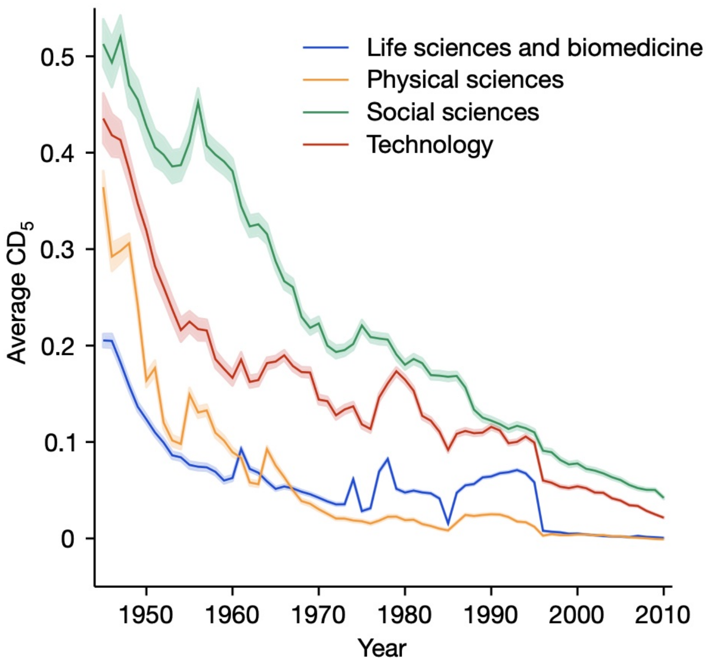



Figure 1: This figure shows the average $CD_5$ index for the four science categories. The subscript 5 indicates a citation window of 5 years.
Adapted from Park et al. (2023).

This decline in the average disruptiveness of a paper is not driven by a fall in the number of disruptive papers, but rather a tremendous increase in the number of incremental papers, Park et al. (2023) find. The number of highly disruptive papers (i.e., $CD_5 > 0.25$) has remained relatively flat, with a minor exception of the exceptionally disruptive papers (i.e., $CD_5 > 0.75$). The number related to the latter grew at a slow but reasonably steady rate in the period 1945-1995 but dropped sharply around 1995, resulting in an absolute number of exceptionally disruptive papers in 2010 that is comparable to that of 1945. The number of slightly disruptive papers (i.e., $0 < CD_5 < 0.25$), on the other hand, has expanded vastly from only a few thousand in 1945 to almost 200,000 per year in 2010.

This dramatic increase coincides with the exponential but stable growth in the number of papers and the number of authors found by D. Wang and Barabási (2021). This pattern of the expansion of scientific knowledge is also documented by Park et al. (2023). They detect the largest increase in research volume in the *life sciences and biomedicine* category, followed by the *physical sciences* and *technology*, respectively. The number of published papers in the *social sciences* appears to be growing at the slowest rate.

To support their found pattern, Park et al. (2023) reference studies investigating novelty in all fields of science (Uzzi et al., 2013; Hofstra et al., 2020). Although the concepts are undeniably correlated, this example of citing illustrates the lax application of the term *disruptive*. Employing the 'atypical combinations' measure of Uzzi et al. (2013) on their data, Park et al., 2023 detect a declining trend in the novelty of papers (i.e., combinations of citing prior studies) that is consistent with their findings using the CD index.

### 3.1.2 Alternative evidence

All the above-mentioned evidence is based on citation counts and networks that are inherently built upon citation practices. These practices form the basis for the methods that scientists use to assess themselves and thus are prone to biases that lead to possibly quantifying scientific impact with noise (Bornmann et al., 2008; Park et al., 2023). A measure that is less prone to such potential biases evaluates the number of unique topics investigated by a certain field.

Milojević (2015) introduced a measurement to assess the number of different research ideas a field is pursuing that is unbiased by the volume of output. It should represent the *cognitive extent of science* and counts the number of unique phrases among 10,000 article title phrases. Milojević characterizes the measure as an indicator of the pace at which new rungs are added to the 'ladder of science' (rather than increasing the width of the ladder). She updated the results by using data until 2020 as part of a recent report from the OECD (2023), finding that there has been stagnation in the cognitive content of scientific publications since the mid-2000s across all of science, see Figure 2. This indicates that an increasing amount of studies are focusing on similar ideas, which in turn suggests that the expansion of the frontiers of knowledge has gotten harder. Nonetheless, the overall long-term trend is increasing, which challenges the findings of Park et al. (2023).

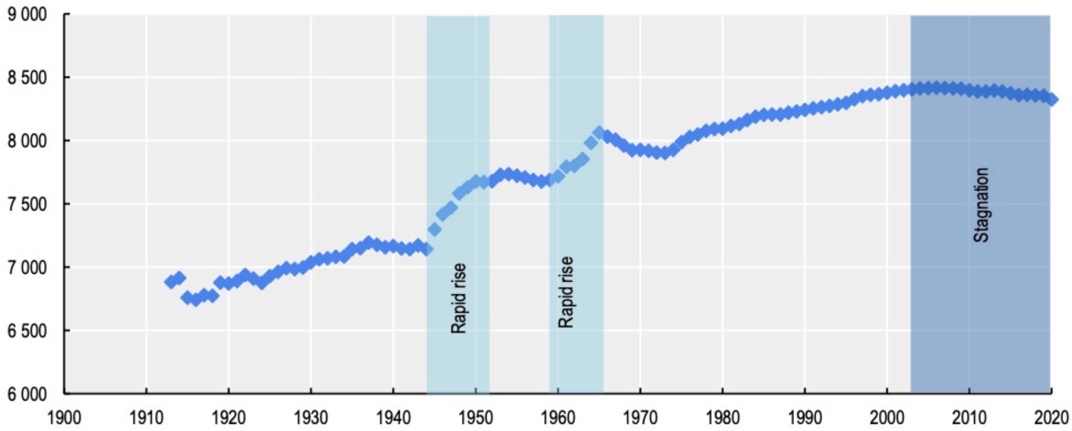



Figure 2: This figure shows the "cognitive extent" of science, referring to the number of unique phrases among 10,000 article title phrases.
Adapted from OECD (2023) report.

#### 3.1.3 Potential explanations

The upcoming section Trend in Marketing provides possible field-specific explanations for a declining trend in disruptiveness. Additionally, this section aims to present several science-wide arguments.

Jones (2009) suggests that the reduction in the innovative capability of individuals can be (partly) attributed to the so-called 'burden of knowledge'; one needs to accumulate an increasing amount of knowledge as technology advances to rise to the frontier of a scientific field. One could overcome this burden by extending education or narrowing expertise. This, in turn, transforms the nature of innovating, for example, being more dependent on working in (larger) teams (Jones, 2009).

This increased need for teamwork subsequently impacts the share of disruptive science as different team sizes tend to produce different types of innovations. L. Wu et al. (2019) find that large teams are likely to build on existing knowledge while smaller teams favor chasing disruptive innovations and novel ideas. They point toward the demand for interdisciplinary approaches to take on the present-day challenges and the above-mentioned specialization of scientific research as reasons for the increase in team size. Moreover, Chu and Evans (2021) point out that an abundance of papers in a field can slow down progress because readers and reviewers have increasingly less time to really grasp novel ideas.

Since trends show that research is progressively done in larger teams across many fields (Wuchty et al., 2007), this larger reliance on teams negatively impacts disruptive science. Correspondingly, Lee et al. (2015) find that an increasing team size has an inverted-U-shaped relation with novelty.

### 3.2 Trend in Marketing

No hard, quantitative evidence exists that the field of Marketing Academia is getting less disruptive, i.e., less destabilization of the existing Marketing research directions. Nonetheless, the previous section Trend in science (overall) hints at a declining trend in Marketing based on the data regarding the *social sciences*, the branch of science encompassing the interdisciplinary Marketing area. The presented citation-based evidence for the *social sciences* indicates a drop in average disruptiveness and (combinatorial) novelty of a paper over time and the slowest growth in scientific publications (out of the four mentioned categories).

Quantitative data concerning the specific field of Marketing with regard to disruption are sparse. Such a sparse example is the study of Yadav (2010), who analyzed 30 years (1978-2007) of publication data from the four major Marketing journals JM, JMR, JCR, MKS, and an additional highly ranked, broad-based Marketing journal; the Journal of the Academy of Marketing Science (JAMS). He finds that the share of conceptual articles, defined as a contribution that concentrates primarily on theory advancement (without relying on data), is falling.

As a consequence, Yadav (2010) argues the advancement of the field of Marketing has stalled because of their provision of new ideas and disproportional influence in terms of, e.g., citations and awards (compared to empirical articles). MacInnis (2011) emphasizes the vital importance of conceptual advances for the continuity of the Marketing discipline by serving as an important part of the knowledge development process.

Various studies highlight the increased need for conceptual articles in the field of Marketing. As part of a virtuous cycle whereby firms can create competitive advantages, Marketing Academia calls for the development of theoretical paradigms on the use of new technologies (Hoffman et al., 2022). Furthermore, Jaakkola (2020) expresses the acknowledgment of the major journals for this demand, influenced by the fact that such papers account for a large share of the most impactful papers. This integrates with the idea of Shugan (2003), who finds that the best way to create impact is through studying those fundamental topics.

Yadav (2010) finds that the sharpest decline in conceptual articles (until 2007) occurred in the Journal of Marketing, the long-time leader in publishing such studies. The average share of conceptual articles was 22.33% during the study period, moderately growing throughout the 1970s and 1980s and reaching a peak in 1988 at 50%. In 1993, the percentage started declining at a reasonable rate, arriving at a low of 6.70% for the most recent period, 2003-2007, included in the study.

Interestingly, and possibly consequently, multiple generations of editorial leadership of JM have curated a series of articles to combat this trend. In light of rapid technological advancements and Marketing practices changing faster than Marketing research being published, Moorman et al. (2019b), JM editors 2018-2022, commissioned a directed series of articles to use "different theoretical and methodological traditions designed to disrupt traditional Marketing doctrine and to open up new areas that we believe deserve the field's attention." They underline the importance of conceptual articles, stating that they have often disrupted the directions of Marketing research. Moreover, Moorman et al. (2019a) adopted the editorial mission[6] to "challenge the boundaries of the Marketing discipline by publishing articles that advance new research questions designed to disrupt traditional Marketing doctrine and to open up new areas of the discipline."

The current editors of JM, Sridhar et al. (2023), appear to continue down this path by encouraging researchers to "take on novel problems, with diverse data and approaches, and move beyond well-established areas and well-known concepts." In addition, they plan to introduce a special issue *New Paradigms for a New World* that seeks researchers to "reimagine the world of Marketing in the wake of recent significant disruptions." These disruptions, such as swift advances in AI and the acknowledgment of broad social and economic disparities, have fundamentally changed consumer behavior and Marketing practices (Sridhar et al., 2023).

The remaining four journals vary considerably in their share of conceptual articles, as reported by Yadav (2010). JMR published the least of such studies out of the major journals at 2.34%, which is unsurprising given its historical concentration on empirical research (Yadav, 2010). The proportion of conceptual articles in JCR fluctuated lightly over the investigated period, averaging 7.47% but was 2.58% for the final term. MKS has the highest average share among the studied journals at 29.22%, which is remarkable considering its quantitative nature. Yadav (2010) finds that these conceptual articles in MKS are often analytical studies focusing on the development of theory instead of theory testing. Finally, JAMS published, on average, 22.88% conceptual papers. However, similar to some of the other journals, it experienced a large drop to 11.48% during the last period of study.

In summary, Yadav (2010) finds a declining trend in the number of conceptual papers in the field of Marketing, although most recent data are not included. Multiple studies recognize the importance of such articles in the development and disruption of knowledge, including the above-mentioned publications of JM's editorial leadership. These editorials strongly express a need for the disruption of the existing research paths pursued by Marketing scholars and thus imply the current streams of knowledge fail to achieve just that.

[6] https://www.ama.org/guiding-editorial-principles-for-the-journal-of-Marketing/

### 3.2.1 Potential explanations

Jaakkola (2020) and Ulaga et al. (2021) find that researchers are struggling with proposals and the development of non-empirical articles in the absence of widely accepted and systematic guidelines. Therefore, researchers often avoid proposing and writing conceptual articles, which is detrimental to those who believe in developing theory (Ulaga et al., 2021). Additionally, focusing on the domain of consumer research, MacInnis et al. (2020) discover that the domain is often unsuccessful at generating a large impact since researchers adhere to implicit boundaries concerning the problems that can be studied, for which reasons, and how to do so. These boundaries hinder the identification of novel ideas that are the essence of growth for the Marketing discipline (MacInnis et al., 2020).

Other studies direct attention to the misaligned incentives that Marketing scholars face. Stremersch et al. (2021) discover that features like creativity, literacy, and relevance hold insufficient weight while the total output receives too much weight. Nonetheless, one should take into account the essential difference between relevance and importance, for which the latter should indisputably be the most influential when assigning weights (Kohli & Haenlein, 2021), which Stremersch (2021) reaffirms.

Moreover, Stremersch et al. (2021) find that Marketing faculty members feel inadequately compensated for their research efforts. Similarly, Reibstein et al. (2009) argue the interests of Marketing academics and practitioners have become increasingly divergent. Better aligning these interests, i.e., theoretical advancement and practical importance, will positively impact the long-term health of the field.

On a different note, Lehmann et al. (2011) suggest rigor has received too much attention from Marketing researchers. Therefore, rigor has often evolved to be the main goal of one's study, disregarding advantageous components similar to those identified by Stremersch et al. (2021) such as relevance and simplicity. The main argument brought forward by Lehmann et al. (2011) is that a paper can be sophisticated while not being complicated, referring to *Occam's razor*, i.e., the simpler solution is preferred. MacInnis et al. (2020) affirm this sentiment but add the nuance that rigor, in addition to relevance, is required to create high-impact, Marketing-relevant research.

## 3.3 Marketing vs. AI comparison

This section provides a succinct comparison of the research background between the fields of AI and Marketing. Section Research Background: Trend in AI in the Appendix offers more in-depth AI-specific evidence.

It appears Marketing researchers are more conscious of the content they put out with regard to advancing the field as a whole, based on the many editorials and review articles published (see section Trend in Marketing). This is not to say AI researchers do not strongly evaluate their potential research topics, there are simply few publications I have been able to find that discuss such trends. Nonetheless, from the evidence that is available, both fields seem to struggle with a narrowing of their research efforts.

Matters in the field of Marketing presumably causing a decline in conceptual advances are mostly related to misaligned interests and incentives (Lehmann et al., 2011; MacInnis et al., 2020; Stremersch et al., 2021). Moreover, Marketing scholars express that there exist few guidelines for identifying and proposing research topics, in addition to rigor being valued too much (over importance) (Jaakkola, 2020; MacInnis et al., 2020; Ulaga et al., 2021). Contrarily, the narrowing of the field of AI looks to be substantially driven by interests from outside the academic sector, i.e., the private sector (Ahmed and Wahed, 2020; Klinger et al., 2020; Whittaker, 2021).

# 4 Theoretical Background: Quantifying Scientific Innovation

This section explores techniques for quantifying scientific advancements. In particular, it covers the relationship between indicators of science and citations, in addition to trends in citation practices. Furthermore, it includes a discussion on the relevant matters to consider when implementing these techniques. These topics involve publication and citation selection criteria, as well as methods to correct for changes in publication, citation, and authorship practices.

## 4.1 Citations

Garfield (1964) introduced the seminal metric of science, the *Citation Index*, now embodied as the proprietary Web of Science[7] (WoS). Many evaluative metrics have been developed following the introduction of the citation index. Metrics that employ such citation data can be categorized into ones that measure research impact, rooted in citation counts, and ones that quantify scientific direction, rooted in citation networks.

## 4.2 Citation-based indicators

### 4.2.1 Impact indicators

The more common application of citations is to assess the impact of scientific research (Abramo and D'Angelo, 2011; Aksnes et al., 2019; D. Wang et al., 2013). Waltman (2016) distinguishes between five basic citation impact indicators, with a distinction between *size-dependent* and *size-independent* indicators, on which the majority of variants or extensions of indicators are built. The size-dependent indicators aim to deliver a comprehensive performance measure and include the *total number of citations*, the *number of highly cited publications*, and the *h-index* (Hirsch, 2005). Conversely, the size-independent indicators address the average performance per publication and consist of the *average number of citations per publication* and the *proportion of highly cited publications*.

Compared to individual publications, citation impact factors more commonly provide information on research entities, including researchers, research institutions, or journals (Waltman, 2016). The above-mentioned *h-index* is such a measure, and possibly an even more widely recognized example is the *journal impact factor* (Garfield, 1972).

[7] https://clarivate.com/products/scientific-and-academic-research/research-discovery-and-workflow-solutions/webofscience-platform/

### 4.2.2 Direction indicators

Contrary to working with mere citation counts, direction indicators make use of the structure of citation networks to measure scientific direction. Garfield (1955) realized at an early stage that citations can be used to assess the influence of a specific work on the literature and intellectual discourse of a given period. The study of Uzzi et al. (2013) is one of the earlier works that employ citation structure to develop an indicator. By measuring the frequency of atypical, pairwise combinations of references, Uzzi et al. (2013) quantify the novelty of a research paper. However, as the definition of novelty implies, such a novelty indicator purely takes into account how researchers integrate existing knowledge from the input side without being able to quantify the extent to which the focal paper (FP) leads to a shift in research directions.

Funk and Owen-Smith (2017) developed an indicator that does measure the extent to which a scientific advancement destabilizes the current paths pursued by researchers; the CD index. It is based on the idea that disruptive innovations will eventually take citation precedence over the papers those innovations are built on. Formally, L. Wu et al. (2019) adapted the CD index for its application on scientific papers and termed it the *disruption index*.

### 4.2.3 Disruption indicators

The prior section Literature Review: What is Disruptive Science?, discussing the definition of disruptive and its related ideas, points out the ambiguities between similar concepts. Therefore, it is crucial to understand and overcome these differences in meaning in order to come up with the appropriate metrics.

Bornmann et al. (2020) suggest disruption indicators have been developed based on the earlier introduced family of novelty indicators. Nonetheless, novelty indicators are not able to capture any shifts in research direction after publication, which is the essence of disruption indicators. This is the case because disruption indicators incorporate both the references of the FP and the references of its citing papers, whereas novelty indicators are based entirely on the cited references of the FP.

Moreover, disruptive research is often wrongly equated to revolutionary science because it encompasses the 'changing the direction of future research' aspect. It does, however, not necessarily involve the second feature of revolutionary science that should generate a large impact, typically measured by the number of citations. It is important to note that the disruption index is unable to assess whether the focal paper disrupts current research; such a determination can only be made after a certain number of years.

The upcoming section Disruption indicators of the Methodology chapter goes into detail on the disruption index and its proposed adjustments.

### 4.3 Relationship between science classifications and citations

Although the classifications of science, such as disruptive and novel have differences in meaning, they do exhibit correlation.

Lin et al. (2022) find that novel papers, measured through atypicality as defined by Uzzi et al. (2013), are almost twice as likely to cause disruption in science compared to conventional papers. However, this is a gradual process that takes ten years or more for disruption scores to converge. Ruan et al. (2023) confirm that such topic-combination novelty has a positive effect on disruption. Additionally, they identify an inverted U-shaped effect of novelty on the impact of scientific papers.

J. Wang et al. (2017) discover that novel papers are more likely to be among the top 1% highly cited papers in the long run. In general, novel papers are two times more likely to be highly cited (Uzzi et al., 2013). Moreover, these novel papers are cited in a wider array of scientific fields and fields that are further removed from their primary domain (J. Wang et al., 2017).

### 4.4 Trends in citation practices

Park et al. (2023) document a downward trend in the use of existing knowledge among researchers. They detect a decline in the level of semantic diversity of cited scientific work in parallel with a significant increase in the share of citations to the top 1 percent of most cited papers. However, Bornmann and Mutz (2015) find that scientists today are likely to cite a higher number of existing works than previous generations of researchers.

Furthermore, Park et al. (2023) identify an increase in the mean number of self-citations, a commonly used proxy for measuring the extent of continuing one's research directions, and the mean age of a cited work, a frequently used proxy for assessing the use of dated knowledge. Nonetheless, a survey by Teplitskiy et al. (2022) suggests citations to older work show no significant distinction from citations to younger work.

In summary, all the above-mentioned patterns point toward the overall trend of a narrowing range of existing knowledge used by scientists.

### 4.5 Selection criteria

For any citation-based indicator, it is essential to rigorously select the publications and citations to be considered its the computation. Waltman (2016) lists the following important criteria.

### 4.5.1 Publication selection

The *document type* (e.g., *article* or *editorial material*) is a standard criterion for the exclusion of publications and holds particular significance to size-independent indicators.

Secondly, the *language* of a publication can serve as a selection criterion (e.g., Van Raan et al., 2011) and is, similar to the document type, mainly relevant for size-independent indicators.

In addition, one can pick out *(inter)national journals* (e.g., Waltman and van Eck, 2013) in order to exclusively make comparisons on the (inter)national level.

### 4.5.2 Citation selection

Regarding the selection of citations, the inclusion or omission of *self-citations* is the heaviest debated topic (Costas et al., 2010). Self-citations are possible at different levels, such as the *journal* and *research institute* levels but more importantly, the *author* and *co-author* levels.

Finally, one should take into account the *citation window* when calculating indicators since they are dynamic measures that change over time. The adoption of a particular citation window can result in the exclusion of both publications (i.e., by requiring the existence of an x-year citation window) and citations (i.e., by omitting citations x-years after publication) from the computation of citation impact indicators.

Bornmann and Tekles (2019) find that the disruption index indeed depends on the length of the citation window and specify that a citation window of at least three years is necessary to generate relevant insights. Park et al. (2023) compute the CD index based on the five years after the year of publication, based on the fact that most works peak in their annual citation number within this time window (Funk and Owen-Smith, 2017; Jaffe and Trajtenberg, 2002). Alternatively, one can consider all forward citations of an FP as of a certain year (while still enforcing a minimum of three years).

The research on citation window length is tightly connected with the study of *delayed recognition* (Van Raan, 2004); the idea that a publication receives a majority of its (forward) citations long after its publication date. Glänzel et al. (2003) find that the length of the window should not have a large influence on citation impact indicators.

Nonetheless, there exist concerns that conclusions based on the CD index could be dependent on the length of their citation windows (Liang et al., 2020) and even produce contradictory results. Therefore, I choose to study multiple citation window lengths in order to test if the outcome is sensitive to such a choice.

## 4.6 Techniques to correct for changes in publication and citation practices over time

The number of publications has sharply expanded (D. Wang & Barabási, 2021) as well as the average number of citations made in a single publication (Bornmann & Mutz, 2015). Therefore, modern papers are increasingly likely to co-cite a focal paper and one of its references. This affects the robustness of citation-based (disruption) indicators. In particular, the comparison of indicator values over time might be biased (Park et al., 2023).

### 4.6.1 Normalization approaches

To correct for changes in these practices, one can employ *normalization*. The practice of normalizing citation-based indicators is an essential concept of citation analysis that facilitates field versus field comparison and, perhaps more importantly, comparisons across time. In their *Manifesto for research metrics*, Hicks et al. (2015) reiterate the need for dealing with variations in publication and citation practices. A simple normalization approach, for example, is dividing the actual number of citations by the expected number (e.g., based on the average number of citations in a field and year).

In general, one can differentiate between cited-side and citing-side normalization methods. *Cited-side* normalization is based on the cited papers, while *citing-side* normalization is based on the citing papers. Many papers in bibliometric research focus on this topic and have produced conflicting results, concluding that citing-side normalization performs better than cited-side normalization and vice versa (Bornmann and Marx, 2015; Waltman, 2016).

The fundamental idea of the more traditional approach of cited-side normalization is to compare a focal paper with an expected value based on, for example, field and publication year (Bornmann & Marx, 2015). These methods require a *field classification system*, for which the 250 WoS subject categories[8] are often used (Waltman, 2016).

In contrast, citing-side methods (Zitt & Small, 2008) do not require a field classification system and can therefore be regarded as advantageous over cited-side methods (Waltman & van Eck, 2013). These approaches aim to correct for differences in the length of reference lists (of citing papers) between fields (Waltman, 2016).

### 4.6.2 Regression

In addition to normalization, one can employ regressions that adjust for changes in publication and citation practices. Such regressions may include controls for field and/or year factors and

[8] https://images.webofknowledge.com/images/help/WOS/hp_subject_category_terms_tasca.html

accordingly provide for more robust comparisons over indicator values over time (Park et al., 2023).

#### 4.6.3 Counting methods

Finally, one can consider how to assign credit to individual co-authors of a publication. Since the average number of authors per paper is rising (Wuchty et al., 2007), Waltman (2016) points out the fact that it has become increasingly difficult to justly assign credit. Most citation-based indicators do not (yet) take differences into account between co-authors and therefore make use of *full counting* and allocate full credit to each individual author. Other counting methods include fractional counting and allocating credit based on an author's position in the author list (Waltman, 2016).

# 5 Methodology

## 5.1 Disruption indicators

### 5.1.1 Original indicator

Funk and Owen-Smith (2017) built the CD index based on the concept of *tripartite networks* or *tripartite graphs*. Such graphs represent a network consisting of three types of node categories, $V_1$ $V_2$, and $V_3$, and their edges (i.e., connections) $E$: $G = (V_1, V_2, V_3, E)$. Following the notation of Funk and Owen-Smith (2017), $V_1$ represents the FP, $f$; $V_2$ includes references cited by the FP, $b$; $V_3$ consists of a set of subsequent papers that cite the FP and/or references cited by the FP, $i$. The (directed) edges in a network serve as the citations between papers, and because of the inherent chronological nature of citations, these networks are *acyclic*.

The nodes $f$ and $b$ remain fixed (at the publication date of the FP), while the content of set $i$ can increase over time as the FP and/or references cited by the FP accumulate new forward citations. A new paper can join set $i$ in three distinct ways: (1) it cites the FP and is therefore of type $f$, (2) it cites one or more of the FP's references, i.e, type $b$, or (3) is of both type $b$ and $f$. Figure 3 displays a graphical representation of three exemplary citation networks.

Accordingly, Funk and Owen-Smith (2017) define the $CD_t$ index at time $t$ for an FP and its corresponding vector of $n$ subsequent papers that cite the FP and/or references cited by the FP $i = (i_1, i_2, \ldots, i_n)$ as

$$CD_t = \frac{1}{n_t} \sum_{i=1}^{n} \frac{-2f_{it}b_{it} + f_{it}}{w_{it}}, w_{it} > 0, \tag{1}$$

where

$$f_{it} = \begin{cases} 1 & \text{if } i \text{ cites the FP (type } f\text{)}, \\ 0 & \text{otherwise}, \end{cases} \tag{2}$$

and

$$b_{it} = \begin{cases} 1 & \text{if } i \text{ cites any reference cited by the FP (type } b\text{)}, \\ 0 & \text{otherwise}, \end{cases} \tag{3}$$

where $w_{it}$ is an optional weighting parameter that, by default, is set to 1 for simplicity but could be employed to incorporate differences in citation importance based on factors such as the age of citations.

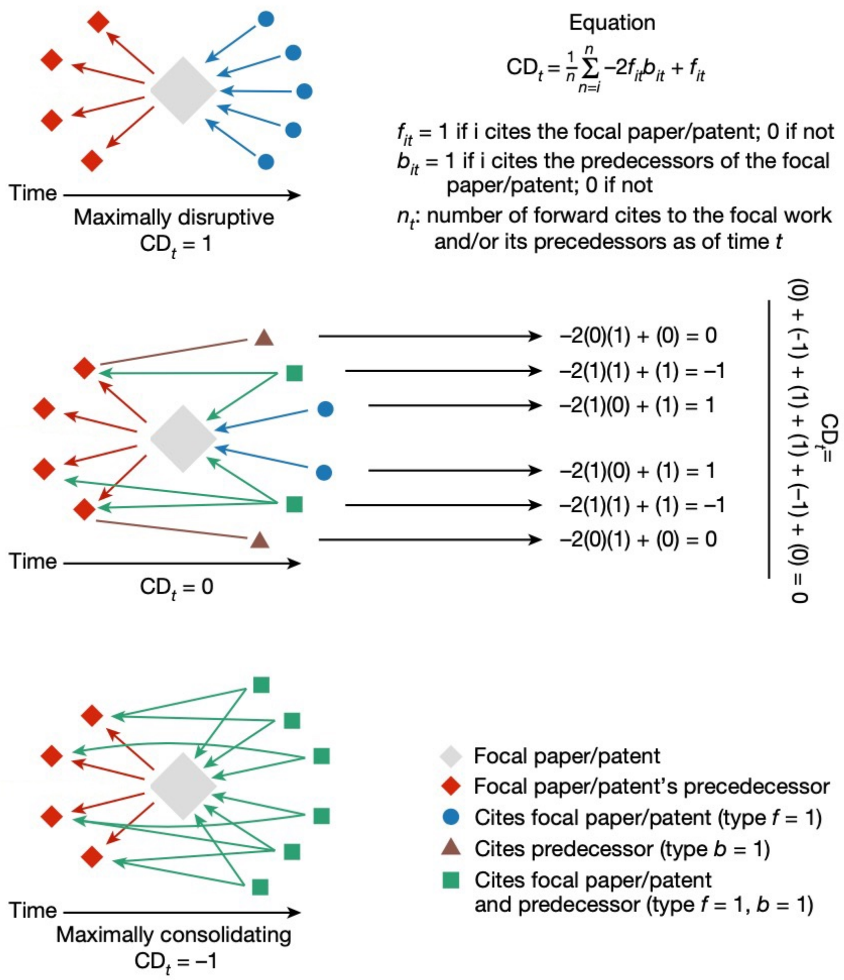



Figure 3: This figure depicts a graphical representation of three exemplary citation networks. Adapted from Park et al. (2023).

L. Wu et al. (2019) rewrote the seemingly complicated Equation 1 in a simpler manner, though denoted following the notation of Funk and Owen-Smith (2017) that makes use of the types $f$ and/or $b$:

$$DI = \frac{N_f - N_{bf}}{N_f + N_{bf} + N_b}, \tag{4}$$

where

$$N_f = \ |\{f_1, f_2, \dots\}|, \quad N_b = \ |\{b_1, b_2, \dots\}|, \text{ and } \quad N_{bf} = \ |\{bf_1, bf_2, \dots\}| \tag{5}$$

Equation 4 now simply expresses, in the nominator, the difference between the number of subsequent papers that cite the FP but not any of its references (i.e., $N_f$) and the number of papers that cite both the FP and at least one of its references (i.e., $N_{bf}$), and in the denominator the total number of papers that cite the FP and/or references cited by the FP. Since Equation

4 allows for a more accessible discussion of the different components, I will continue using this form.

The $N_{bf}$ term in Equation 4 measures the extent to which papers that cite the FP *bibliographically couple* (Kessler, 1963) the FP with its references, otherwise known as a *co-citation* (i.e., two papers cite the same paper). Leydesdorff et al. (2021) refer to this coupling as a signal of *continuity* throughout multiple generations of citations, and a higher value of the $N_{bf}$ term is, therefore, a sign of consolidation. Subsequently, if ,e.g., $N_f = 0$ and $N_b = 0$ with $N_{bf} > 0$ in Equation 4, $DI = -\frac{N_{bf}}{N_{bf}} = -1$ and therefore maximally consolidating. In the other extreme example that $N_b = 0$ and $N_{bf} = 0$ with $N_f > 0$, $DI = \frac{N_f}{N_f} = 1$ and therefore maximally disruptive. See Figure 3 for a graphical representation of these extreme cases.

#### 5.1.2 Distinct disruption and consolidation indicators: $CD^*$ and $CD^{\#}$

Leydesdorff et al. (2021) argue that the CD index is more akin to an indicator of continuity rather than disruption since it is rooted in bibliographic couplings. From that viewpoint, disruption is achieved when continuity is not sufficiently produced by the bibliographic coupling of an FP with the references that it cites. Nonetheless, Leydesdorff et al. (2021) suggest that the use of two words for a single indicator, with an opposite sign for the two dimensions, can lead to confusion in semantics.

Chen et al. (2021) deal with this issue by re-conceptualizing disruption and consolidation as two distinct dimensions of a scientific innovation instead of compressing the two dimensions into one variable as the CD index does. They do this by exclusively retaining (the positive) term $N_f$ or $N_{bf}$ in the numerator of Equation 4, thereby respectively obtaining indicators for disruption and consolidation:

$$CD^* = \frac{N_f}{N_f + N_{bf} + N_b}, \tag{6}$$

$$CD^{\#} = \frac{N_{bf}}{N_f + N_{bf} + N_b} \tag{7}$$

Consequently, Chen et al. (2021) are able to identify dual characteristics of technologies by unfolding the two dimensions.

In a similar spirit, S. Wu and Wu (2019) call attention to a confusing characteristic of the CD index. They find that as the sign of the original index changes, the effect of $N_b$ (i.e., papers that cite the FP's references but not the FP itself) on the index value is contradictory. In the case that the numerator is positive, i.e., $N_f - N_{bf} > 0$, $N_b$ reduces disruption. For a negative value of the numerator, i.e., $N_f - N_{bf} < 0$, on the other hand, $N_b$ enhances disruption. This inconsistency can be eliminated by employing the independent indicators in the two-dimensional model of Chen et al. (2021).

### 5.1.3 Exclusion of the $N_b$ term: $CD^{nob}$

Taking the discussion around the $N_b$ term a step further, Bornmann et al. (2020) assess its validity for inclusion. Q. Wu and Yan (2019) find that the value for $N_b$ is often significantly larger than the other terms in Equation 4. Consequently, the CD index often generates disruption values of modest size and a high citation impact, i.e., a large value for $N_f$, is needed to achieve a high disruption score. As a result, Bornmann et al. (2020) question if $N_b$ is too dominant for adequately capturing nuances in disruption. This finding contradicts Funk and Owen-Smith (2017), who claim that disruptive papers, according to the CD index, do not need to have a large following;

It is worth mentioning that Funk and Owen-Smith (2017) subsequently introduce a variant of the CD index, the *mCD index*, that is able to capture the magnitude of disruption:

$$mCD_t = \frac{m_t}{n_t} \sum_{i=1}^{n} \frac{-2f_{it}b_{it} + f_{it}}{w_{it}}, w_{it} > 0, \tag{8}$$

where $m_t$ is the parameter capturing the magnitude and equals the number of (forward) citations of the FP. However, the citation impact is (now even more) a deciding factor in this mCD index. Thus, it is not relevant to the discussion on disruption indicators in this study.

Furthermore, Bornmann et al. (2020) express that the inclusion of $N_b$ does not capture the fundamental idea of disruption indicators that differentiates between subsequent papers that exclusively cite the FP or cite both the FP and its references. Instead, $N_b$ incorporates the relative citation impact of an FP in comparison to comparable papers, that is, papers that cite the same references the FP cites. In the case that one is specifically looking to find a high-impact disruptive paper, there might have a legitimate theoretical argument in favor of the inclusion of $N_b$ term.

Contrarily, when one aims to find a paper that is disruptive but not necessarily has a high impact (i.e., disruption as defined in this paper), they should not include the $N_b$ term (Bornmann et al., 2020). Accordingly, Bu et al. (2021) define a multi-dimensional framework on citation impact that draws a differentiation between "(1) the level, (2) the depth and breadth, and (3) the dependence and independence of the citation impact of a publication". They present an indicator of *independence*, $\frac{N_f}{N_f+N_{bf}}$, as part of a family of disruption (and consolidation) indicators. It is worth noting that it is highly comparable to the disruption indicator from Chen et al. (2021)

in Equation 6; it merely does not include the summation of the $N_b$ term in the denominator. However, this independence (i.e., disruption) indicator of Bu et al. (2021) does not have the same range of values as the original CD index. Therefore, Bornmann et al. (2020) introduce an analogous variant that does have the same range of values, denoted as

$$CD^{nob} = \frac{N_f - N_{bf}}{N_f + N_{bf}}, \tag{9}$$

where *nob* displays the exclusion of the $N_b$ term from the denominator.

#### 5.1.4 Co-citation threshold: $CD_l$

In addition to their independence indicator, Bu et al. (2021) introduce one indicator measuring *dependence* (i.e., consolidation). This dependence indicator takes into account how strong the relationship is between the FP's references and its citing papers, that is, the number of bibliographic couplings. It is defined as the average number of references of the FP that are cited by its citing papers. This indicator does not have an upper bound and is designed to decrease when there is more disruption since it captures a paper's dependency on previous work. For that reason, and such that it has the same range of values as the original indicator, Bornmann et al. (2020) propose the $CD_l$ variant, denoted as

$$CD_l = \frac{N_f - N_{bf}^l}{N_f + N_{bf}^l + N_b}, \tag{10}$$

where

$$N_{bf}^l = \quad | \{bf_i \mid bf_i \text{ cites FP} \wedge bf_i \text{ cites} \geq l \text{ of FP's references}\} | \tag{11}$$

Please be aware of the fact that the $CD_l$ variant in Equation 10 does include the $N_b$ term in the denominator, as proposed initially by Bornmann et al. (2020). Thus, when $l = 1$, $CD_l$ is equal to the original disruption indicator.

Nonetheless, the idea of requiring a threshold for a citing paper to be considered in the calculation of the indicator can similarly be applied to other variants. Correspondingly, the $CD^{nob}$ variant in Equation 9 can be denoted as

$$CD_l^{nob} = \frac{N_f - N_{bf}^l}{N_f + N_{bf}^l} \tag{12}$$

## 5.2 Correcting for changes in publication and citation practices over time

### 5.2.1 Normalization approaches

Park et al. (2023) propose two normalized variants of the CD index based on the idea of citing-side normalization of citation impact. The $N_b$ term (of the original indicator, see Equation 4) proves to be most representative of citation impact and thus is likely to scale with changes in publication and citation practices over time, relative to the other terms (Bornmann et al., 2020). Since this can generate a descending bias, Park et al. (2023) aim attention at the $N_b$ term for both their indicators.

Firstly, they introduce a 'Paper normalized' variant in which the number of (backward) citations made by the FP is subtracted from $N_b$ in order to offset the increased likelihood of $N_b$ being large due to a larger reference list. The 'Paper normalized' variant can be written as

$$CD_{norm}^{paper} = \frac{N_f - N_{bf}}{N_f + N_{bf} + (N_b - N_c)}, \tag{13}$$

where

$$(N_b - N_c) = \begin{cases} (N_b - N_c) & \text{if } N_c <= N_b, \\ 0 & \text{otherwise,} \end{cases} \tag{14}$$

such that the $(N_b - N_c)$ term is non-negative and where $N_c$ denotes the number of citations made by the FP.

Secondly, Park et al. (2023) present a 'Field x year normalized' indicator that field- and time-normalizes based on the average number of backward citations of a publication in the FP's field from the same publication year. This 'field x year normalized' variant can be written as

$$CD_{norm}^{field*year} = \frac{N_f - N_{bf}}{N_f + N_{bf} + (N_b - N_c^{mean})}, \tag{15}$$

where

$$(N_b - N_c^{mean}) = \begin{cases} (N_b - N_c^{mean}) & \text{if } N_c^{mean} <= N_b, \\ 0 & \text{otherwise,} \end{cases} \tag{16}$$

such that the $(N_b - N_c^{mean})$ term is non-negative and where $N_c^{mean}$ denotes the average number of backward citations of a publication in the FP's field in the publication year.

Since the components of the CD index include no literal citation counts or reference counts, but derivative measures of these counts, it makes it more difficult for one to design normalized indicators that correct for variations in publication and citation practices between publications from different fields and different years. Because the FP's citation behavior inherently determines the value of the CD index, I believe it is the right choice of Park et al. (2023) to correct based on average reference list length (and not based on average citation counts).

However, because of the fact that $N_b$ indeed proves to be too dominant to adequately capture nuances (i.e., very similar scores), these normalized variants fail to achieve their purpose of correcting for changes in publication and citation practices over time (see future Results section). This is because $N_b$ is multiple orders of magnitude larger than the correction terms (and the other $N_f$ and $N_{bf}$ terms).

In sum, there does not yet exist a robust field-normalized indicator to my knowledge.

### 5.3 Do the indicators capture what they promise?

To summarize, previous sections 5.1 and 5.2 present many (normalized) indicators for measuring disruption. Table 2 below displays the seven different indicators.

Table 2: This table reports a summary of the different disruption indicators.

| Indicator | Remarks | Main source(s) |
|---|---|---|
| $CD$ / $DI$ | Original index | Funk and Owen-Smith (2017) |
| $CD^*$ | Distinct disruption indicator (separate from consolidation) | Chen et al. (2021) |
| $CD^{nob}$ | Exclusion of $N_b$ | Bu et al. (2021); Bornmann et al. (2020) |
| $CD_l$ | Co-citation threshold | Bu et al. (2021); Bornmann et al. (2020) |
| $CD_l^{nob}$ | Exclusion of $N_b$ and Co-citation threshold | Bu et al. (2021); Bornmann et al. (2020) |
| $CD_{norm}^{paper}$ | 'Paper normalized' | Park et al. (2023) |
| $CD_{norm}^{field*year}$ | 'Field x year normalized' | Park et al. (2023) |

The original CD index, i.e., $CD_1$, is heavily criticized for its inclusion of the $N_b$ term that is said not to capture the fundamental idea of disruption indicators and primarily reflect citation impact (i.e., relative to other publications in a comparable context) (Bornmann et al., 2020; Leydesdorff et al., 2021).

Firstly, the inclusion of the $N_b$ term is responsible for confusing characteristics, i.e., contradictory effects, of the CD index dependent on the sign of the nominator (S. Wu & Wu, 2019). Moreover, because the FP's citation behavior inherently determines the value for $N_b$, minor changes in the FP's reference list are able to have a big impact on the resulting disruption score.

Accordingly, Bornmann et al. (2020) perform a factor analysis that shows $CD_1$ scores low on the 'disruptiveness' dimension. However, it does load strongly on the second dimension, which resembles citation impact. Contrarily, the other variants $CD_5$, $CD_1^{nob}$ and $CD_5^{nob}$ (where the subscript indicates the co-citation threshold) score highly on the 'disruptiveness' dimension, presumably because of their co-citation threshold and/or exclusion of the 'citation impact' term $N_b$.

In other efforts to diminish the effect of the $N_b$ term, the normalized variants $CD_{norm}^{paper}$ and $CD_{norm}^{field*year}$ were introduced (Park et al., 2023). However, the $N_b$ term that is still present in these normalized indicators indeed proves to be too dominant to adequately capture nuances. For that reason, I do not think the $N_b$ term should be included in the principal indicator for this study.

Furthermore, I support the idea of employing a larger-than-one co-citation threshold because it allows for capturing the strength of the FP's citing papers' reliance on the FP's references (Bornmann et al., 2020). In the case that one of the FP's references is highly cited , the probability for a citing paper to also cite one of the FP's references is substantially larger compared to a lower cited reference (reflected by a larger value for the $N_{bf}$ term). Since a higher probability reflects a larger value for $N_{bf}$, a sign of consolidation, not every citing paper of type *bf* is uniformly indicative of a consolidating FP. Therefore, only taking into account those reliable citing papers (e.g., $l = 5$) leads to a more robust measure.

In conclusion, I will employ the $CD_{l=5}^{nob}$ as the primary indicator of disruption throughout this study. This indicator, constructed based on fundamental bibliographic theory and the resulting citation networks, should be a robust measure of disruption and accordingly complies with the encouragement of Stremersch et al. (2015) to "move beyond mere citation counts to assess a paper's scientific contribution".

# 6 Data

The data in this study are derived from the WoS and consist of two distinct lists. Firstly, a list of focal papers, which also contains the references for each focal paper. Secondly, a list of articles that cite the focal papers and/or (one or more of) the focal paper's references. The articles that comprise the latter list can be obtained by performing a so-called *Cited Reference Search* in the WoS for the focal publications and all their (unique) references.

The lists contain each article's unique identifier (i.e., DOI; Digital Object Identifier), which is needed to establish citation links between papers and construct the citation networks. Any articles in the lists retrieved from the WoS containing duplicate or missing DOIs are discarded from the samples. Furthermore, the lists include additional publication information, such as journal title, document type, and publication date.

With the help of the publication date, different citation windows can be studied. Since a citation window of at least three years is necessary to generate relevant insights (Bornmann & Tekles, 2019), the most recent publication year in the deployed samples is 2019. Consequently, the number of focal articles in the relevant samples will decrease when a citation window of more than three years is used. I study four different window lengths, three of which are 'regular': three, five, and ten years after publication. Lastly, I employ a citation window that includes all the citations up to and including the year 2022. Contrary to the other three windows, the '2022' window doesn't use the publication date to enforce which citations can be included. Nonetheless, the included focal papers in this window are still published in 2019 or earlier to comply with the minimal window length of three years. Therefore, it includes the same number of publications as the three-year window (but encompasses more citations). The idea behind such a citation window is that it allows for the delayed recognition (of older papers). I will refer (in many tables and figures) to this window as '2022'.

The final selection criterion I employ is regarding the focal paper's document type, for which I only consider *articles* and not types such as *editorial material* or *review article*.

Finally, it is important to highlight that the above-mentioned lists encompass vast amounts of data that are rather cumbersome to obtain from the WoS. The WoS does have an API available, however, because of the extremely large number of publications one needs information on, one would need a great number of credits for API calls. To combat this issue, I employ RPA (Robotic Process Automation), which is able to automatically download the data in batches of 1,000 papers (out of millions). Nonetheless, this technique still requires days of running since it replicates user actions (e.g., clicking and typing).

The following sections provide more insight into the exact publication numbers, citation counts, and reference figures for the Marketing data and the AI data on a journal level. Moreover, I offer a Marketing vs. AI comparison (on the field level) regarding citation and reference counts.

## 6.1 Marketing data

Table 3 reports the number of focal papers per journal for the four Marketing samples (i.e., the four different citation windows). Since the 3-year and '2022' samples contain the same number of publications, they are displayed in the same column. It is evident that MKS accounts for the fewest publications, followed by JM. On the other hand, JCR and JMR have the highest number of papers in all the samples.

The 7,093 focal papers (the full sample) contain 60,429 references with a unique DOI. The number of unique publications that cite a focal paper and/or (one or more of) the focal paper's references is 3,688,818. In total, these citing papers make 141,357,933 unique citations to the focal paper and/or (one or more of) the focal paper's references.

Table 3: This table reports the number of focal publications per journal included in the four Marketing samples.

| **Journal** | **Number of publications** | | |
|---|---|---|---|
| | Citation window | | |
| | 3-year/2022 | 5-year | 10-year |
| JCR | 2113 | 1978 | 1631 |
| JM | 1667 | 1578 | 1365 |
| JMR | 2052 | 1934 | 1649 |
| MKS | 1261 | 1162 | 907 |
| **Total** | 7093 | 6652 | 5552 |

Figure 4 below plots the publication count across time for the full sample. For the exact corresponding numbers, see Table 12 in the Appendix. Across all journals, the number of publications was relatively steady until the turn of the century but started to increase in the early 2000s and continued to do so throughout that decade. However, all journals appear to have reduced their publication count again after 2010.

Focusing on individual journals, JCR consistently publishes the largest number of papers starting around 1985, with the exception of the 1995-2003 and 2010-2011 periods. During the first decade included in the data, JM and JMR produced the most articles. JMR continues this trend to some extent, while JM tumbles to the lowest-publishing journal. JM shared this qualification with MKS from the inception of MKS in 1987 till around 2005, when MKS started to consistently publish more than JM.

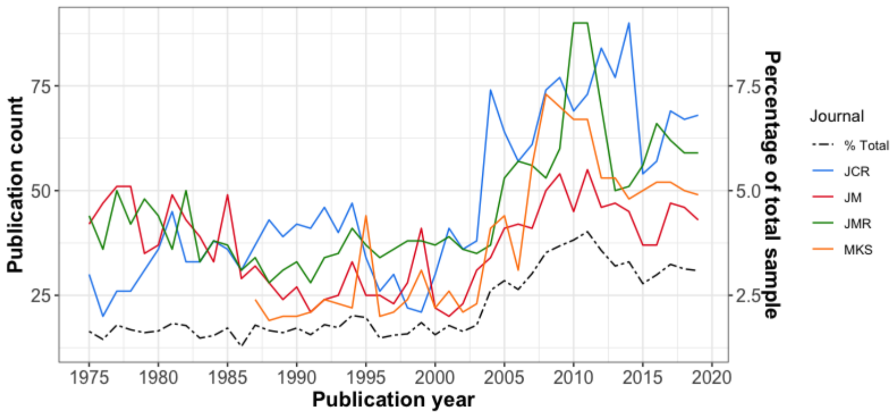
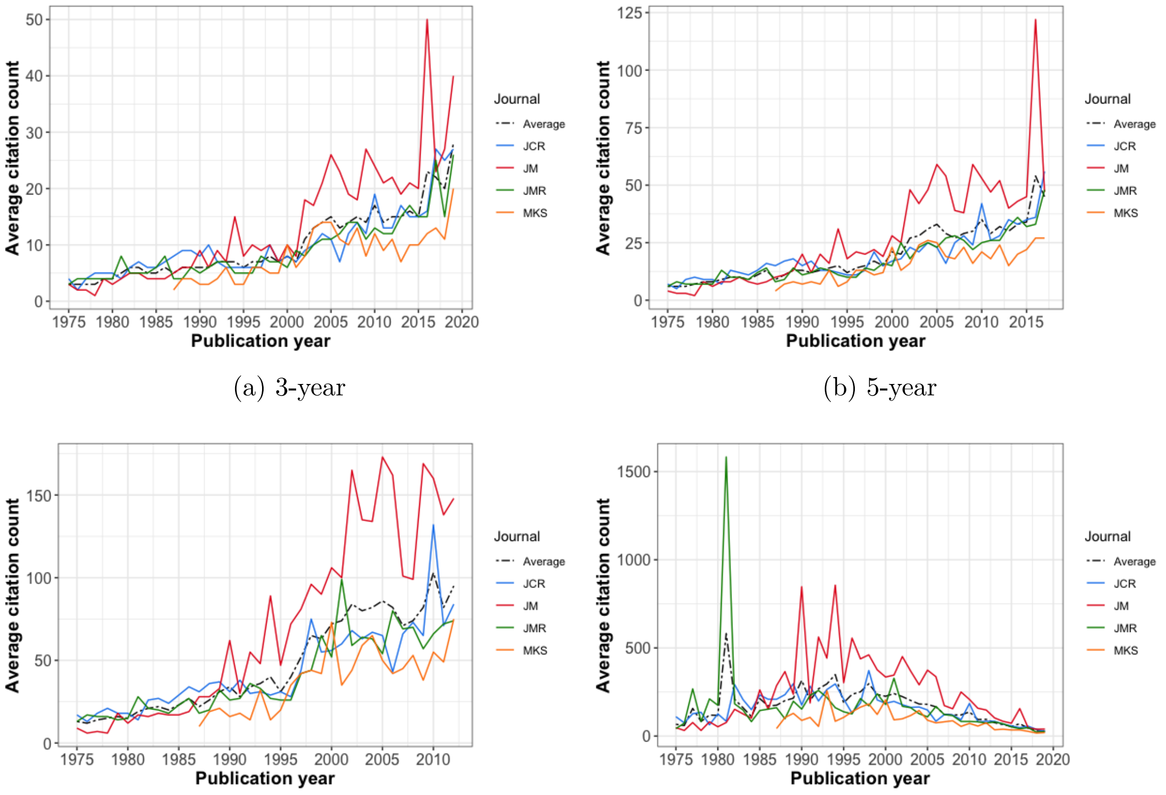



Figure 4: This figure displays the publication count per journal for the full Marketing sample.

Figure 5 displays the average number of citations per journal for the different citation windows in the respective sub-figures. Starting with the 3-year window in Figure 5a, there is clearly a trend encompassing a higher citation count for more recent publications. Moreover, JM publications show to have considerably more citations than the average, while MKS articles often fall short of the average. The steep peak in 2016 for JM can be (partly) attributed to the paper *Understanding Customer Experience Throughout the Customer Journey*, which received 399 citations during the first three years after publication. JCR and JM follow the average trend closely. Figure 5b shows that the 5-year citation window is, unsurprisingly, not much different compared to the 3-year window.

However, shifting attention to the 10-year window in Figure 5c, one can see larger differences between journals without any ranking changes. A similar overall trend of increasing citations for younger publications is visible. Nevertheless, this trend is inconsistent with the trend in the smaller citation windows because there is a steady increase from 1995-2002 and stagnation thereafter.




(a) 3-year

(b) 5-year

(c) 10-year

(d) 2022

Figure 5: This figure displays the average citation count per journal for the four Marketing samples.

Finally, the '2022' citation window that takes into account all citations up to and including 2022 shows a very distinct trend in Figure 5d. Publications between 1990-1995 have amassed the highest number of average citations. Papers before this window have received few citations, except for JMR articles in 1981. The reason for that is the article *Evaluating Structural Equation Models With Unobservable Variables And Measurement Error*, which has 51,321 citations as of 2022. The average citation count after the 1990-1995 period slowly declines. Like the other windows, JM strongly ranks first in average citation count, while JCR, JMR, and MKS display little variation among them.

Summary statistics on the number of citations of a single focal publication are displayed in Table 4, confirming the above-mentioned patterns among citation windows and journals. A notable highlight is the mean being persistently larger than the median due to those publications that receive an outlier-like number of citations.

Table 4: This table reports summary statistics on the number of (forward) citations of a single publication per journal in the four Marketing samples. The minimum number of citations per publication is 0 in all cases, thus not reported.

| **Journal** | **Citation summary statistic** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Citation window | | | | | | | | | | | |
| | 3-year | | | 5-year | | | 10-year | | | 2022 | | |
| | Median | Mean | Max. | Median | Mean | Max. | Median | Mean | Max. | Median | Mean | Max. |
| JCR | 8 | 12 | 368 | 15 | 22 | 950 | 32 | 50 | 3760 | 72 | 141 | 5772 |
| JM | 8 | 14 | 399 | 15 | 29 | 1126 | 31 | 72 | 1040 | 78 | 219 | 9521 |
| JMR | 6 | 10 | 616 | 12 | 19 | 1005 | 25 | 44 | 1299 | 52 | 153 | 51321 |
| MKS | 6 | 9 | 386 | 12 | 18 | 617 | 27 | 43 | 1213 | 43 | 79 | 1959 |

To conclude the exploratory data analysis for the Marketing data, I study the number of references a single focal publication makes. Figure 6 shows the average reference count and exhibits an unambiguous, increasing long-term trend. While the average number of references is quite low at the start of the data, it has more than doubled in 2019. I have been unable to find (journal-specific) hard evidence for reasons as to why the reference count has increased this greatly. A simple explanation could be the exponential growth of the scientific literature (D. Wang & Barabási, 2021), which in turn has allowed researchers to build upon more existing knowledge.

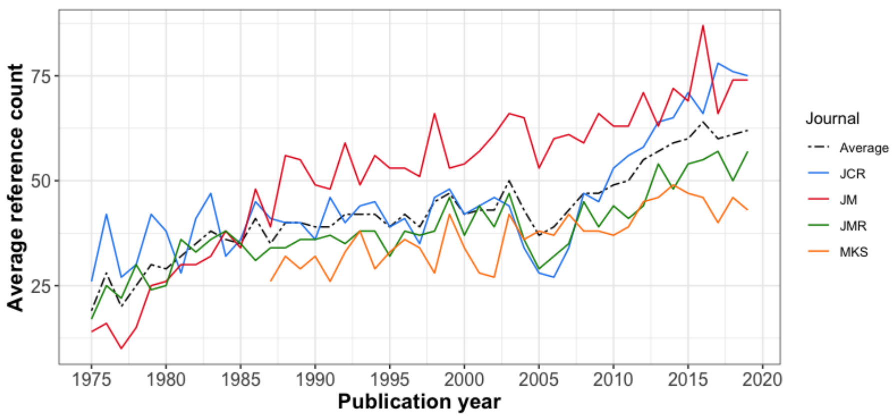



Figure 6: This figure displays the average reference count per journal for the full Marketing sample.

JM, in particular, has almost quadrupled its average reference count. This vast increase could be related to JM's shift to focus more on conceptual articles around the 1970s (Yadav, 2010), which involves a greater reliance on previous scientific works. JCR's average remained relatively flat from 1975 to 2005 but greatly increased in recent years. Correspondingly, the summary statistics of these two journals, see Table 5, are greater compared to JMR and MKS.

Specifically, the maximum number of references is considerably larger for JCR and JM. JMR closely follows the average, while MKS is consistently below average, particularly in the final years of the sample. Additionally, both journals report very similar summary statistics. Finally, and interestingly, it is important to highlight that there exist publications in the sample that do not make any references.

Table 5: This table reports summary statistics on the number of references of a single publication per journal in the full Marketing sample.

| **Journal** | **Reference summary statistic** | | | |
|---|---|---|---|---|
| | Min. | Median | Mean | Max. |
| JCR | 0 | 42 | 48 | 327 |
| JM | 0 | 51 | 51 | 252 |
| JMR | 0 | 37 | 39 | 146 |
| MKS | 0 | 36 | 39 | 139 |

### 6.2 AI data

Table 6 displays the number of focal papers per journal for the four AI samples. TPAML (Transactions On Pattern Analysis And Machine Intelligence) articles constitute roughly half of each sample. Contrarily, AIR (Artificial Intelligence Review) accounts for the lowest number of publications by a great margin. AI (Artificial Intelligence; the journal) has the second-highest publication count, shortly followed by IJCV (International Journal Of Computer Vision). Figure 15 in the Appendix shows the distribution across time, and Table 13 in the Appendix reports the exact corresponding numbers.

Table 6: This table reports the number of focal publications per journal included in the four Marketing samples.

| **Journal** | **Number of publications** | | |
|---|---|---|---|
| | Citation window | | |
| | 3-year/2022 | 5-year | 10-year |
| AI | 2365 | 2228 | 1860 |
| AIR | 837 | 724 | 512 |
| IJCV | 1978 | 1817 | 1375 |
| TPAML | 4968 | 4525 | 3547 |
| **Total** | 10148 | 9294 | 7294 |

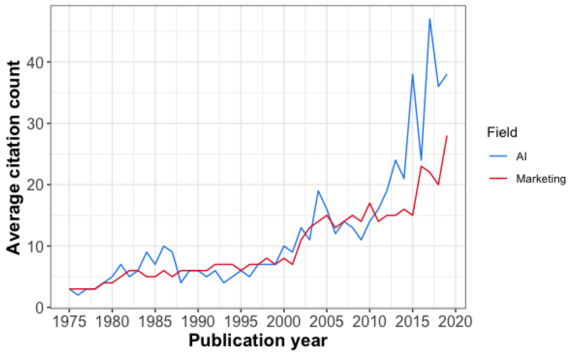
The 10,148 AI focal papers (the full sample) contain 86,222 references with a unique DOI. The number of unique publications that cite a focal paper and/or (one or more of) the focal paper's references is 4,346,409. In total, these citing papers make 157,798,462 unique citations to the focal paper and/or (one or more of) the focal paper's references.

Since the AI sample principally serves as a comparison for the Marketing field, I present the journal-level data regarding citations and references in the Appendix; Figure 16 displays the average citation count, Table 14 reports citation summary statistics, Figure 17 shows the average reference and Table 15 presents the reference summary statistics.

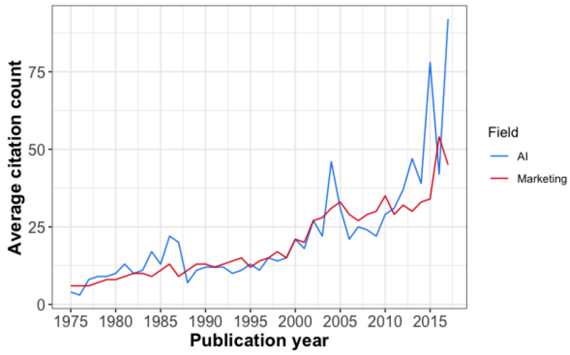
### 6.3 Marketing vs. AI comparison

As the final part of the exploratory data analysis, I compare the citation and reference counts of the Marketing and AI fields.

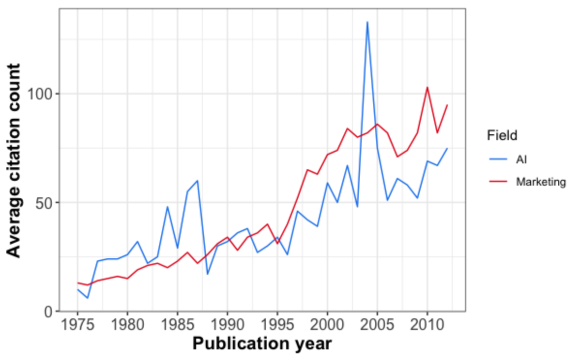
Figure 7 shows the four citation windows' average counts. In the 3-year window in Figure 7a, the fields basically followed the same trend until 2010, when the count for AI started to increase at a faster rate. The trend in the 5-year citation window in Figure 7b is very much alike. Considering the above, it appears that recent AI papers secure citations noticeably faster than recent Marketing papers. This is confirmed by the corresponding summary statistics in Table 7.



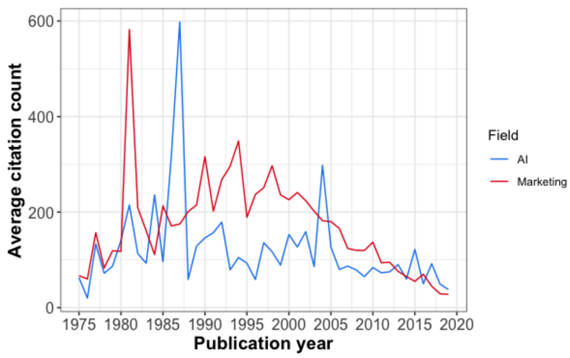

(a) 3-year




(b) 5-year




(c) 10-year




(d) 2022

Figure 7: This figure displays the average citation count per field for the four citation window lengths.

One can see substantial changes in the 10-year window in Figure 7c. During the first decade included in the data, AI outperforms the field of Marketing. However, after this period, the average citation count is roughly equal for about ten years, and after that, Marketing consistently surpasses AI with the exception of 2004. The paper *Distinctive image features from scale-invariant key points* is largely responsible for that exception, which has amassed 12,455 in the 10-year window.

Figure 7d takes into account all citations up to and including 2022. Marketing papers have frequently received more citations than AI papers; see also Table 7. Exceptions are the above-mentioned paper from 2004 and the article *A Computational Approach To Edge-detection* from 1986, which has received 15,597 citations as of 2022. On average, the most cited Marketing papers were published between 1990 and 2000, while for AI, it is more difficult to define such a period.

Table 7: This table reports summary statistics on the number of (forward) citations of a single publication per field for the four citation window-length samples. The minimum number of citations per publication is 0 in all cases, thus not reported.

| **Field** | **Citation summary statistic** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Citation window | | | | | | | | | | | |
| | 3-year | | | 5-year | | | 10-year | | | 2022 | | |
| | Median | Mean | Max. | Median | Mean | Max. | Median | Mean | Max. | Median | Mean | Max. |
| AI | 8 | 22 | 12294 | 14 | 38 | 25752 | 24 | 68 | 12455 | 33 | 128 | 32117 |
| Marketing | 7 | 11 | 616 | 14 | 22 | 1126 | 29 | 52 | 3760 | 61 | 152 | 51321 |

To conclude this section, I compare the reference counts. Figure 8 shows the trend across time, and Table 8 the accompanying summary statistics. The two fields exhibit a similar, increasing trend in the average number of references starting around 1992. Before this, AI papers made substantially fewer references.

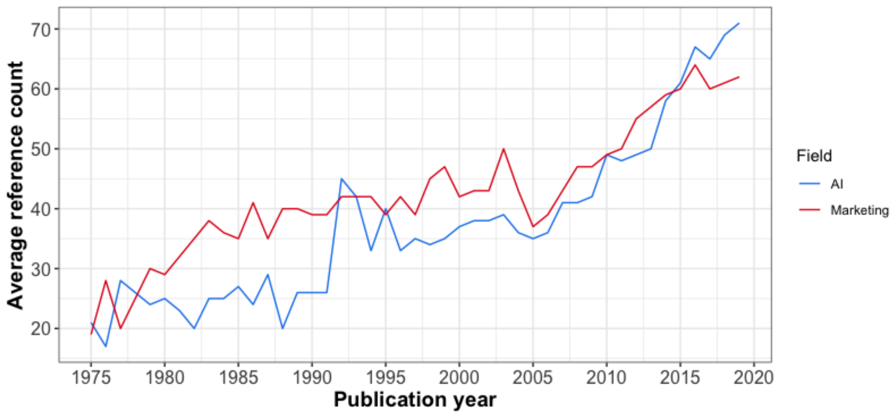



Figure 8: This figure displays the average reference count per field for the full samples.

Table 8: This table reports summary statistics on the number of references of a single publication per field in the full AI sample.

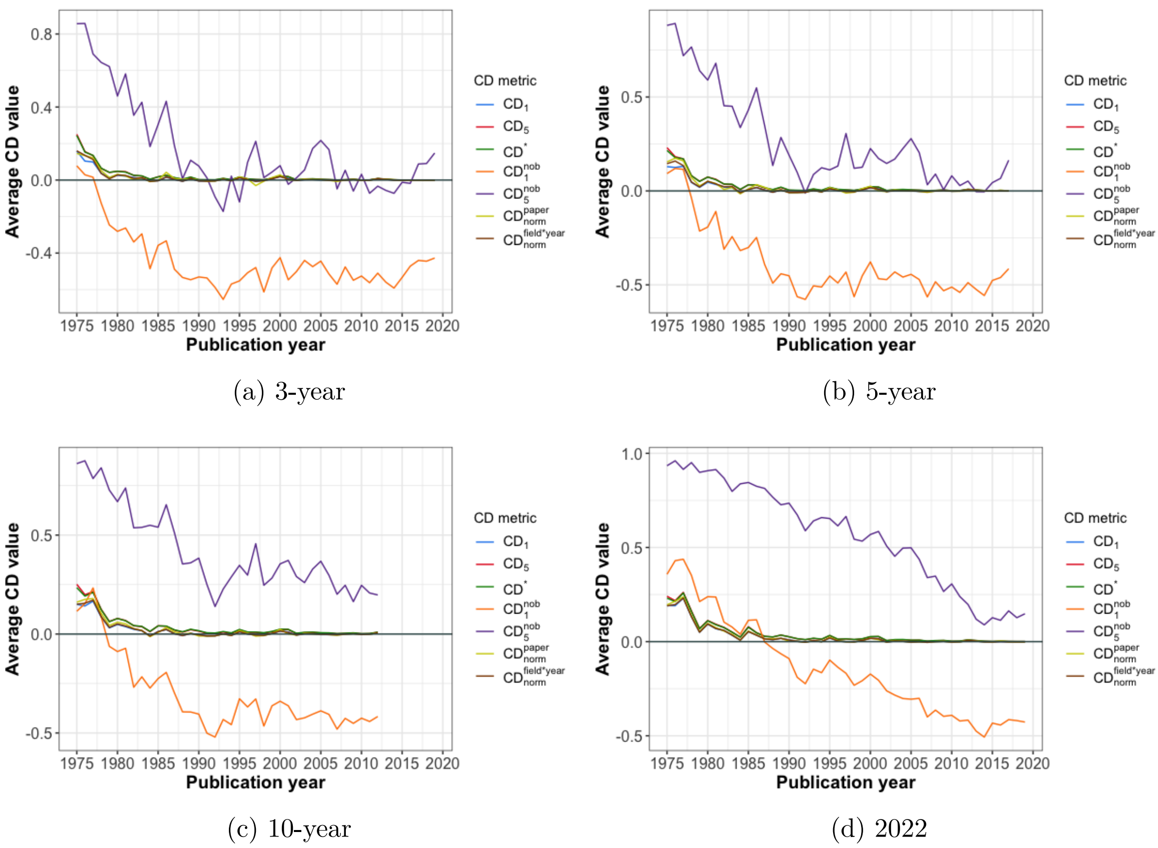
| **Field** | **Reference summary statistic** | | | |
|---|---|---|---|---|
| | Min. | Median | Mean | Max. |
| AI | 0 | 36 | 40 | 724 |
| Marketing | 0 | 41 | 44 | 327 |

# 7 Results

## 7.1 Marketing

### 7.1.1 Initial results

Figure 9 shows the average CD values for the seven different indicators of disruption (see Table 2). It is apparent that the $CD^{nob}$ indicators have more distinct values over a greater range compared to the other five indicators that do include the $N_b$ term (in the denominator). The cause for this is the strong influence of the $N_b$ term that is multiple orders of magnitude larger than the other $N_f$ and $N_{bf}$ terms. Figure 18 in the Appendix displays these five indicators separately from the $CD^{nob}$ indicators. Although on a smaller scale, they still follow almost identical trends; downward and converging to zero.




(a) 3-year

(b) 5-year

(c) 10-year

(d) 2022

Figure 9: This figure displays the seven different disruption indicators' values, averaged by publication year for the four Marketing samples.

Irrespective of the choice of indicator, it appears that early papers have a higher disruption score compared to more recent ones. Upon inspection of these initial results, it turns out that many of the early papers (1975-1990) have, on average, less than half of their cited references available for the construction of their citation networks due to limitations of the WoS, see Figure

10. This is especially relevant for the field of Marketing since the difference between available and actual publications included in bibliometric databases is far greater for social sciences compared to natural and life sciences (Bornmann et al., 2020; Moed, 2006). Moed (2006) finds that this is partially due to the larger prevalence of non-article publications like books as cited references, which are not included in the WoS. Moreover, there are hundreds of publications that have 0 references available in the WoS, and thus will receive a maximum disruption score in case they have any forward citations.

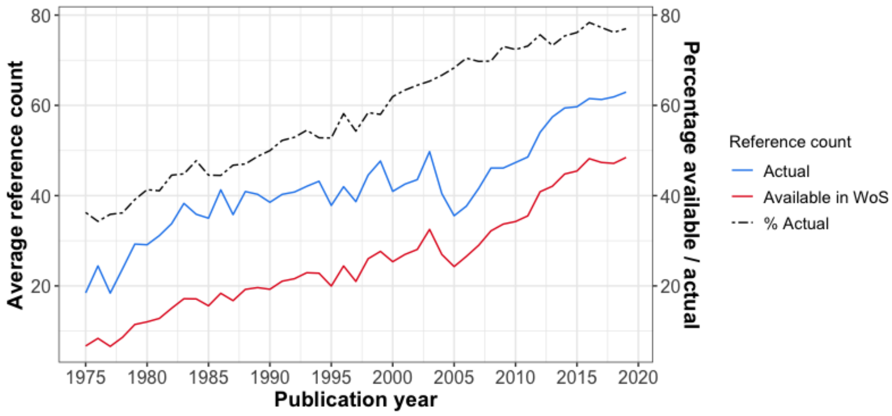
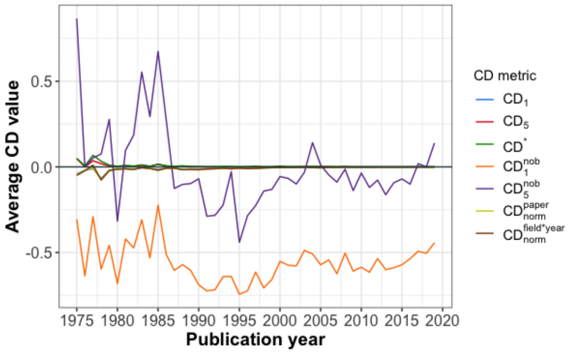



Figure 10: This figure displays the average reference counts, i.e., actual and available, for the full Marketing sample.

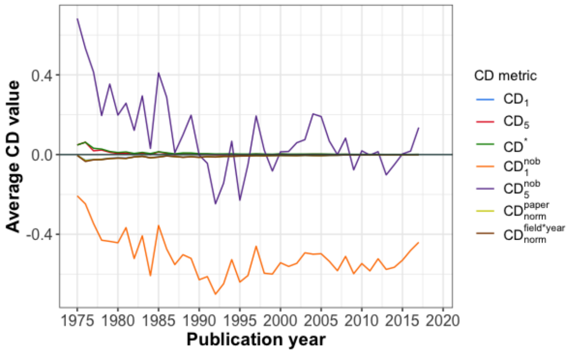
To make my samples more robust, I choose to continue this analysis with a subset of the data such that robust citation networks can be constructed for each individual paper. Following Bornmann et al. (2020), I include only the publications with at least ten references and citations to strengthen the validity of the different indicators. This is also in line with other studies (Deng and Zeng, 2023; Ruan et al., 2021) that recommended imposing a minimum on the number (forward) of citations and cited references.

Nonetheless, this does not fully eliminate the tendency of early publications to have disproportionally fewer references available compared to recent papers (due to the low coverage in WoS). This allows for a possible bias resulting in the artificial inflation of disruption scores of older papers. Because the data is more complete for younger publications, using data from, e.g., 2000 and later will make the results even more robust. However, this would not allow for studying long-term trends, which is one of the main goals of this paper. Note that the robustness samples with a lower citation window contain fewer articles compared to the original samples due to the imposed condition of a minimum of ten citations (within the window). Tables 16 (Marketing) and 17 (AI) in the Appendix report the specifics of these more robust data.

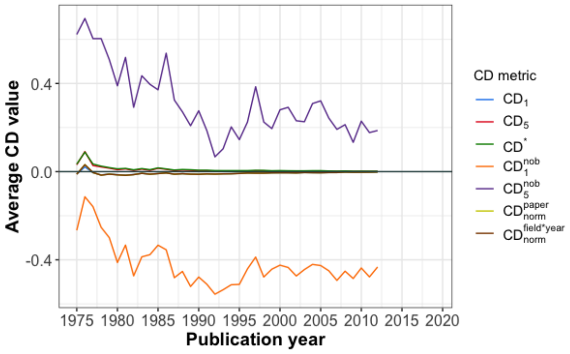
Figure 11 below shows the recalculated results for the four robust Marketing samples.



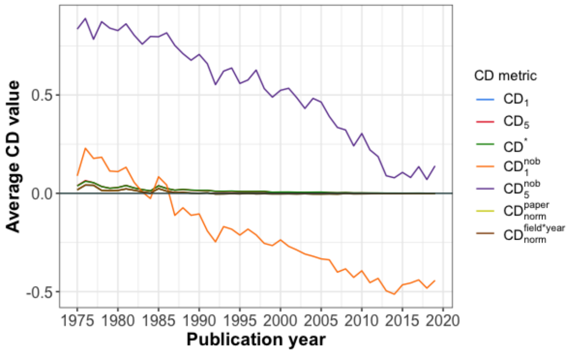

(a) 3-year ($n = 2243$)




(b) 5-year ($n = 3633$)




(c) 10-year ($n = 3944$)



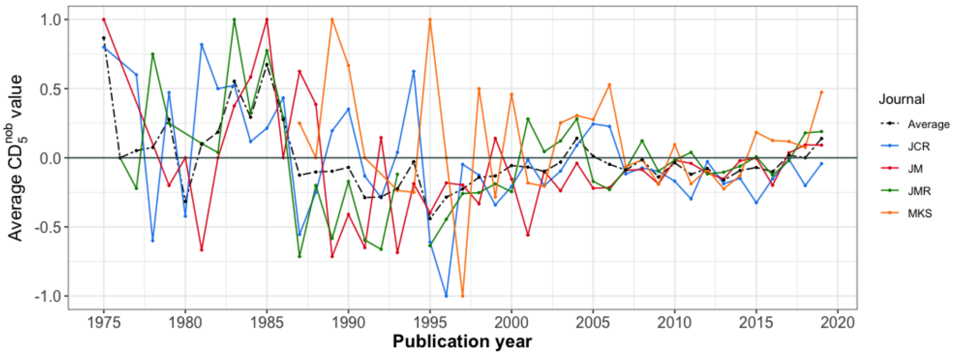

(d) 2022 ($n = 5699$)

Figure 11: This figure displays the seven different disruption indicators' values, averaged by publication year for the four robust Marketing samples.

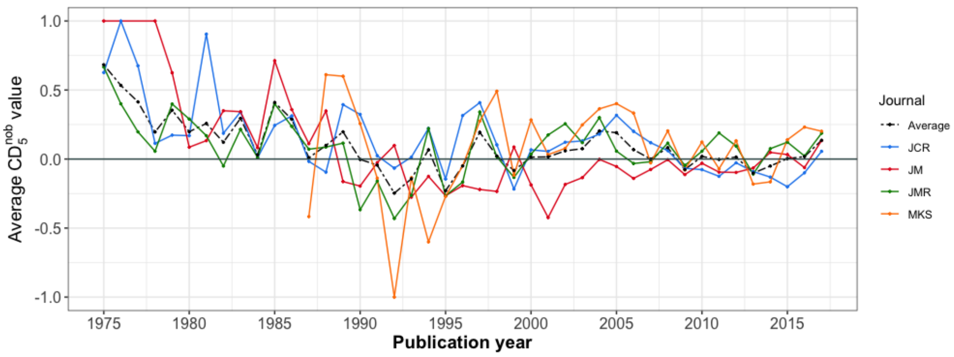
Still, the $CD^{nob}$ indicators display more nuanced values compared to the other five indicators that do include the $N_b$ term. Additionally, Figure 19 in the Appendix displays the five non-$CD^{nob}$ indicators for the robust Marketing samples. Adverse to the complete sample, these indicators now have diverging trends. $CD_{l=5}$ and $CD^*$ display downward trends while $CD_{l=1}$, i.e., the original index, $CD^{paper}_{norm}$ and $CD^{field*year}_{norm}$ display upward trends. What remains the same is the fact that all trends converge toward zero. Nonetheless, the $CD^{nob}_{l=5}$ measure will be the indicator of interest throughout the remainder of this analysis since it allows for more nuanced disruption scores, especially for the years 2000 and later.

### 7.1.2 Sub-field: journal vs. journal comparison

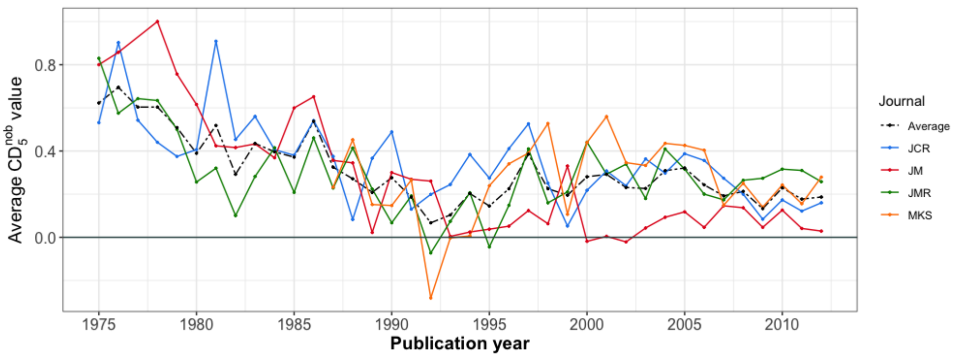
Figure 12 on the next page displays the disruption scores, based on $CD^{nob}_{l=5}$, for the JCR, JM, JMR, and MKS journals individually (using the robust data). Since the trends for each journal follow each other quite closely, the plots are increased in size to enhance interpretation. Table 10 reports the accompanying summary statistics.



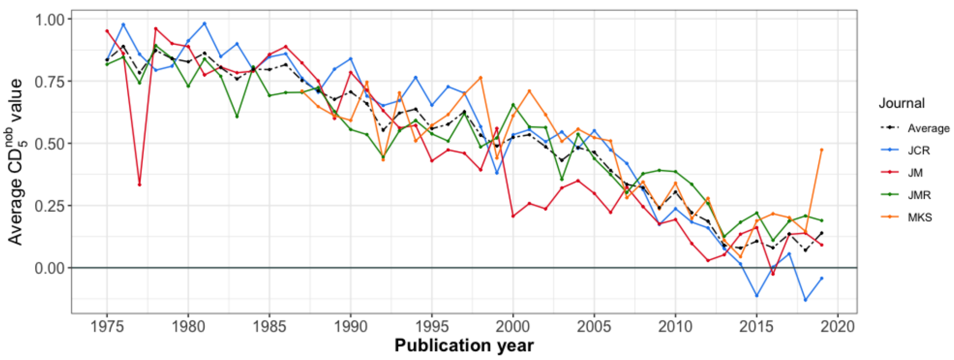

(a) 3-year ($n$ = 2243)




(b) 5-year ($n$ = 3633)




(c) 10-year ($n$ = 3944)




(d) 2022 ($n$ = 5699)

Figure 12: This figure displays the average $CD_5^{nob}$ value per journal, for the four robust Marketing samples.

Starting with Figure 12a, which corresponds to the sample with a citation window of three years, one will notice rather volatile values. This is partly because of the low number of publications that are included in the samples for some of the journals, particularly in earlier publication years. Nonetheless, there are some insights to be drawn from this. JCR and JM have the lowest average disruption score, followed by JMR. These three journals alternated for the highest average value until around 1987, the year of inception of MKS. Starting then, it is often MKS that produces the most disruptive studies, occasionally changing back and forth with primarily JMR.

Figure 12b shows similar trends, with MKS leading the journals in disruptiveness in the majority of recent years. JMR appears to score below average in the first half of the study period but scores relatively consistently above after the turn of the century. Conversely, JM appears to persistently be the least disruptive journal starting around 1995, although it scored better in the final years.

The trends in Figure 12c do not deviate much from those in Figure 12b, while the volatility does decrease. JCR, JMR, and MKS have almost identical averages over the whole period. Compared to the 3-year and 5-year samples, JCR more clearly is the most disruptive journal, for the most part in the period 1988-1997.

Finally, the '2022' citation window that takes into account all citations up to and including 2022, shows, in Figure 12d, a distinct downward trend over the complete study period, contrary to the other three samples. Yet, the patterns among journals remain comparable, while their average disruption scores are more alike, see Table 10 below.

Table 9: This table reports summary statistics on the $CD_5^{nob}$ indicator per Marketing journal for the four robust citation window-length samples. The minimum and maximum values are respectively -1 and 1 in all cases, thus not reported.

| **Journal** | $CD_5^{nob}$ **summary statistic** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Citation window | | | | | | | | | | | |
| | 3-year | | | 5-year | | | 10-year | | | 2022 | | |
| | Median | Mean | Std. | Median | Mean | Std. | Median | Mean | Std. | Median | Mean | Std. |
| JCR | -0.2 | -0.09 | 0.61 | 0 | 0.04 | 0.61 | 0.33 | 0.28 | 0.55 | 0.56 | 0.42 | 0.53 |
| JM | -0.2 | -0.09 | 0.61 | -0.12 | -0.04 | 0.6 | 0.2 | 0.15 | 0.58 | 0.5 | 0.36 | 0.55 |
| JMR | 0 | -0.03 | 0.64 | 0 | 0.07 | 0.61 | 0.33 | 0.27 | 0.58 | 0.6 | 0.45 | 0.51 |
| MKS | 0.1 | 0.12 | 0.62 | 0.06 | 0.1 | 0.62 | 0.3 | 0.26 | 0.56 | 0.5 | 0.38 | 0.54 |

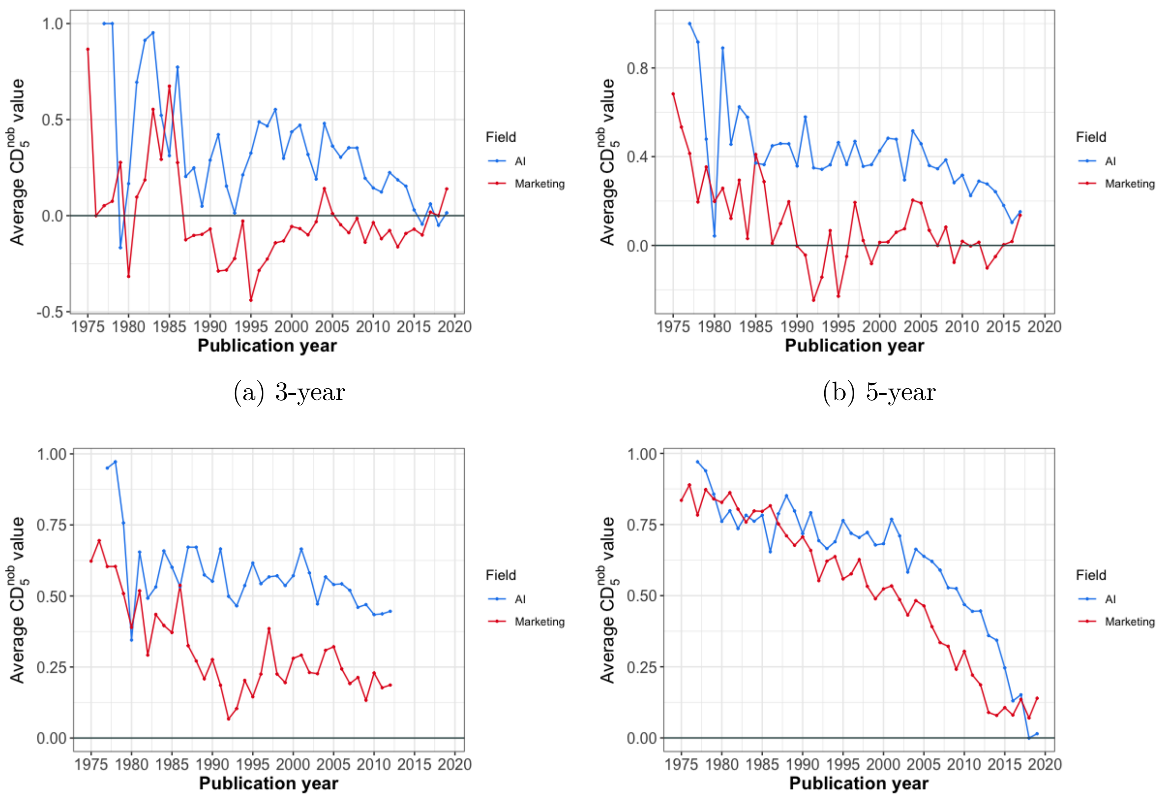
## 7.2 AI

Figure 20 and Table 18 in the Appendix report the subfield, i.e., journal vs. journal, results for the field of AI.

## 7.3 Marketing vs. AI comparison

Since there does not yet exist a robust field-normalized indicator, to my knowledge (see section Normalization approaches), the $CD_{l=5}^{nob}$ index represents the best-known measure because it is the least susceptible to citation and reference trends compared to other CD variants. More importantly, because the fields of Marketing and AI have comparable citation and reference trends (see section 6.3), the $CD_{l=5}^{nob}$ values should be worthy of comparison among these two field (but still less so over time).

Figure 13 below shows the average $CD_5^{nob}$ value for both fields, for the four different citation windows, and Table 10 the accompanying summary statistics.




(a) 3-year

(b) 5-year

(c) 10-year

(d) 2022

Figure 13: This figure displays the average $CD_5^{nob}$ value per field, for the four robust samples.

Across all four different samples, it is clear that the field of AI produces, on average, more disruptive papers compared to the field of Marketing. The distance in disruptiveness scores between the two fields is the largest in the 1990-2002 period in Figures 13a, 13b, and 13c.

However, while AI papers still appear to have declined in disruptiveness in the most recent years, Marketing papers appear to have reversed their downward trend, particularly in the three-year and five-year samples (Figures 13a and 13b).

The upcoming section 7.3.1 will test the robustness of these trends.

Table 10: This table reports summary statistics on the $CD_5^{nob}$ indicator per field for the four robust citation window-length samples. The minimum and maximum values are respectively -1 and 1 in all cases, thus not reported.

| **Field** | $CD_5^{nob}$ **summary statistic** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Citation window | | | | | | | | | | | |
| | 3-year | | | 5-year | | | 10-year | | | 2022 | | |
| | Median | Mean | Std. | Median | Mean | Std. | Median | Mean | Std. | Median | Mean | Std. |
| AI | 0.67 | 0.35 | 0.74 | 0.8 | 0.46 | 0.67 | 0.99 | 0.63 | 0.55 | 0.89 | 0.58 | 0.58 |
| Marketing | -0.1 | -0.05 | 0.62 | 0 | 0.04 | 0.61 | 0.33 | 0.25 | 0.57 | 0.56 | 0.41 | 0.53 |

#### 7.3.1 Robustness check using regression

Section 4.6 highlights the importance of correcting for changes in publication, citation, and authorship practices over time. Because of such changes, the comparison of disruption indicator values over time can be biased.

To correct for the aforementioned changes, I employ different regression models that include relevant, paper-level control variables. For each of the four citation windows, I estimate a baseline model that simply includes an indicator variable for each year (of publication). Additionally, for each of the four citation windows, I estimate a model that controls for the number of references, the number of authors, and the article length (in pages). Furthermore, I also include the number of missing or unlinked references (i.e., the difference between the actual number and the number available in the WoS) as a control.

Table 11 below reports results for these regression models, where $CD_5^{nob}$ is the dependent variable. The baseline category is 1975 in all models. Results show that the incorporation of the control variables substantially improves the fit of the models. This is confirmed by the Wald tests, which test against the null hypothesis of all coefficients being equal to zero (by means of a two-sided t-test). Nonetheless, only the control variables involving the number of references seem significant. The coefficient for the number of authors is significant in none of the models, while the coefficient for the article length is only significant for the '2022' model.

Table 11: This table reports the results of the OLS regression models, with $CD_5^{nob}$ as the dependent variable, for the four robust Marketing samples. The paper-level control variables include the number of references, the number of authors, the article length (in pages), and the number of missing or unlinked references (i.e., the difference between the actual number and the number available in the WoS). The baseline category is the year 1975 in all models.

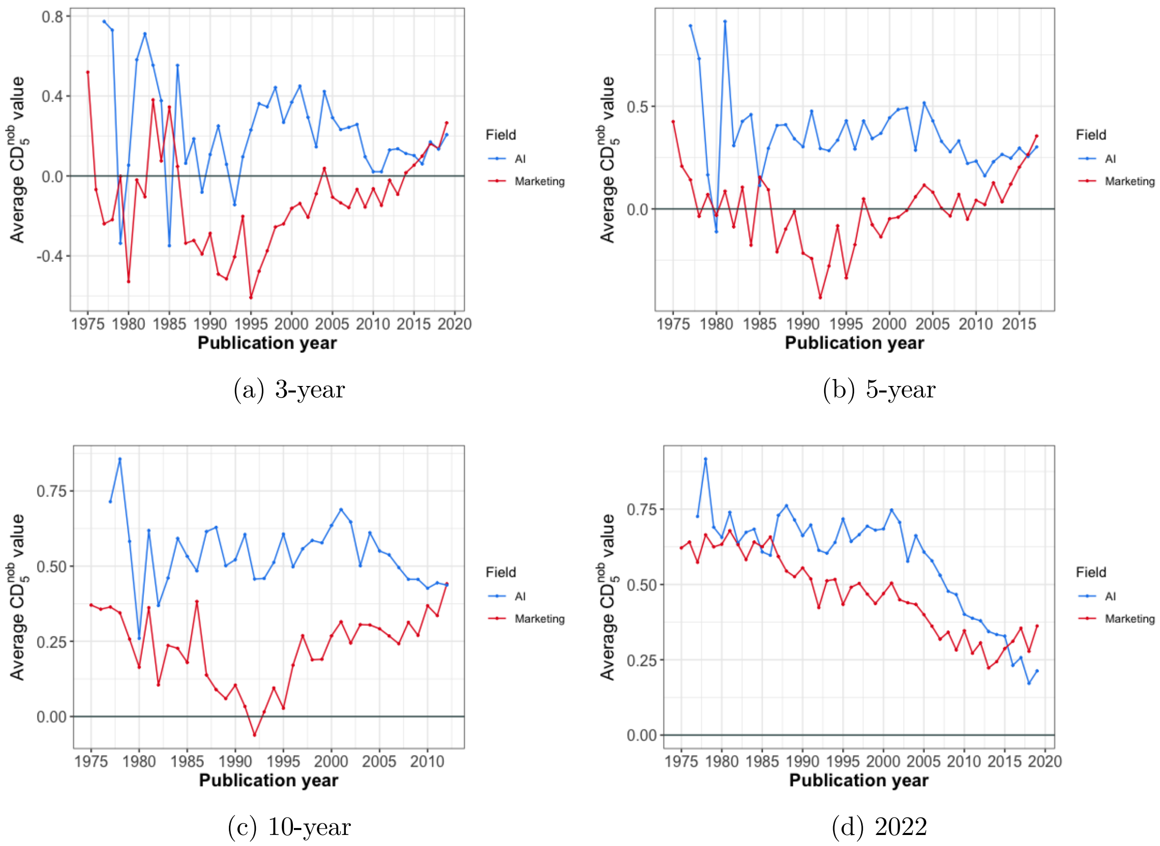
| **Citation window** | 3-year | | 5-year | | 10-year | | 2022 | |
|---|---|---|---|---|---|---|---|---|
| **Model** | Baseline | Controlled | Baseline | Controlled | Baseline | Controlled | Baseline | Controlled |
| (Intercept) | **0.87** (0.36)$^{*}$ | **0.92** (0.34)$^{**}$ | **0.68** (0.21)$^{**}$ | **0.86** (0.20)$^{***}$ | **0.62** (0.14)$^{***}$ | **0.82** (0.13)$^{***}$ | **0.84** (0.10)$^{***}$ | **0.93** (0.09)$^{***}$ |
| year=1976 | −0.87 (0.71) | −0.59 (0.67) | −0.15 (0.29) | −0.22 (0.27) | 0.07 (0.20) | −0.01 (0.18) | 0.05 (0.13) | 0.02 (0.12) |
| year=1977 | −0.81 (0.50) | −0.76 (0.47) | −0.27 (0.28) | −0.28 (0.26) | −0.02 (0.18) | −0.01 (0.17) | −0.05 (0.13) | −0.05 (0.12) |
| year=1978 | −0.79 (0.56) | −0.74 (0.53) | −0.49 (0.30) | −0.46 (0.28) | −0.02 (0.18) | −0.03 (0.16) | 0.04 (0.13) | 0.04 (0.12) |
| year=1979 | −0.59 (0.42) | −0.52 (0.39) | −0.33 (0.25) | −0.36 (0.24) | −0.11 (0.17) | −0.11 (0.15) | 0.01 (0.12) | 0.00 (0.11) |
| year=1980 | **−1.18** (0.47)$^{*}$ | **−1.05** (0.44)$^{*}$ | −0.49 (0.25) | −0.46 (0.24) | −0.23 (0.17) | −0.21 (0.16) | −0.01 (0.12) | 0.01 (0.11) |
| year=1981 | −0.77 (0.41) | −0.54 (0.39) | −0.43 (0.26) | −0.34 (0.25) | −0.10 (0.17) | −0.01 (0.15) | 0.03 (0.12) | 0.06 (0.11) |
| year=1982 | −0.68 (0.40) | −0.62 (0.37) | **−0.56** (0.24)$^{*}$ | **−0.51** (0.22)$^{*}$ | **−0.33** (0.16)$^{*}$ | −0.27 (0.15) | −0.03 (0.11) | 0.01 (0.10) |
| year=1983 | −0.31 (0.44) | −0.14 (0.41) | −0.38 (0.24) | −0.32 (0.23) | −0.18 (0.16) | −0.13 (0.15) | −0.07 (0.11) | −0.04 (0.11) |
| year=1984 | −0.57 (0.41) | −0.44 (0.38) | **−0.65** (0.24)$^{**}$ | **−0.60** (0.22)$^{**}$ | −0.23 (0.16) | −0.14 (0.15) | −0.04 (0.11) | 0.02 (0.11) |
| year=1985 | −0.19 (0.41) | −0.17 (0.38) | −0.27 (0.24) | −0.27 (0.22) | −0.25 (0.16) | −0.19 (0.14) | −0.04 (0.11) | 0.00 (0.10) |
| year=1986 | −0.59 (0.40) | −0.47 (0.38) | −0.40 (0.24) | −0.33 (0.22) | −0.09 (0.16) | 0.01 (0.15) | −0.02 (0.11) | 0.04 (0.11) |
| year=1987 | **−0.99** (0.43)$^{*}$ | **−0.86** (0.40)$^{*}$ | **−0.67** (0.24)$^{**}$ | **−0.63** (0.22)$^{**}$ | −0.30 (0.16) | −0.23 (0.14) | −0.08 (0.11) | −0.03 (0.10) |
| year=1988 | **−0.97** (0.39)$^{*}$ | **−0.84** (0.37)$^{*}$ | **−0.58** (0.24)$^{*}$ | **−0.52** (0.22)$^{*}$ | **−0.35** (0.15)$^{*}$ | **−0.28** (0.14)$^{*}$ | −0.12 (0.11) | −0.08 (0.10) |
| year=1989 | **−0.96** (0.40)$^{*}$ | **−0.91** (0.38)$^{*}$ | **−0.48** (0.23)$^{*}$ | **−0.44** (0.22)$^{*}$ | **−0.41** (0.15)$^{**}$ | **−0.31** (0.14)$^{*}$ | −0.16 (0.11) | −0.10 (0.10) |
| year=1990 | **−0.94** (0.38)$^{*}$ | **−0.81** (0.36)$^{*}$ | **−0.69** (0.23)$^{**}$ | **−0.64** (0.22)$^{**}$ | **−0.35** (0.15)$^{*}$ | −0.27 (0.14) | −0.13 (0.11) | −0.07 (0.10) |
| year=1991 | **−1.15** (0.38)$^{**}$ | **−1.01** (0.36)$^{**}$ | **−0.73** (0.23)$^{**}$ | **−0.67** (0.22)$^{**}$ | **−0.44** (0.15)$^{**}$ | **−0.34** (0.14)$^{*}$ | −0.18 (0.11) | −0.10 (0.10) |
| year=1992 | **−1.15** (0.38)$^{**}$ | **−1.03** (0.36)$^{**}$ | **−0.93** (0.23)$^{***}$ | **−0.86** (0.22)$^{***}$ | **−0.56** (0.15)$^{***}$ | **−0.43** (0.14)$^{**}$ | **−0.28** (0.11)$^{**}$ | **−0.20** (0.10)$^{*}$ |
| year=1993 | **−1.09** (0.38)$^{**}$ | **−0.92** (0.36)$^{**}$ | **−0.83** (0.23)$^{***}$ | **−0.70** (0.21)$^{**}$ | **−0.52** (0.15)$^{***}$ | **−0.36** (0.14)$^{*}$ | **−0.21** (0.11)$^{*}$ | −0.11 (0.10) |
| year=1994 | **−0.89** (0.38)$^{*}$ | **−0.72** (0.36)$^{*}$ | **−0.62** (0.23)$^{**}$ | **−0.51** (0.21)$^{*}$ | **−0.42** (0.15)$^{**}$ | **−0.28** (0.14)$^{*}$ | −0.20 (0.11) | −0.10 (0.10) |
| year=1995 | **−1.31** (0.39)$^{***}$ | **−1.13** (0.36)$^{**}$ | **−0.88** (0.23)$^{***}$ | **−0.76** (0.22)$^{***}$ | **−0.48** (0.15)$^{**}$ | **−0.34** (0.14)$^{*}$ | **−0.28** (0.11)$^{*}$ | −0.19 (0.10) |
| year=1996 | **−1.15** (0.40)$^{**}$ | **−1.00** (0.37)$^{**}$ | **−0.73** (0.23)$^{**}$ | **−0.60** (0.21)$^{**}$ | **−0.40** (0.15)$^{**}$ | −0.20 (0.14) | **−0.26** (0.11)$^{*}$ | −0.13 (0.10) |
| year=1997 | **−1.09** (0.39)$^{**}$ | **−0.89** (0.36)$^{*}$ | **−0.49** (0.23)$^{*}$ | −0.38 (0.21) | −0.24 (0.15) | −0.10 (0.14) | −0.21 (0.11) | −0.12 (0.10) |
| year=1998 | **−1.01** (0.37)$^{**}$ | **−0.77** (0.35)$^{*}$ | **−0.66** (0.23)$^{**}$ | **−0.50** (0.21)$^{*}$ | **−0.40** (0.15)$^{**}$ | −0.18 (0.14) | **−0.30** (0.11)$^{**}$ | −0.15 (0.10) |
| year=1999 | **−1.00** (0.38)$^{**}$ | **−0.76** (0.35)$^{*}$ | **−0.77** (0.22)$^{***}$ | **−0.56** (0.21)$^{**}$ | **−0.43** (0.15)$^{**}$ | −0.18 (0.14) | **−0.35** (0.11)$^{**}$ | −0.18 (0.10) |
| year=2000 | **−0.92** (0.38)$^{*}$ | −0.68 (0.36) | **−0.67** (0.23)$^{**}$ | **−0.47** (0.21)$^{*}$ | **−0.34** (0.15)$^{*}$ | −0.10 (0.14) | **−0.31** (0.11)$^{**}$ | −0.15 (0.10) |
| year=2001 | **−0.93** (0.38)$^{*}$ | −0.66 (0.35) | **−0.67** (0.22)$^{**}$ | **−0.47** (0.21)$^{*}$ | **−0.33** (0.15)$^{*}$ | −0.06 (0.14) | **−0.30** (0.11)$^{**}$ | −0.12 (0.10) |
| year=2002 | **−0.97** (0.37)$^{**}$ | **−0.73** (0.35)$^{*}$ | **−0.62** (0.22)$^{**}$ | **−0.43** (0.21)$^{*}$ | **−0.39** (0.15)$^{**}$ | −0.13 (0.14) | **−0.35** (0.11)$^{**}$ | −0.17 (0.10) |
| year=2003 | **−0.90** (0.37)$^{*}$ | −0.61 (0.34) | **−0.61** (0.22)$^{**}$ | −0.37 (0.21) | **−0.40** (0.15)$^{**}$ | −0.07 (0.14) | **−0.40** (0.11)$^{***}$ | −0.18 (0.10) |
| year=2004 | **−0.73** (0.36)$^{*}$ | −0.48 (0.34) | **−0.48** (0.22)$^{*}$ | −0.31 (0.20) | **−0.31** (0.15)$^{*}$ | −0.07 (0.13) | **−0.35** (0.10)$^{***}$ | −0.19 (0.10) |
| year=2005 | **−0.86** (0.36)$^{*}$ | −0.63 (0.34) | **−0.49** (0.22)$^{*}$ | −0.34 (0.20) | **−0.30** (0.15)$^{*}$ | −0.08 (0.13) | **−0.37** (0.10)$^{***}$ | **−0.22** (0.10)$^{*}$ |
| year=2006 | **−0.91** (0.36)$^{*}$ | −0.65 (0.34) | **−0.62** (0.22)$^{**}$ | **−0.42** (0.21)$^{*}$ | **−0.38** (0.15)$^{**}$ | −0.10 (0.13) | **−0.44** (0.11)$^{***}$ | **−0.26** (0.10)$^{**}$ |
| year=2007 | **−0.96** (0.36)$^{**}$ | **−0.68** (0.34)$^{*}$ | **−0.68** (0.22)$^{**}$ | **−0.46** (0.20)$^{*}$ | **−0.43** (0.15)$^{**}$ | −0.13 (0.13) | **−0.50** (0.10)$^{***}$ | **−0.30** (0.10)$^{**}$ |
| year=2008 | **−0.88** (0.36)$^{*}$ | −0.59 (0.34) | **−0.60** (0.22)$^{**}$ | −0.35 (0.20) | **−0.41** (0.14)$^{**}$ | −0.06 (0.13) | **−0.51** (0.10)$^{***}$ | **−0.28** (0.10)$^{**}$ |
| year=2009 | **−1.01** (0.36)$^{**}$ | **−0.68** (0.34)$^{*}$ | **−0.76** (0.22)$^{***}$ | **−0.48** (0.20)$^{*}$ | **−0.49** (0.14)$^{***}$ | −0.10 (0.13) | **−0.59** (0.10)$^{***}$ | **−0.34** (0.10)$^{***}$ |
| year=2010 | **−0.90** (0.36)$^{*}$ | −0.58 (0.34) | **−0.66** (0.22)$^{**}$ | −0.38 (0.20) | **−0.39** (0.14)$^{**}$ | −0.00 (0.13) | **−0.53** (0.10)$^{***}$ | **−0.27** (0.10)$^{**}$ |
| year=2011 | **−1.00** (0.36)$^{**}$ | **−0.67** (0.34)$^{*}$ | **−0.72** (0.22)$^{***}$ | **−0.40** (0.20)$^{*}$ | **−0.47** (0.14)$^{**}$ | −0.03 (0.13) | **−0.64** (0.10)$^{***}$ | **−0.35** (0.10)$^{***}$ |
| year=2012 | **−0.94** (0.36)$^{**}$ | −0.54 (0.34) | **−0.67** (0.22)$^{**}$ | −0.30 (0.20) | **−0.44** (0.14)$^{**}$ | 0.07 (0.13) | **−0.65** (0.10)$^{***}$ | **−0.32** (0.10)$^{**}$ |
| year=2013 | **−1.03** (0.36)$^{**}$ | −0.61 (0.34) | **−0.78** (0.22)$^{***}$ | −0.39 (0.20) | | | **−0.75** (0.10)$^{***}$ | **−0.40** (0.10)$^{***}$ |
| year=2014 | **−0.96** (0.36)$^{**}$ | −0.50 (0.34) | **−0.73** (0.22)$^{***}$ | −0.30 (0.20) | | | **−0.76** (0.10)$^{***}$ | **−0.38** (0.10)$^{***}$ |
| year=2015 | **−0.94** (0.36)$^{**}$ | −0.47 (0.34) | **−0.68** (0.22)$^{**}$ | −0.22 (0.21) | | | **−0.73** (0.10)$^{***}$ | **−0.33** (0.10)$^{***}$ |
| year=2016 | **−0.97** (0.36)$^{**}$ | −0.42 (0.34) | **−0.67** (0.22)$^{**}$ | −0.16 (0.21) | | | **−0.75** (0.10)$^{***}$ | **−0.31** (0.10)$^{**}$ |
| year=2017 | **−0.85** (0.36)$^{*}$ | −0.36 (0.34) | **−0.55** (0.22)$^{*}$ | −0.07 (0.21) | | | **−0.70** (0.10)$^{***}$ | **−0.27** (0.10)$^{**}$ |
| year=2018 | **−0.87** (0.36)$^{*}$ | −0.38 (0.34) | | | | | **−0.77** (0.11)$^{***}$ | **−0.34** (0.10)$^{***}$ |
| year=2019 | **−0.73** (0.36)$^{*}$ | −0.25 (0.34) | | | | | **−0.70** (0.11)$^{***}$ | **−0.26** (0.10)$^{*}$ |
| nRefs | | **−0.01** (0.00)$^{***}$ | | **−0.01** (0.00)$^{***}$ | | **−0.02** (0.00)$^{***}$ | | **−0.01** (0.00)$^{***}$ |
| nRefs_Missing | | **0.02** (0.00)$^{***}$ | | **0.02** (0.00)$^{***}$ | | **0.03** (0.00)$^{***}$ | | **0.02** (0.00)$^{***}$ |
| nAuthors | | 0.00 (0.01) | | 0.00 (0.01) | | −0.00 (0.01) | | 0.01 (0.01) |
| articleLength | | −0.00 (0.00) | | −0.00 (0.00) | | −0.00 (0.00) | | **−0.00** (0.00)$^{*}$ |
| $R^2$ | 0.04 | 0.16 | 0.04 | 0.16 | 0.04 | 0.20 | 0.21 | 0.32 |
| Adj. $R^2$ | 0.02 | 0.14 | 0.03 | 0.15 | 0.03 | 0.19 | 0.21 | 0.32 |
| Num. obs. | 2234 | 2234 | 3608 | 3605 | 3919 | 3917 | 5673 | 5670 |
| F statistic (Wald) | **2.08**$^{***}$ | **8.44**$^{***}$ | **3.33**$^{***}$ | **14.58**$^{***}$ | **3.75**$^{***}$ | **23.19**$^{***}$ | **34.49**$^{***}$ | **55.45**$^{***}$ |

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

On average, the number of missing references has a positive effect on the disruption score, ceteris paribus. More importantly, including more references has, on average, a negative effect on the disruption score, ceteris paribus. Combined, these coefficient signs confirm the earlier-raised suspicion that disruption scores of earlier papers are inflated.

Across all baseline models, except for the '2022' window, every coefficient of a year indicator later than 1988 is negative and statistically significant (at $\alpha = 0.05$). For the '2022' baseline model, this year is 1992. However, many of these previously statistically significant yearly indicators are not significant any longer. Across the 3-year, 5-year, and 10-year controlled models, only the indicators for the years 1988-1995 are uninterruptedly significant. This means that only those years have significantly different disruption scores compared to the baseline year of 1975, ceteris paribus.

In order to aid interpretation and comparison, I present in Figure 14 below the individual predictions for the year indicators in a single line (for the field of Marketing). These predictions are made by fixing the value of the control variables, at their respective sample means. It is important to note that even though many of the yearly indicator variables are not significant, they can still be employed to make predictions.




(a) 3-year (b) 5-year

(c) 10-year (d) 2022

Figure 14: This figure displays the predicted $CD_5^{nob}$ value based on the controlled models in Tables 11 and 19 (Appendix), for the four robust Marketing and AI samples. These predictions are made by fixing the value of the control variables, at their respective sample means.

These regression-adjusted generally follow similar trends compared to the unadjusted graphs above in Figure 13. However, the surge in disruptiveness during recent years in the field of Marketing is substantially more pronounced in these time-corrected graphs.

# 8 Conclusion

## 8.1 Main findings

Firstly, regarding the different variations of the CD index, results show that the $N_b$ term indeed primarily reflects citation impact, since it is multiple orders of magnitude larger than the other $N_f$ and $N_{bf}$ terms. Therefore, it is too dominant for adequately capturing nuances in disruption. This is in line with the findings from Bornmann et al. (2020) and Q. Wu and Yan (2019), but contradicts the claim of Funk and Owen-Smith (2017) that disruptive papers, according to the CD index, do not need to have a large following. Consequently, I use the $CD_{l=5}^{nob}$ variant throughout the analysis. Nonetheless, this variant has its limitations which will be discussed below (section Limitations & further research).

Secondly, in general, the 3-year, 5-year, and 10-year citation windows samples have similar results but are different than the results for the '2022' window. The results for the '2022' window are likely divergent because, by construction, older publications have a longer period of time to amass citations. One could argue that it is, thus, not fair to compare the most recent publication years with older ones. Presumably, for that reason, the coefficient values of the year indicator for recent years are still statistically significant in the '2022' controlled model (Table 11), compared to the other three controlled models. Therefore, I will base my conclusion principally on the 3-year, 5-year, and 10-year citation windows samples.

Furthermore, controlled regressions (Table 11) point to a negative effect of the number of references on a paper's disruptiveness score, in accordance with Ruan et al. (2021). Additionally, this effect is strengthened by the number of missing references, which has a positive effect on the publication score (i.e., when there are more references missing a paper is more disruptive). These regressions imply the discovered trends are robust.

On the field level, throughout most of the sample period, AI led the way in disruption compared to Marketing. For Marketing, the least disruptive period is around 1990-1995. Nevertheless, after this trough, the average disruption score fairly consistently increases up until and including the final year of all three samples. During these final years, the average level of disruption is approximately back at the level from the start of the sample period. This roughly u-shaped trend is not in line with previously mentioned findings such as the downward trend discovered by Park et al. (2023) (although on a more aggregate level; for the social sciences) or the stagnating trend in recent years (in science overall) found by Milojević (OECD, 2023).

Compared to the field of AI, trends indicate the field of Marketing will have a higher average disruption score in the immediate years post-sample, although the trend in AI in the final years remains somewhat inconclusive due to divergent trends among citation windows (Figures 14a, 14b, and 14c). Interestingly, this is divergent compared to trends of the mere citation count from both fields (Figure 7), which indicate a higher citation impact for the field of AI. Nevertheless, this is evidence in favor of publications not needing a large impact to be disruptive.

On the sub-field level, among the four prominent Marketing journals, JM appears to be the least disruptive, especially in the years after 2000. This trend generally coincides with the study of Yadav (2010), which finds that JM experienced a large decline in conceptual papers. Similar to the field-level findings, the apparent higher citation impact of JM compared to the other journals (Figure 5) does not seem to translate to its disruption scores. MKS looks to be consistently among the most disruptive papers while having the fewest citations among journals (Table 4). Likewise, this is in line with the findings of Yadav (2010). He discovers that these conceptual articles in MKS are mainly focused on theory development, instead of theory testing, which could explain their disruptive nature. JCR and JMR, on the other hand, remain close to the field's average to a great extent.

### 8.2 Managerial and academic implications

In sum, this study employs (variations of) the CD index, a measure based on the fundamental theories of science, to quantify the degree to which a publication disrupts future research. A comprehensive review of the existing disruption indicators is provided, preceded by an introduction to the techniques for quantifying scientific advancements based on citation data. The CD index is a relatively new indicator that can complement existing measures by providing data on the way publications cause shifts in the research paths pursued by scientists. It is inherently distinct from existing measures, such as novelty, since it does not take into account how existing knowledge is integrated from the input side but assesses a paper's citation precedence years after publication.

Thereby, this paper offers new insights into how the field of Marketing forms new knowledge, as a field as a whole but also on a journal level. Although the CD index undeniably has its limitations (see Limitations & further research), I think it can aid Marketing science policymakers, journal editors, and other relevant stakeholders in evaluating and directing research.

Furthermore, this study provides a detailed overview of the citation and reference patterns over time, on the field and sub-field levels. These data are relevant to anyone who intends to review or bibliometrically analyze the field of Marketing and can easily be operationalized in subsequent studies.

Finally, even though not one of the main goals of this thesis and therefore not directly presented, one can use individual disruption scores to guide one's research. By means of looking at high-scoring studies, a researcher can find disruptive publications that challenge fundamental theories and problems within the field.

## 8.3 Limitations & further research

The CD index variant $CD_{l=5}^{nob}$ that I use throughout is, in my opinion, the most robust disruption indicator currently available. However, it does not correct for changes in publication and citation practices over time, and neither the normalized indicators proposed by Park et al. (2023) achieve this goal. For that reason, I employ a regression analysis, but ideally, one prefers to be able to deploy an indicator that is inherently normalized to aid the comparison of disruption scores across time and fields. Concurrently, future research should aim to improve the robustness of the CD index and its normalized variants. Moreover, researchers need to precisely define disruption in the context that they are studying it because many of the CD index variants seem to be heavily influenced by citation impact.

Secondly, the use of Web of Science data proves to suffer from strong limitations. Early publications are missing a substantial number of their references which leads to selection bias and consequently inflated disruption scores. This could particularly influence Marketing scores since the social sciences appear to have inferior coverage in the WoS. For that reason, future studies can aim to minimize this selection bias by combining data from multiple bibliometric databases or focusing on more recent publications so that the best possible coverage is achieved.

Finally, future research could aim to highlight more in-depth the disruptiveness of individual Marketing publications. Accordingly, one can examine these publications and study their relationship with different characteristics such as (citation) impact and novelty.

# 9 References


Abernathy, W. J., & Clark, K. B. (1985). Innovation: Mapping the winds of creative destruction. *Research policy*, *14*(1), 3–22.

Abramo, G., & D'Angelo, C. A. (2011). Evaluating research: From informed peer review to bibliometrics. *Scientometrics*, *87*(3), 499–514.

Ahmed, N., & Wahed, M. (2020). The de-democratization of ai: Deep learning and the compute divide in artificial intelligence research. *arXiv preprint arXiv:2010.15581*.

Aksnes, D. W., Langfeldt, L., & Wouters, P. (2019). Citations, citation indicators, and research quality: An overview of basic concepts and theories. *Sage Open*, *9*(1), 2158244019829575.

Bloom, N., Jones, C. I., Van Reenen, J., & Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, *110*(4), 1104–1144.

Bornmann, L., Devarakonda, S., Tekles, A., & Chacko, G. (2020). Are disruption index indicators convergently valid? the comparison of several indicator variants with assessments by peers. *Quantitative Science Studies*, *1*(3), 1242–1259.

Bornmann, L., & Marx, W. (2015). Methods for the generation of normalized citation impact scores in bibliometrics: Which method best reflects the judgements of experts? *Journal of Informetrics*, *9*(2), 408–418.

Bornmann, L., & Mutz, R. (2015). Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. *Journal of the association for information science and technology*, *66*(11), 2215–2222.

Bornmann, L., Mutz, R., Neuhaus, C., & Daniel, H.-D. (2008). Citation counts for research evaluation: Standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in science and environmental politics*, *8*(1), 93–102.

Bornmann, L., & Tekles, A. (2019). Disruption index depends on length of citation window. *Profesional de la información*, *28*(2).

Bower, J. L., & Christensen, C. M. (1995). Disruptive technologies: Catching the wave.

Bu, Y., Waltman, L., & Huang, Y. (2021). A multidimensional framework for characterizing the citation impact of scientific publications. *Quantitative science studies*, *2*(1), 155–183.

Casadevall, A., & Fang, F. C. (2016). Revolutionary science.

Chen, J., Shao, D., & Fan, S. (2021). Destabilization and consolidation: Conceptualizing, measuring, and validating the dual characteristics of technology. *Research policy*, *50*(1), 104115.

Christensen, C., Raynor, M. E., & McDonald, R. (2015). *What is disruptive innovation?* (Vol. 93). Harvard Business Review.

Christensen, C., & Raynor, M. (2003). *The innovator's solution: Creating and sustaining successful growth*. Harvard Business Review Press.

Christensen, C. M. (1997). *The innovator's dilemma: When new technologies cause great firms to fail*. Harvard Business Review Press.

Chu, J. S., & Evans, J. A. (2021). Slowed canonical progress in large fields of science. *Proceedings of the National Academy of Sciences*, *118*(41), e2021636118.

Ciarli, T., Kenney, M., Massini, S., & Piscitello, L. (2021). Digital technologies, innovation, and skills: Emerging trajectories and challenges. *Research Policy*, *50*(7), 104289.

Costas, R., van Leeuwen, T., & Bordons, M. (2010). Self-citations at the meso and individual levels: Effects of different calculation methods. *Scientometrics*, *82*(3), 517–537.

Dan, Y., & Chieh, H. C. (2008). A reflective review of disruptive innovation theory. *PICMET'08-2008 Portland International Conference on Management of Engineering & Technology*, 402–414.

Davis, M. S. (1971). That's interesting! towards a phenomenology of sociology and a sociology of phenomenology. *Philosophy of the social sciences*, *1*(2), 309–344.

Deng, N., & Zeng, A. (2023). Enhancing the robustness of the disruption metric against noise. *Scientometrics*, *128*(4), 2419–2428.

Funk, R. J., & Owen-Smith, J. (2017). A dynamic network measure of technological change. *Management science*, *63*(3), 791–817.

Garfield, E. (1955). Citation indexes for science: A new dimension in documentation through association of ideas. *Science*, *122*(3159), 108–111.

Garfield, E. (1964). " science citation index"—a new dimension in indexing: This unique approach underlies versatile bibliographic systems for communicating and evaluating information. *Science*, *144*(3619), 649–654.

Garfield, E. (1972). Citation analysis as a tool in journal evaluation: Journals can be ranked by frequency and impact of citations for science policy studies. *Science*, *178*(4060), 471–479.

Glänzel, W., Schlemmer, B., & Thijs, B. (2003). Better late than never? on the chance to become highly cited only beyond the standard bibliometric time horizon. *Scientometrics*, *58*, 571–586.

Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). Bibliometrics: The leiden manifesto for research metrics. *Nature*, *520*(7548), 429–431.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences*, *102*(46), 16569–16572.

Hoffman, D. L., Moreau, C. P., Stremersch, S., & Wedel, M. (2022). The rise of new technologies in marketing: A framework and outlook.

Hofstra, B., Kulkarni, V. V., Munoz-Najar Galvez, S., He, B., Jurafsky, D., & McFarland, D. A. (2020). The diversity–innovation paradox in science. *Proceedings of the National Academy of Sciences*, *117*(17), 9284–9291.

Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS review*, *10*(1-2), 18–26.

Jaffe, A. B., & Trajtenberg, M. (2002). *Patents, citations, and innovations: A window on the knowledge economy*. MIT press.

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.

Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American documentation*, *14*(1), 10–25.

Klinger, J., Mateos-Garcia, J., & Stathoulopoulos, K. (2020). A narrowing of ai research? *arXiv preprint arXiv:2009.10385*.

Kohli, A. K., & Haenlein, M. (2021). Factors affecting the study of important marketing issues: Implications and recommendations. *International Journal of Research in Marketing*, *38*(1), 1–11.

Kuhn, T. S. (1962). The structure of scientific revolutions.

Lee, Y.-N., Walsh, J. P., & Wang, J. (2015). Creativity in scientific teams: Unpacking novelty and impact. *Research policy*, *44*(3), 684–697.

Lehmann, D. R., McAlister, L., & Staelin, R. (2011). Sophistication in research in marketing. *Journal of Marketing*, *75*(4), 155–165.

Leydesdorff, L., Tekles, A., & Bornmann, L. (2021). A proposal to revise the disruption index. *Profesional de la información (EPI)*, *30*(1).

Liang, G., Jiang, Y., & Hou, H. (2020). Same data may bring conflict results: A caution to use the disruptive index. *arXiv preprint arXiv:2009.06888*.

Lin, Y., Evans, J. A., & Wu, L. (2022). New directions in science emerge from disconnection and discord. *Journal of Informetrics*, *16*(1), 101234.

Lu, C., Bu, Y., Dong, X., Wang, J., Ding, Y., Larivière, V., Sugimoto, C. R., Paul, L., & Zhang, C. (2019). Analyzing linguistic complexity and scientific impact. *Journal of Informetrics*, *13*(3), 817–829.

MacInnis, D. J. (2011). A framework for conceptual contributions in marketing. *Journal of Marketing*, *75*(4), 136–154.

MacInnis, D. J., Morwitz, V. G., Botti, S., Hoffman, D. L., Kozinets, R. V., Lehmann, D. R., Lynch Jr, J. G., & Pechmann, C. (2020). Creating boundary-breaking, marketing-relevant consumer research. *Journal of Marketing*, *84*(2), 1–23.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization science*, *2*(1), 71–87.

Marx, K., & Engels, F. (1848). *The communist manifesto*. the Workers' Educational Association.

Mela, C. F., Roos, J., & Deng, Y. (2013). A keyword history of marketing science. *Marketing Science*, *32*(1), 8–18.

Milojević, S. (2015). Quantifying the cognitive extent of science. *Journal of Informetrics*, *9*(4), 962–973.

Moed, H. F. (2006). *Citation analysis in research evaluation* (Vol. 9). Springer Science & Business Media.

Moorman, C., van Heerde, H. J., Moreau, C. P., & Palmatier, R. W. (2019a). Challenging the boundaries of marketing.

Moorman, C., van Heerde, H. J., Moreau, C. P., & Palmatier, R. W. (2019b). Jm as a marketplace of ideas.

Muller, E. (2020). Delimiting disruption: Why uber is disruptive, but airbnb is not. *International Journal of Research in Marketing*, *37*(1), 43–55.

OECD. (2021). Effective policies to foster high-risk/high-reward research. (112). https://www.oecdilibrary.org/content/paper/06913b3b-en

OECD. (2023). *Artificial intelligence in science*. https://www.oecdilibrary.org/content/publication/a8d820bd-en

Park, M., Leahey, E., & Funk, R. J. (2023). Papers and patents are becoming less disruptive over time. *Nature*, *613*(7942), 138–144.

Reibstein, D. J., Day, G., & Wind, J. (2009). Guest editorial: Is marketing academia losing its way? *Journal of marketing*, *73*(4), 1–3.

Ruan, X., Ao, W., Lyu, D., Cheng, Y., & Li, J. (2023). Effect of the topic-combination novelty on the disruption and impact of scientific articles: Evidence from pubmed. *Journal of Information Science*, 01655515231161133.

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.

Schumpeter, J. A. (1942). *Capitalism, socialism and democracy*. routledge.

Segerstrom, P. S. (1991). Innovation, imitation, and economic growth. *Journal of political economy*, *99*(4), 807–827.

Shugan, S. M. (2003). Defining interesting research problems.

Si, S., & Chen, H. (2020). A literature review of disruptive innovation: What it is, how it works and where it goes. *Journal of Engineering and Technology Management*, *56*, 101568.

Sridhar, S., Lamberton, C., Marinova, D., & Swaminathan, V. (2023). Jm: Promoting catalysis in marketing scholarship.

Staudt, J., Yu, H., Light, R. P., Marschke, G., Börner, K., & Weinberg, B. A. (2018). High-impact and transformative science (hits) metrics: Definition, exemplification, and comparison. *PloS one*, *13*(7), e0200597.

Stremersch, S. (2021). The study of important marketing issues: Reflections. *International Journal of Research in Marketing*, *38*(1), 12–17.

Stremersch, S., Camacho, N., Vanneste, S., & Verniers, I. (2015). Unraveling scientific impact: Citation types in marketing journals. *International Journal of Research in Marketing*, *32*(1), 64–77.

Stremersch, S., Winer, R. S., & Camacho, N. (2021). Faculty research incentives and business school health: A new perspective from and for marketing. *Journal of Marketing*, *85*(5), 1–21.

Teplitskiy, M., Duede, E., Menietti, M., & Lakhani, K. R. (2022). How status of research papers affects the way they are read and cited. *Research Policy*, *51*(4), 104484.

Tushman, M. L., & Anderson, P. (1986). Technological discontinuities and organizational environments. *Administrative science quarterly*, 439–465.

Ulaga, W., Kleinaltenkamp, M., Kashyap, V., & Eggert, A. (2021). Advancing marketing theory and practice: Guidelines for crafting research propositions. *AMS Review*, 1–12.

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, *342*(6157), 468–472.

Van Raan, A. F. (2004). Sleeping beauties in science. *Scientometrics*, *59*(3), 467–472.

Van Raan, A. F., Van Leeuwen, T. N., & Visser, M. S. (2011). Severe language effect in university rankings: Particularly germany and france are wronged in citation-based rankings. *Scientometrics*, *88*(2), 495–498.

Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of informetrics*, *10*(2), 365–391.

Waltman, L., & van Eck, N. J. (2013). Source normalized indicators of citation impact: An overview of different approaches and an empirical comparison. *Scientometrics*, *96*, 699–716.

Wang, D., & Barabási, A.-L. (2021). *The science of science*. Cambridge University Press.

Wang, D., Song, C., & Barabási, A.-L. (2013). Quantifying long-term scientific impact. *Science*, *342*(6154), 127–132.

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.

Weitzman, M. L. (1998). Recombinant growth. *The Quarterly Journal of Economics*, *113*(2), 331–360.

Whittaker, M. (2021). The steep cost of capture. *Interactions*, *28*(6), 50–55.

Whittlestone, J., & Clark, J. (2021). Why and how governments should monitor ai development. *arXiv preprint arXiv:2108.12427*.

Wu, L., Kittur, A., Youn, H., Milojević, S., Leahey, E., Fiore, S. M., & Ahn, Y.-Y. (2022). Metrics and mechanisms: Measuring the unmeasurable in the science of science. *Journal of Informetrics*, *16*(2), 101290.

Wu, L., Wang, D., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, *566*(7744), 378–382.

Wu, Q., & Yan, Z. (2019). Solo citations, duet citations, and prelude citations: New measures of the disruption of academic papers. *arXiv preprint arXiv:1905.03461*.

Wu, S., & Wu, Q. (2019). A confusing definition of disruption.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036–1039.

Yadav, M. S. (2010). The decline of conceptual articles and implications for knowledge development. *Journal of Marketing*, *74*(1), 1–19.

Zitt, M., & Small, H. (2008). Modifying the journal impact factor by fractional citation weighting: The audience factor. *Journal of the American Society for Information Science and technology*, *59*(11), 1856–1860.

# 10 Appendix

## 10.1 Research Background: Trend in AI

Since the field of AI is primarily employed as a benchmark in this paper, I will keep this section concise. There exist no studies that I am aware of on trends that explicitly reflect disruptiveness in Artificial Intelligence research or related concepts. Nonetheless, there exists some alternative evidence of a probable narrowing of AI research.

Klinger et al. (2020) perform a semantic analysis of AI studies in the broadly used pre-print repository *arXiv* by measuring thematic diversity in paper abstracts. They find that even though technological diversity in AI research has increased throughout the early 2000s, the growth has reached a standstill and, starting from the mid-2010s, has even begun to decline. One could be surprised by this finding because the corpus of AI literature has vastly expanded in recent years and appears to keep growing at faster rates (Klinger et al., 2020). Sixty percent of the papers included in the study of Klinger et al. (2020) were published after 2018.

In the case that one wants to learn more about AI-specific trends, they could look at papers by Ciarli et al. (2021) and Whittlestone and Clark (2021).

### 10.1.1 Potential explanations

Klinger et al. (2020) discover that large technology companies such as Google and Microsoft are more likely to narrow their focus on a more limited set of cutting-edge methods and techniques compared to universities, specifically the current paradigm of the compute-intensive *deep learning*. However, some of the top institutions, such as MIT, Stanford University, and the University of California, Berkeley, have lower-than-expected levels of thematic diversity. Ahmed and Wahed (2020) uncover that such elite academic institutions are the top collaborators of private companies. In contrast, they find that there is an increasing *compute divide* between non-elite universities and the private sector.

Whittaker (2021) points out that the increasing need for access to extensive datasets and infrastructures, essential for cutting-edge research, is a likely driver of these skewed research priorities. Moreover, Ahmed and Wahed (2020) detect that AI research with involvement from the private sector tends to be more highly cited. Therefore, private companies could be directly influencing the field's evolution by means of the research they publish, a trend Ahmed and Wahed (2020) refer as the *de-democratisation* of AI research.

## 10.2 Data

### 10.2.1 Marketing

Table 12: This table reports the number of focal publications included in the Marketing samples per year per journal.

| Pub. year | Number of publications | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Citation window | | | | | | | | | | | |
| | 3-year/2022 | | | | 5-year | | | | 10-year | | | |
| | JCR | JM | JMR | MKS | JCR | JM | JMR | MKS | JCR | JM | JMR | MKS |
| 1975 | 30 | 42 | 44 | | 30 | 42 | 44 | | 30 | 42 | 44 | |
| 1976 | 20 | 47 | 36 | | 20 | 47 | 36 | | 20 | 47 | 36 | |
| 1977 | 26 | 51 | 50 | | 26 | 51 | 50 | | 26 | 51 | 50 | |
| 1978 | 26 | 51 | 42 | | 26 | 51 | 42 | | 26 | 51 | 42 | |
| 1979 | 31 | 35 | 48 | | 31 | 35 | 48 | | 31 | 35 | 48 | |
| 1980 | 36 | 37 | 44 | | 36 | 37 | 44 | | 36 | 37 | 44 | |
| 1981 | 45 | 49 | 36 | | 45 | 49 | 36 | | 45 | 49 | 36 | |
| 1982 | 33 | 43 | 50 | | 33 | 43 | 50 | | 33 | 43 | 50 | |
| 1983 | 33 | 39 | 33 | | 33 | 39 | 33 | | 33 | 39 | 33 | |
| 1984 | 38 | 33 | 38 | | 38 | 33 | 38 | | 38 | 33 | 38 | |
| 1985 | 36 | 49 | 37 | | 36 | 49 | 37 | | 36 | 49 | 37 | |
| 1986 | 31 | 29 | 31 | | 31 | 29 | 31 | | 31 | 29 | 31 | |
| 1987 | 37 | 32 | 34 | 24 | 37 | 32 | 34 | 24 | 37 | 32 | 34 | 24 |
| 1988 | 43 | 28 | 28 | 19 | 43 | 28 | 28 | 19 | 43 | 28 | 28 | 19 |
| 1989 | 39 | 24 | 31 | 20 | 39 | 24 | 31 | 20 | 39 | 24 | 31 | 20 |
| 1990 | 42 | 27 | 33 | 20 | 42 | 27 | 33 | 20 | 42 | 27 | 33 | 20 |
| 1991 | 41 | 21 | 28 | 21 | 41 | 21 | 28 | 21 | 41 | 21 | 28 | 21 |
| 1992 | 46 | 24 | 34 | 24 | 46 | 24 | 34 | 24 | 46 | 24 | 34 | 24 |
| 1993 | 40 | 25 | 35 | 23 | 40 | 25 | 35 | 23 | 40 | 25 | 35 | 23 |
| 1994 | 47 | 33 | 41 | 22 | 47 | 33 | 41 | 22 | 47 | 33 | 41 | 22 |
| 1995 | 34 | 25 | 37 | 44 | 34 | 25 | 37 | 44 | 34 | 25 | 37 | 44 |
| 1996 | 26 | 25 | 34 | 20 | 26 | 25 | 34 | 20 | 26 | 25 | 34 | 20 |
| 1997 | 30 | 23 | 36 | 21 | 30 | 23 | 36 | 21 | 30 | 23 | 36 | 21 |
| 1998 | 22 | 28 | 38 | 24 | 22 | 28 | 38 | 24 | 22 | 28 | 38 | 24 |
| 1999 | 21 | 41 | 38 | 31 | 21 | 41 | 38 | 31 | 21 | 41 | 38 | 31 |
| 2000 | 30 | 22 | 37 | 22 | 30 | 22 | 37 | 22 | 30 | 22 | 37 | 22 |
| 2001 | 41 | 20 | 39 | 26 | 41 | 20 | 39 | 26 | 41 | 20 | 39 | 26 |
| 2002 | 36 | 23 | 36 | 21 | 36 | 23 | 36 | 21 | 36 | 23 | 36 | 21 |
| 2003 | 38 | 31 | 35 | 23 | 38 | 31 | 35 | 23 | 38 | 31 | 35 | 23 |
| 2004 | 74 | 34 | 37 | 41 | 74 | 34 | 37 | 41 | 74 | 34 | 37 | 41 |
| 2005 | 64 | 41 | 53 | 44 | 64 | 41 | 53 | 44 | 64 | 41 | 53 | 44 |
| 2006 | 57 | 42 | 57 | 31 | 57 | 42 | 57 | 31 | 57 | 42 | 57 | 31 |
| 2007 | 61 | 41 | 56 | 56 | 61 | 41 | 56 | 56 | 61 | 41 | 56 | 56 |
| 2008 | 74 | 50 | 53 | 73 | 74 | 50 | 53 | 73 | 74 | 50 | 53 | 73 |
| 2009 | 77 | 54 | 60 | 70 | 77 | 54 | 60 | 70 | 77 | 54 | 60 | 70 |
| 2010 | 69 | 45 | 90 | 67 | 69 | 45 | 90 | 67 | 69 | 45 | 90 | 67 |
| 2011 | 73 | 55 | 90 | 67 | 73 | 55 | 90 | 67 | 73 | 55 | 90 | 67 |
| 2012 | 84 | 46 | 70 | 53 | 84 | 46 | 70 | 53 | 84 | 46 | 70 | 53 |
| 2013 | 77 | 47 | 50 | 53 | 77 | 47 | 50 | 53 | | | | |
| 2014 | 90 | 45 | 51 | 48 | 90 | 45 | 51 | 48 | | | | |
| 2015 | 54 | 37 | 56 | 50 | 54 | 37 | 56 | 50 | | | | |
| 2016 | 57 | 37 | 66 | 52 | 57 | 37 | 66 | 52 | | | | |
| 2017 | 69 | 47 | 62 | 52 | 69 | 47 | 62 | 52 | | | | |
| 2018 | 67 | 46 | 59 | 50 | | | | | | | | |
| 2019 | 68 | 43 | 59 | 49 | | | | | | | | |
| **Total** | 2113 | 1667 | 2052 | 1261 | 1978 | 1578 | 1934 | 1162 | 1631 | 1365 | 1649 | 907 |
| | 7093 | | | | 6652 | | | | 5552 | | | |

### 10.2.2 AI

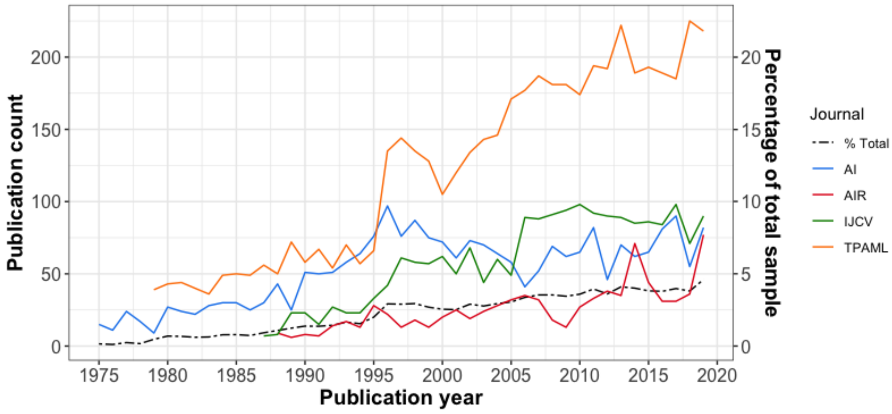
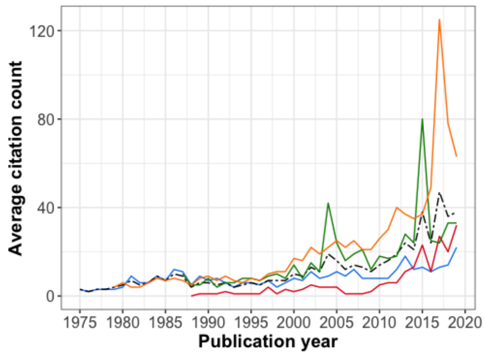


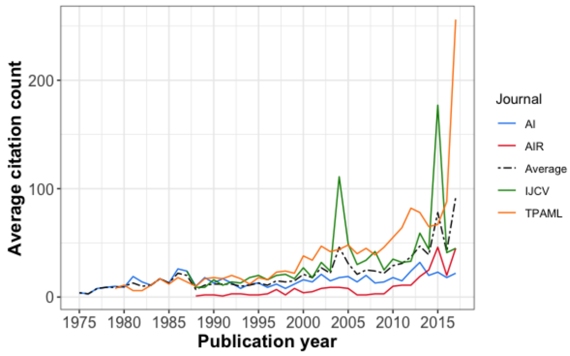

Figure 15: This figure displays the publication count per journal for the full AI sample.

Table 13: This table reports the number of focal publications included in the AI samples per year per journal.

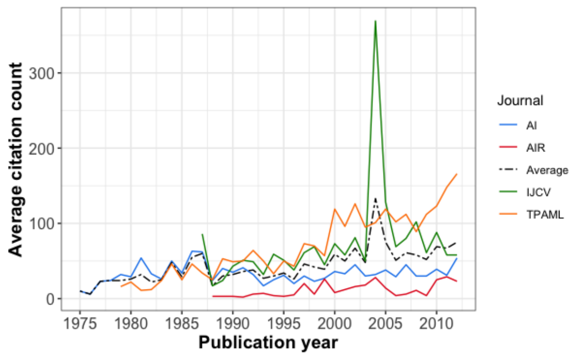
| **Pub. year** | **Number of publications** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Citation window | | | | | | | | | | | |
| | 3-year/2022 | | | | 5-year | | | | 10-year | | | |
| | AI | AIR | IJCV | TPAML | AI | AIR | IJCV | TPAML | AI | AIR | IJCV | TPAML |
| 1975 | 15 | | | | 15 | | | | 15 | | | |
| 1976 | 11 | | | | 11 | | | | 11 | | | |
| 1977 | 24 | | | | 24 | | | | 24 | | | |
| 1978 | 17 | | | | 17 | | | | 17 | | | |
| 1979 | 9 | | | 39 | 9 | | | 39 | 9 | | | 39 |
| 1980 | 27 | | | 43 | 27 | | | 43 | 27 | | | 43 |
| 1981 | 24 | | | 44 | 24 | | | 44 | 24 | | | 44 |
| 1982 | 22 | | | 40 | 22 | | | 40 | 22 | | | 40 |
| 1983 | 28 | | | 36 | 28 | | | 36 | 28 | | | 36 |
| 1984 | 30 | | | 49 | 30 | | | 49 | 30 | | | 49 |
| 1985 | 30 | | | 50 | 30 | | | 50 | 30 | | | 50 |
| 1986 | 25 | | | 49 | 25 | | | 49 | 25 | | | 49 |
| 1987 | 30 | | 7 | 56 | 30 | | 7 | 56 | 30 | | 7 | 56 |
| 1988 | 43 | 9 | 8 | 50 | 43 | 9 | 8 | 50 | 43 | 9 | 8 | 50 |
| 1989 | 25 | 6 | 23 | 72 | 25 | 6 | 23 | 72 | 25 | 6 | 23 | 72 |
| 1990 | 51 | 8 | 23 | 58 | 51 | 8 | 23 | 58 | 51 | 8 | 23 | 58 |
| 1991 | 50 | 7 | 15 | 67 | 50 | 7 | 15 | 67 | 50 | 7 | 15 | 67 |
| 1992 | 51 | 14 | 27 | 54 | 51 | 14 | 27 | 54 | 51 | 14 | 27 | 54 |
| 1993 | 58 | 17 | 23 | 70 | 58 | 17 | 23 | 70 | 58 | 17 | 23 | 70 |
| 1994 | 64 | 13 | 23 | 57 | 64 | 13 | 23 | 57 | 64 | 13 | 23 | 57 |
| 1995 | 76 | 28 | 33 | 66 | 76 | 28 | 33 | 66 | 76 | 28 | 33 | 66 |
| 1996 | 97 | 22 | 42 | 135 | 97 | 22 | 42 | 135 | 97 | 22 | 42 | 135 |
| 1997 | 76 | 13 | 61 | 144 | 76 | 13 | 61 | 144 | 76 | 13 | 61 | 144 |
| 1998 | 87 | 18 | 58 | 135 | 87 | 18 | 58 | 135 | 87 | 18 | 58 | 135 |
| 1999 | 75 | 13 | 57 | 128 | 75 | 13 | 57 | 128 | 75 | 13 | 57 | 128 |
| 2000 | 72 | 20 | 62 | 105 | 72 | 20 | 62 | 105 | 72 | 20 | 62 | 105 |
| 2001 | 61 | 25 | 50 | 120 | 61 | 25 | 50 | 120 | 61 | 25 | 50 | 120 |
| 2002 | 73 | 19 | 68 | 134 | 73 | 19 | 68 | 134 | 73 | 19 | 68 | 134 |
| 2003 | 70 | 24 | 44 | 143 | 70 | 24 | 44 | 143 | 70 | 24 | 44 | 143 |
| 2004 | 64 | 28 | 60 | 146 | 64 | 28 | 60 | 146 | 64 | 28 | 60 | 146 |
| 2005 | 58 | 32 | 49 | 171 | 58 | 32 | 49 | 171 | 58 | 32 | 49 | 171 |
| 2006 | 41 | 35 | 89 | 177 | 41 | 35 | 89 | 177 | 41 | 35 | 89 | 177 |
| 2007 | 52 | 32 | 88 | 187 | 52 | 32 | 88 | 187 | 52 | 32 | 88 | 187 |
| 2008 | 69 | 18 | 91 | 181 | 69 | 18 | 91 | 181 | 69 | 18 | 91 | 181 |
| 2009 | 62 | 13 | 94 | 181 | 62 | 13 | 94 | 181 | 62 | 13 | 94 | 181 |
| 2010 | 65 | 27 | 98 | 174 | 65 | 27 | 98 | 174 | 65 | 27 | 98 | 174 |
| 2011 | 82 | 33 | 92 | 194 | 82 | 33 | 92 | 194 | 82 | 33 | 92 | 194 |
| 2012 | 46 | 38 | 90 | 192 | 46 | 38 | 90 | 192 | 46 | 38 | 90 | 192 |
| 2013 | 70 | 35 | 89 | 222 | 70 | 35 | 89 | 222 | | | | |
| 2014 | 62 | 71 | 85 | 189 | 62 | 71 | 85 | 189 | | | | |
| 2015 | 65 | 44 | 86 | 193 | 65 | 44 | 86 | 193 | | | | |
| 2016 | 81 | 31 | 84 | 189 | 81 | 31 | 84 | 189 | | | | |
| 2017 | 90 | 31 | 98 | 185 | 90 | 31 | 98 | 185 | | | | |
| 2018 | 55 | 36 | 71 | 225 | | | | | | | | |
| 2019 | 82 | 77 | 90 | 218 | | | | | | | | |
| **Total** | 2365 | 837 | 1978 | 4968 | 2228 | 724 | 1817 | 4525 | 1860 | 512 | 1375 | 3547 |
| | 10148 | | | | 9294 | | | | 7294 | | | |



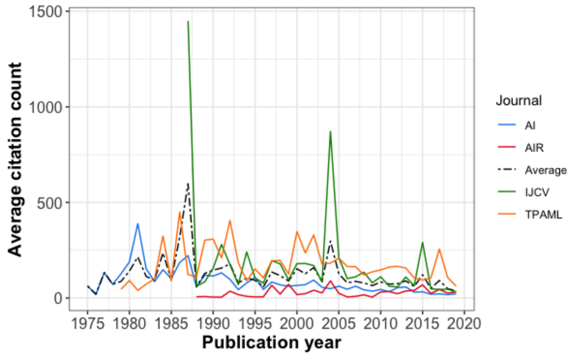

(a) 3-year




(b) 5-year




(c) 10-year




(d) 2022

Figure 16: This figure displays the average citation count per journal for the four AI samples.

Table 14: This table reports summary statistics on the number of (forward) citations of a single publication per journal in the four AI samples. The minimum number of citations per publication is 0 in all cases, thus not reported.

| **Journal** | **Citation summary statistic** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Citation window | | | | | | | | | | | |
| | 3-year | | | 5-year | | | 10-year | | | 2022 | | |
| | Median | Mean | Max. | Median | Mean | Max. | Median | Mean | Max. | Median | Mean | Max. |
| AI | 5 | 9 | 1060 | 9 | 16 | 311 | 17 | 33 | 654 | 25 | 70 | 6560 |
| AIR | 2 | 10 | 346 | 3 | 13 | 658 | 5 | 13 | 464 | 9 | 31 | 1742 |
| IJCV | 9 | 21 | 4339 | 15 | 41 | 10301 | 27 | 81 | 12455 | 32 | 142 | 32117 |
| TPAML | 11 | 30 | 12294 | 18 | 52 | 25752 | 32 | 89 | 5830 | 46 | 167 | 25752 |

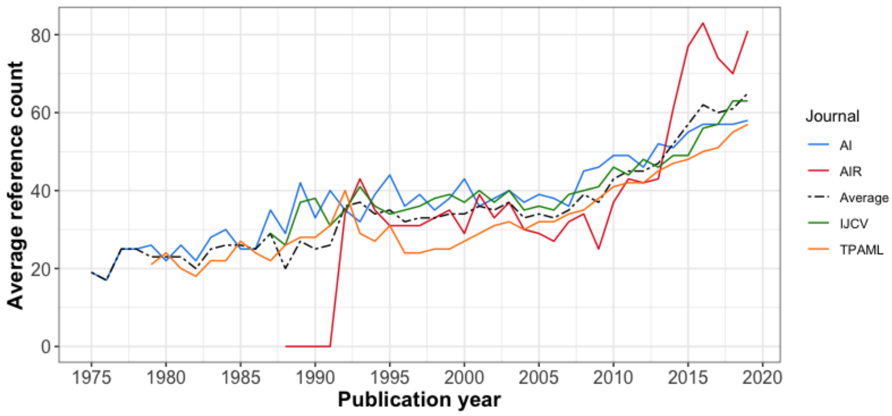


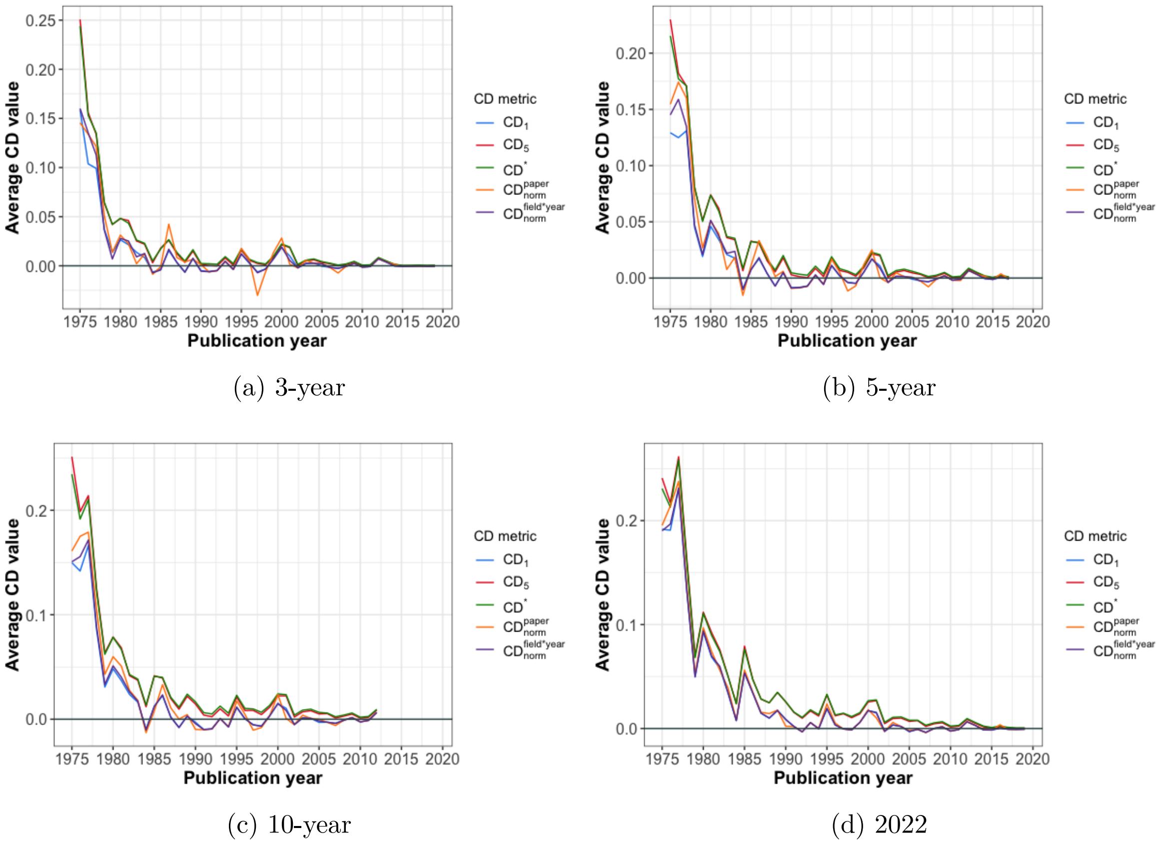

Figure 17: This figure displays the average reference count per journal for the full AI sample.

Table 15: This table reports summary statistics on the number of references of a single publication per journal in the full AI sample.

| **Journal** | **Reference summary statistic** | | | |
|---|---|---|---|---|
| | Min. | Median | Mean | Max. |
| AI | 0 | 37 | 41 | 289 |
| AIR | 0 | 35 | 47 | 724 |
| IJCV | 0 | 41 | 44 | 253 |
| TPAML | 0 | 34 | 37 | 297 |

## 10.3 Results

### 10.3.1 Marketing




(a) 3-year

(b) 5-year

(c) 10-year

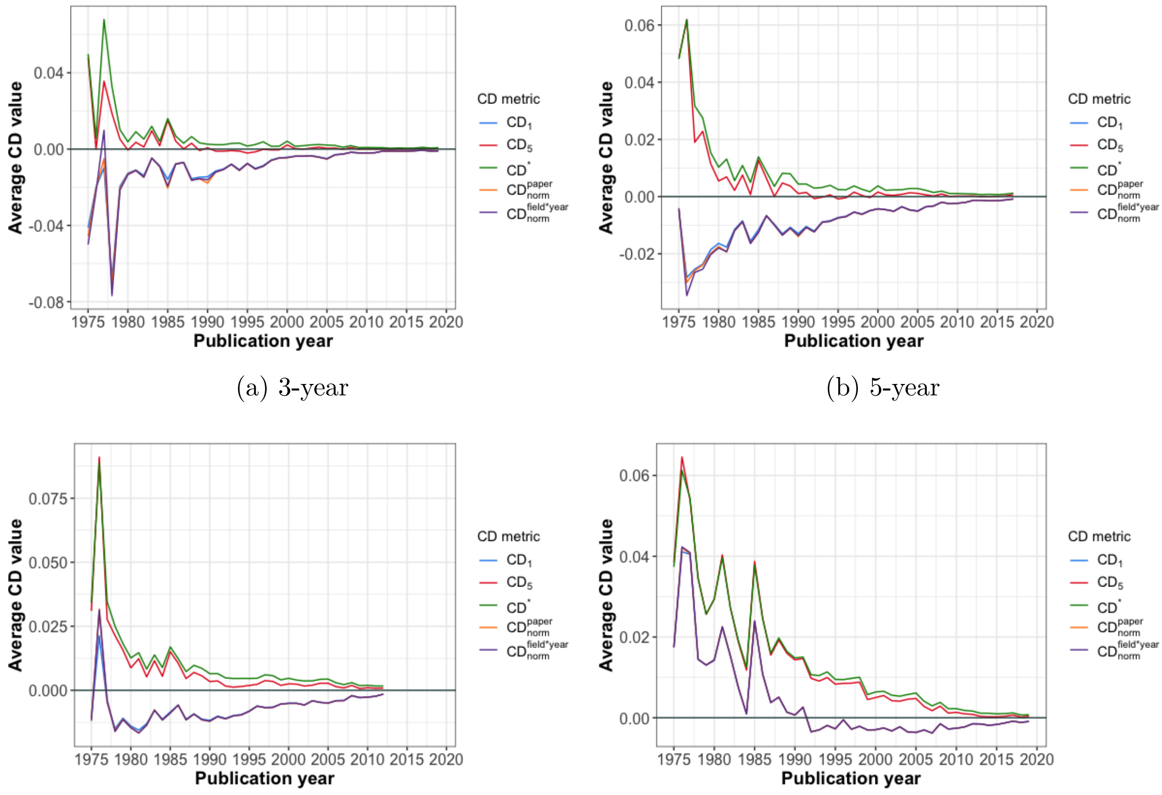
(d) 2022

Figure 18: This figure displays five different disruption indicators' (excluding $CD^{nob}$) values, averaged by publication year for the four Marketing samples.

### 10.3.2 Marketing robustness sample

Table 16: This table reports the number of focal publications included in the robust Marketing samples per year per journal.

| **Pub. year** | **Number of publications** | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Citation window | | | | | | | | | | | | | | | |
| | 3-year | | | | 5-year | | | | 10-year | | | | 2022 | | | |
| | JCR | JM | JMR | MKS | JCR | JM | JMR | MKS | JCR | JM | JMR | MKS | JCR | JM | JMR | MKS |
| 1975 | 2 | 1 | | | 5 | 1 | 2 | | 11 | 1 | 4 | | 15 | 1 | 7 | |
| 1976 | | | 1 | | 2 | | 7 | | 5 | 1 | 10 | | 9 | 5 | 15 | |
| 1977 | 1 | | 2 | | 5 | | 6 | | 9 | | 14 | | 16 | 1 | 18 | |
| 1978 | 1 | | 1 | | 3 | 1 | 4 | | 8 | 2 | 17 | | 9 | 3 | 23 | |
| 1979 | 3 | 1 | 4 | | 8 | 4 | 8 | | 16 | 9 | 11 | | 25 | 12 | 19 | |
| 1980 | 3 | 1 | | | 6 | 5 | 8 | | 11 | 8 | 15 | | 22 | 10 | 25 | |
| 1981 | 1 | 1 | 7 | | 2 | 4 | 9 | | 11 | 12 | 16 | | 19 | 21 | 19 | |
| 1982 | 4 | 2 | 6 | | 9 | 8 | 15 | | 18 | 12 | 23 | | 24 | 18 | 32 | |
| 1983 | 3 | 2 | 1 | | 6 | 9 | 9 | | 21 | 11 | 17 | | 26 | 17 | 27 | |
| 1984 | 4 | 2 | 4 | | 10 | 8 | 11 | | 16 | 18 | 18 | | 26 | 19 | 25 | |
| 1985 | 3 | 3 | 4 | | 12 | 7 | 12 | | 19 | 15 | 22 | | 27 | 24 | 27 | |
| 1986 | 3 | 2 | 6 | | 15 | 4 | 13 | | 23 | 12 | 19 | | 25 | 19 | 22 | |
| 1987 | 3 | 2 | 1 | 1 | 15 | 8 | 7 | 2 | 23 | 18 | 13 | 6 | 29 | 20 | 25 | 12 |
| 1988 | 10 | 3 | 1 | 1 | 18 | 8 | 5 | 3 | 32 | 19 | 16 | 12 | 38 | 22 | 19 | 15 |
| 1989 | 5 | 1 | 4 | 1 | 19 | 11 | 12 | 3 | 25 | 20 | 22 | 10 | 32 | 21 | 27 | 14 |
| 1990 | 6 | 9 | 2 | 1 | 20 | 14 | 11 | 2 | 27 | 17 | 22 | 12 | 32 | 21 | 27 | 16 |
| 1991 | 13 | 3 | 4 | 1 | 23 | 9 | 10 | 6 | 36 | 13 | 23 | 10 | 38 | 18 | 27 | 14 |
| 1992 | 4 | 8 | 9 | 0 | 16 | 11 | 16 | 5 | 37 | 17 | 26 | 13 | 40 | 21 | 30 | 20 |
| 1993 | 7 | 5 | 5 | 4 | 18 | 11 | 12 | 9 | 30 | 20 | 25 | 14 | 36 | 23 | 28 | 19 |
| 1994 | 5 | 15 | 1 | 1 | 21 | 22 | 15 | 2 | 38 | 27 | 29 | 11 | 45 | 30 | 38 | 16 |
| 1995 | 3 | 6 | 6 | 1 | 14 | 12 | 11 | 8 | 24 | 19 | 17 | 23 | 29 | 20 | 26 | 30 |
| 1996 | 1 | 6 | 3 | 2 | 11 | 16 | 15 | 8 | 23 | 24 | 24 | 15 | 25 | 24 | 29 | 17 |
| 1997 | 5 | 4 | 7 | 1 | 14 | 14 | 13 | 11 | 28 | 18 | 27 | 18 | 30 | 20 | 29 | 18 |
| 1998 | 8 | 10 | 8 | 4 | 15 | 20 | 22 | 8 | 22 | 26 | 32 | 19 | 22 | 26 | 32 | 19 |
| 1999 | 5 | 8 | 9 | 4 | 14 | 20 | 20 | 14 | 20 | 33 | 34 | 24 | 21 | 35 | 35 | 28 |
| 2000 | 5 | 9 | 5 | 5 | 18 | 17 | 18 | 8 | 27 | 20 | 33 | 15 | 27 | 20 | 35 | 17 |
| 2001 | 6 | 7 | 10 | 3 | 28 | 13 | 28 | 9 | 38 | 18 | 35 | 17 | 38 | 18 | 37 | 17 |
| 2002 | 8 | 17 | 9 | 5 | 26 | 22 | 18 | 11 | 34 | 23 | 33 | 17 | 35 | 23 | 35 | 18 |
| 2003 | 14 | 21 | 12 | 12 | 28 | 29 | 28 | 19 | 36 | 31 | 32 | 21 | 37 | 31 | 32 | 22 |
| 2004 | 23 | 23 | 12 | 23 | 59 | 32 | 20 | 38 | 70 | 34 | 32 | 39 | 72 | 34 | 33 | 40 |
| 2005 | 18 | 32 | 20 | 25 | 47 | 38 | 33 | 37 | 58 | 40 | 40 | 43 | 60 | 40 | 44 | 43 |
| 2006 | 9 | 27 | 17 | 9 | 30 | 36 | 31 | 25 | 48 | 42 | 42 | 29 | 51 | 42 | 45 | 31 |
| 2007 | 22 | 29 | 28 | 17 | 39 | 37 | 41 | 38 | 57 | 38 | 49 | 48 | 58 | 41 | 50 | 52 |
| 2008 | 29 | 31 | 29 | 13 | 60 | 43 | 39 | 31 | 70 | 48 | 48 | 57 | 72 | 48 | 52 | 61 |
| 2009 | 41 | 40 | 25 | 13 | 65 | 49 | 40 | 37 | 74 | 52 | 55 | 57 | 74 | 53 | 58 | 62 |
| 2010 | 32 | 28 | 39 | 23 | 54 | 37 | 69 | 42 | 67 | 44 | 86 | 58 | 67 | 44 | 88 | 62 |
| 2011 | 39 | 34 | 34 | 18 | 59 | 47 | 72 | 34 | 66 | 51 | 88 | 51 | 67 | 52 | 89 | 53 |
| 2012 | 44 | 33 | 25 | 12 | 72 | 45 | 52 | 33 | 81 | 46 | 70 | 48 | 81 | 46 | 70 | 48 |
| 2013 | 47 | 34 | 22 | 6 | 66 | 45 | 37 | 27 | | | | | 76 | 46 | 48 | 45 |
| 2014 | 52 | 36 | 29 | 15 | 73 | 41 | 41 | 29 | | | | | 84 | 43 | 43 | 43 |
| 2015 | 34 | 26 | 27 | 19 | 49 | 30 | 42 | 30 | | | | | 50 | 36 | 51 | 39 |
| 2016 | 35 | 29 | 29 | 24 | 48 | 32 | 49 | 37 | | | | | 53 | 32 | 56 | 43 |
| 2017 | 47 | 31 | 36 | 23 | 64 | 40 | 53 | 37 | | | | | 64 | 40 | 53 | 37 |
| 2018 | 44 | 30 | 28 | 14 | | | | | | | | | 55 | 40 | 46 | 25 |
| 2019 | 49 | 37 | 33 | 27 | | | | | | | | | 49 | 37 | 33 | 27 |
| **Total** | 701 | 649 | 565 | 328 | 1186 | 860 | 984 | 603 | 1259 | 859 | 1139 | 687 | 1830 | 1217 | 1629 | 1023 |
| | 2243 | | | | 3633 | | | | 3944 | | | | 5699 | | | |




(a) 3-year

(b) 5-year

(c) 10-year

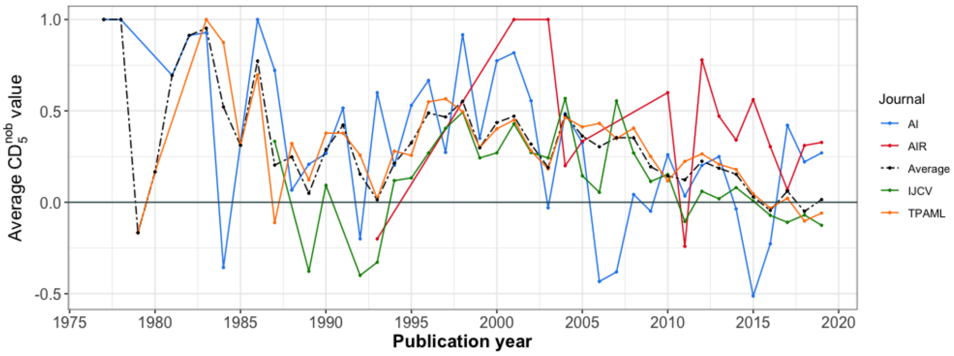
(d) 2022

Figure 19: This figure displays the seven different disruption indicators' values, averaged by publication year for the four Marketing samples.

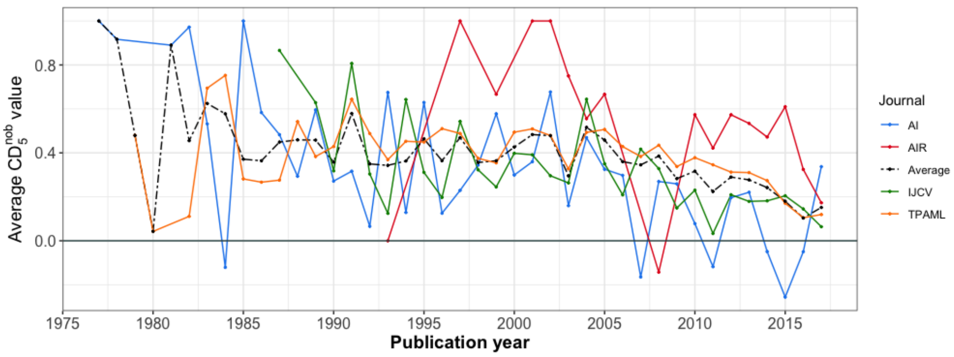
### 10.3.3 AI robustness sample

Table 17: This table reports the number of focal publications included in the robust AI samples per year per journal.

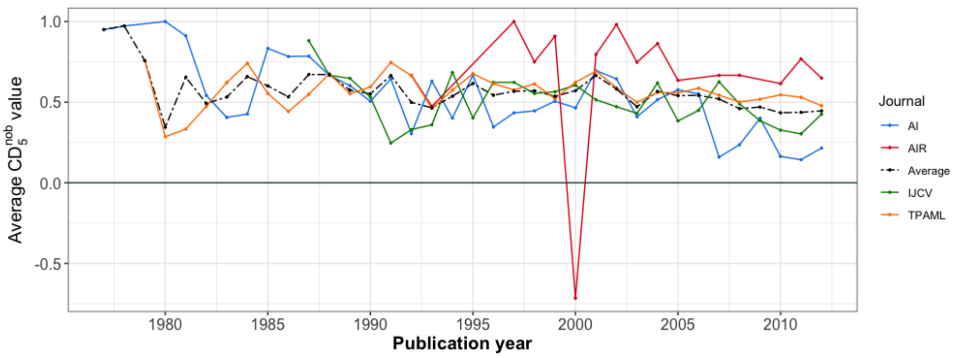
| Pub. year | Number of publications | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Citation window | | | | | | | | | | | | | | | |
| | 3-year | | | | 5-year | | | | 10-year | | | | 2022 | | | |
| | AI | AIR | IJCV | TPAML | AI | AIR | IJCV | TPAML | AI | AIR | IJCV | TPAML | AI | AIR | IJCV | TPAML |
| 1975 | | | | | | | | | | | | | | | | |
| 1976 | | | | | | | | | | | | | | | | |
| 1977 | 1 | | | | 2 | | | | 4 | | | | 4 | | | |
| 1978 | 1 | | | | 2 | | | | 2 | | | | 4 | | | |
| 1979 | | | | 2 | | | | 4 | | | | 6 | | | | 7 |
| 1980 | | | | 2 | | | | 9 | 1 | | | 11 | 1 | | | 16 |
| 1981 | 4 | | | | 4 | | | | 5 | | | 4 | 5 | | | 7 |
| 1982 | 1 | | | | 2 | | | 3 | 3 | | | 8 | 3 | | | 12 |
| 1983 | 2 | | | 1 | 3 | | | 4 | 5 | | | 7 | 5 | | | 11 |
| 1984 | 2 | | | 5 | 2 | | | 8 | 5 | | | 14 | 6 | | | 15 |
| 1985 | | | | 3 | 1 | | | 7 | 2 | | | 10 | 4 | | | 11 |
| 1986 | 2 | | | 6 | 4 | | | 9 | 4 | | | 11 | 5 | | | 14 |
| 1987 | 4 | | 1 | 7 | 7 | | 4 | 11 | 12 | | 4 | 18 | 12 | | 4 | 21 |
| 1988 | 2 | | | 5 | 5 | | | 10 | 9 | | 1 | 15 | 12 | | 2 | 18 |
| 1989 | 3 | | 3 | 11 | 9 | | 5 | 28 | 12 | | 7 | 40 | 14 | | 8 | 42 |
| 1990 | 5 | | 4 | 10 | 13 | | 11 | 22 | 17 | | 14 | 25 | 20 | | 15 | 25 |
| 1991 | 7 | | | 15 | 10 | | 3 | 30 | 18 | | 7 | 41 | 21 | | 7 | 44 |
| 1992 | 1 | | 2 | 14 | 14 | | 7 | 29 | 22 | 1 | 12 | 37 | 24 | 1 | 13 | 42 |
| 1993 | 2 | 1 | 3 | 11 | 4 | 1 | 7 | 21 | 7 | 1 | 10 | 28 | 10 | 2 | 11 | 31 |
| 1994 | 6 | | 3 | 5 | 14 | | 5 | 21 | 20 | | 11 | 29 | 24 | | 12 | 36 |
| 1995 | 10 | | 6 | 13 | 18 | | 16 | 25 | 28 | | 18 | 40 | 41 | | 20 | 46 |
| 1996 | 1 | | 5 | 15 | 12 | | 9 | 31 | 27 | | 18 | 54 | 40 | 2 | 19 | 60 |
| 1997 | 7 | | 7 | 18 | 12 | 1 | 20 | 46 | 27 | 4 | 31 | 68 | 35 | 5 | 39 | 76 |
| 1998 | 4 | | 8 | 17 | 7 | | 17 | 36 | 20 | 1 | 28 | 63 | 31 | 3 | 36 | 76 |
| 1999 | 4 | | 4 | 19 | 9 | 2 | 18 | 41 | 21 | 2 | 29 | 66 | 32 | 4 | 37 | 72 |
| 2000 | 8 | | 12 | 21 | 17 | | 26 | 43 | 31 | 1 | 41 | 62 | 38 | 1 | 44 | 70 |
| 2001 | 2 | 1 | 6 | 30 | 6 | 1 | 13 | 55 | 18 | 4 | 21 | 76 | 21 | 11 | 27 | 81 |
| 2002 | 8 | | 11 | 38 | 13 | 2 | 20 | 64 | 30 | 3 | 35 | 89 | 37 | 4 | 38 | 101 |
| 2003 | 5 | 1 | 11 | 38 | 15 | 1 | 21 | 68 | 32 | 2 | 28 | 94 | 36 | 7 | 30 | 102 |
| 2004 | 5 | 1 | 11 | 48 | 15 | 1 | 19 | 77 | 25 | 1 | 33 | 108 | 33 | 4 | 41 | 117 |
| 2005 | 5 | 1 | 13 | 62 | 18 | 3 | 23 | 89 | 28 | 10 | 26 | 118 | 34 | 11 | 30 | 123 |
| 2006 | 3 | | 24 | 64 | 12 | | 40 | 100 | 22 | | 58 | 120 | 24 | 3 | 61 | 129 |
| 2007 | 8 | | 31 | 70 | 15 | | 45 | 104 | 28 | 4 | 61 | 129 | 29 | 4 | 61 | 135 |
| 2008 | 4 | | 24 | 61 | 15 | 1 | 44 | 97 | 30 | 1 | 57 | 117 | 37 | 4 | 60 | 121 |
| 2009 | 7 | | 26 | 69 | 19 | | 40 | 105 | 34 | | 50 | 135 | 39 | | 56 | 139 |
| 2010 | 8 | 2 | 34 | 77 | 19 | 7 | 54 | 120 | 36 | 10 | 70 | 140 | 38 | 11 | 70 | 141 |
| 2011 | 15 | 3 | 34 | 101 | 27 | 8 | 45 | 135 | 42 | 17 | 59 | 157 | 44 | 17 | 60 | 157 |
| 2012 | 12 | 5 | 43 | 113 | 20 | 11 | 56 | 145 | 27 | 12 | 63 | 162 | 27 | 12 | 63 | 162 |
| 2013 | 26 | 5 | 36 | 135 | 38 | 8 | 53 | 159 | | | | | 55 | 14 | 62 | 182 |
| 2014 | 15 | 16 | 42 | 113 | 24 | 26 | 57 | 140 | | | | | 33 | 37 | 64 | 153 |
| 2015 | 16 | 15 | 40 | 102 | 29 | 29 | 59 | 138 | | | | | 34 | 32 | 64 | 151 |
| 2016 | 24 | 11 | 36 | 120 | 42 | 19 | 50 | 150 | | | | | 47 | 19 | 52 | 154 |
| 2017 | 35 | 17 | 45 | 130 | 55 | 21 | 66 | 152 | | | | | 55 | 21 | 66 | 152 |
| 2018 | 13 | 16 | 42 | 162 | | | | | | | | | 18 | 16 | 49 | 174 |
| 2019 | 18 | 46 | 53 | 156 | | | | | | | | | 18 | 46 | 53 | 156 |
| **Total** | 306 | 141 | 620 | 1889 | 553 | 142 | 853 | 2336 | 654 | 74 | 792 | 2102 | 1050 | 291 | 1274 | 3382 |
| | 2956 | | | | 3884 | | | | 3622 | | | | 5997 | | | |



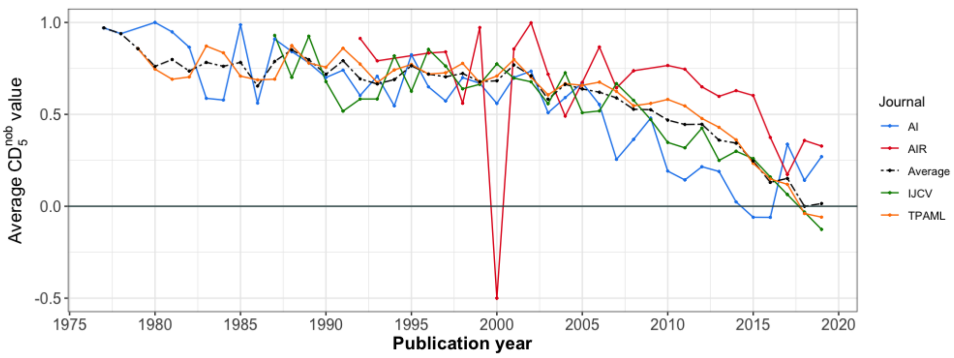

(a) 3-year ($n = 2956$)




(b) 5-year ($n = 3884$)




(c) 10-year ($n = 3622$)




(d) 2022 ($n = 5997$)

Figure 20: This figure displays the average $CD_5^{nob}$ value per journal, for the four robust AI samples.

Table 18: This table reports summary statistics on the $CD_5^{nob}$ indicator per AI journal for the four robust citation window-length samples. The minimum and maximum values are respectively -1 and 1 in all cases, thus not reported.

| **Journal** | $CD_5^{nob}$ **summary statistic** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Citation window | | | | | | | | | | | |
| | 3-year | | | 5-year | | | 10-year | | | 2022 | | |
| | Median | Mean | Std. | Median | Mean | Std. | Median | Mean | Std. | Median | Mean | Std. |
| AI | 0.33 | 0.21 | 0.71 | 0.33 | 0.22 | 0.7 | 0.67 | 0.44 | 0.62 | 0.71 | 0.45 | 0.63 |
| AIR | 0.43 | 0.34 | 0.57 | 0.66 | 0.47 | 0.57 | 0.94 | 0.7 | 0.51 | 0.75 | 0.55 | 0.52 |
| IJCV | 0 | 0.08 | 0.65 | 0.33 | 0.25 | 0.64 | 0.68 | 0.47 | 0.57 | 0.64 | 0.44 | 0.58 |
| TPAML | 0.22 | 0.18 | 0.64 | 0.48 | 0.34 | 0.58 | 0.71 | 0.56 | 0.48 | 0.71 | 0.49 | 0.54 |

Table 19: This table reports the results of the OLS regression models, with $CD_5^{nob}$ as the dependent variable, for the four robust AI samples. The paper-level control variables include the number of references, the number of authors, the article length (in pages), and the number of missing or unlinked references (i.e., the difference between the actual number and the number available in the WoS). The baseline category is the year 1975 in all models.

| **Citation window** | 3-year | | 5-year | | 10-year | | 2022 | |
|---|---|---|---|---|---|---|---|---|
| **Model** | Baseline | Controlled | Baseline | Controlled | Baseline | Controlled | Baseline | Controlled |
| (Intercept) | 1.00 (0.63) | 0.95 (0.59) | **1.00** (0.43)$^{*}$ | **1.07** (0.40)$^{**}$ | **0.95** (0.26)$^{***}$ | **0.96** (0.25)$^{***}$ | **0.97** (0.26)$^{***}$ | **0.88** (0.24)$^{***}$ |
| year=1978 | 0.00 (0.89) | −0.04 (0.83) | −0.08 (0.61) | −0.16 (0.57) | 0.02 (0.46) | 0.14 (0.42) | −0.03 (0.36) | 0.19 (0.34) |
| year=1979 | −1.17 (0.77) | −1.11 (0.72) | −0.52 (0.53) | −0.73 (0.49) | −0.19 (0.34) | −0.13 (0.32) | −0.11 (0.32) | −0.04 (0.30) |
| year=1980 | −0.83 (0.77) | −0.72 (0.72) | **−0.96** (0.47)$^{*}$ | **−1.00** (0.44)$^{*}$ | **−0.61** (0.30)$^{*}$ | −0.45 (0.28) | −0.21 (0.29) | −0.07 (0.27) |
| year=1981 | −0.31 (0.70) | −0.19 (0.66) | −0.11 (0.53) | 0.02 (0.49) | −0.30 (0.32) | −0.10 (0.29) | −0.17 (0.30) | 0.01 (0.28) |
| year=1982 | −0.09 (0.89) | −0.06 (0.83) | −0.54 (0.51) | −0.58 (0.47) | −0.46 (0.31) | −0.35 (0.29) | −0.23 (0.29) | −0.09 (0.27) |
| year=1983 | −0.05 (0.73) | −0.22 (0.68) | −0.38 (0.49) | −0.47 (0.45) | −0.42 (0.30) | −0.25 (0.28) | −0.19 (0.29) | −0.05 (0.27) |
| year=1984 | −0.48 (0.67) | −0.40 (0.63) | −0.42 (0.47) | −0.43 (0.44) | −0.29 (0.29) | −0.12 (0.27) | −0.21 (0.28) | −0.04 (0.26) |
| year=1985 | −0.69 (0.73) | −1.12 (0.68) | −0.63 (0.48) | −0.78 (0.45) | −0.35 (0.30) | −0.18 (0.28) | −0.19 (0.29) | −0.12 (0.27) |
| year=1986 | −0.23 (0.67) | −0.22 (0.63) | −0.64 (0.46) | −0.60 (0.43) | −0.42 (0.30) | −0.23 (0.28) | −0.32 (0.28) | −0.13 (0.27) |
| year=1987 | −0.80 (0.66) | −0.71 (0.61) | −0.55 (0.45) | −0.49 (0.42) | −0.28 (0.28) | −0.10 (0.26) | −0.18 (0.27) | 0.00 (0.26) |
| year=1988 | −0.75 (0.67) | −0.59 (0.63) | −0.54 (0.46) | −0.48 (0.43) | −0.28 (0.28) | −0.09 (0.26) | −0.12 (0.27) | 0.04 (0.26) |
| year=1989 | −0.95 (0.65) | −0.85 (0.61) | −0.54 (0.44) | −0.55 (0.41) | −0.38 (0.27) | −0.21 (0.25) | −0.17 (0.27) | −0.01 (0.25) |
| year=1990 | −0.71 (0.65) | −0.67 (0.60) | −0.64 (0.44) | −0.59 (0.41) | −0.40 (0.27) | −0.19 (0.25) | −0.25 (0.27) | −0.06 (0.25) |
| year=1991 | −0.58 (0.64) | −0.52 (0.60) | −0.42 (0.44) | −0.42 (0.41) | −0.28 (0.27) | −0.11 (0.25) | −0.18 (0.26) | −0.03 (0.25) |
| year=1992 | −0.85 (0.65) | −0.71 (0.61) | −0.65 (0.44) | −0.60 (0.41) | −0.45 (0.27) | −0.26 (0.25) | −0.28 (0.26) | −0.11 (0.25) |
| year=1993 | −0.99 (0.65) | −0.92 (0.61) | −0.66 (0.44) | −0.61 (0.41) | −0.48 (0.27) | −0.26 (0.26) | −0.31 (0.27) | −0.12 (0.25) |
| year=1994 | −0.79 (0.65) | −0.68 (0.61) | −0.64 (0.44) | −0.56 (0.41) | −0.41 (0.27) | −0.20 (0.25) | −0.28 (0.26) | −0.09 (0.25) |
| year=1995 | −0.67 (0.64) | −0.54 (0.60) | −0.54 (0.44) | −0.46 (0.41) | −0.33 (0.27) | −0.11 (0.25) | −0.21 (0.26) | −0.01 (0.25) |
| year=1996 | −0.51 (0.64) | −0.41 (0.60) | −0.64 (0.44) | −0.60 (0.41) | −0.41 (0.27) | −0.22 (0.25) | −0.25 (0.26) | −0.08 (0.25) |
| year=1997 | −0.53 (0.64) | −0.43 (0.60) | −0.53 (0.43) | −0.46 (0.41) | −0.38 (0.27) | −0.16 (0.25) | −0.27 (0.26) | −0.06 (0.25) |
| year=1998 | −0.45 (0.64) | −0.33 (0.60) | −0.64 (0.44) | −0.55 (0.41) | −0.38 (0.27) | −0.13 (0.25) | −0.25 (0.26) | −0.03 (0.25) |
| year=1999 | −0.70 (0.64) | −0.50 (0.60) | −0.64 (0.44) | −0.52 (0.41) | −0.41 (0.27) | −0.14 (0.25) | −0.29 (0.26) | −0.05 (0.25) |
| year=2000 | −0.56 (0.64) | −0.40 (0.60) | −0.57 (0.43) | −0.45 (0.40) | −0.38 (0.27) | −0.08 (0.25) | −0.29 (0.26) | −0.04 (0.25) |
| year=2001 | −0.53 (0.64) | −0.32 (0.60) | −0.52 (0.44) | −0.41 (0.41) | −0.28 (0.27) | −0.03 (0.25) | −0.20 (0.26) | 0.02 (0.25) |
| year=2002 | −0.68 (0.63) | −0.48 (0.59) | −0.52 (0.43) | −0.40 (0.40) | −0.37 (0.27) | −0.07 (0.25) | −0.26 (0.26) | −0.02 (0.25) |
| year=2003 | −0.81 (0.64) | −0.63 (0.59) | −0.70 (0.43) | −0.61 (0.40) | −0.48 (0.27) | −0.21 (0.25) | −0.39 (0.26) | −0.15 (0.25) |
| year=2004 | −0.52 (0.63) | −0.35 (0.59) | −0.48 (0.43) | −0.38 (0.40) | −0.38 (0.27) | −0.10 (0.25) | −0.31 (0.26) | −0.06 (0.25) |
| year=2005 | −0.64 (0.63) | −0.48 (0.59) | −0.54 (0.43) | −0.46 (0.40) | −0.41 (0.27) | −0.16 (0.25) | −0.33 (0.26) | −0.12 (0.25) |
| year=2006 | −0.70 (0.63) | −0.54 (0.59) | −0.64 (0.43) | −0.56 (0.40) | −0.41 (0.27) | −0.18 (0.25) | −0.35 (0.26) | −0.15 (0.24) |
| year=2007 | −0.65 (0.63) | −0.53 (0.59) | −0.65 (0.43) | −0.61 (0.40) | −0.43 (0.27) | −0.22 (0.25) | −0.38 (0.26) | −0.20 (0.24) |
| year=2008 | −0.65 (0.63) | −0.51 (0.59) | −0.61 (0.43) | −0.56 (0.40) | −0.49 (0.27) | −0.26 (0.25) | −0.44 (0.26) | −0.25 (0.24) |
| year=2009 | −0.81 (0.63) | −0.68 (0.59) | −0.72 (0.43) | −0.67 (0.40) | −0.48 (0.27) | −0.26 (0.25) | −0.45 (0.26) | −0.26 (0.24) |
| year=2010 | −0.86 (0.63) | −0.75 (0.59) | −0.68 (0.43) | −0.66 (0.40) | −0.52 (0.27) | −0.29 (0.25) | −0.50 (0.26) | −0.32 (0.24) |
| year=2011 | −0.88 (0.63) | −0.75 (0.59) | −0.78 (0.43) | −0.73 (0.40) | −0.51 (0.27) | −0.27 (0.25) | **−0.53** (0.26)$^{*}$ | −0.34 (0.24) |
| year=2012 | −0.78 (0.63) | −0.64 (0.59) | −0.71 (0.43) | −0.66 (0.40) | −0.50 (0.27) | −0.28 (0.25) | **−0.52** (0.26)$^{*}$ | −0.35 (0.24) |
| year=2013 | −0.81 (0.63) | −0.64 (0.59) | −0.72 (0.43) | −0.63 (0.40) | | | **−0.61** (0.26)$^{*}$ | −0.38 (0.24) |
| year=2014 | −0.85 (0.63) | −0.66 (0.59) | −0.76 (0.43) | −0.64 (0.40) | | | **−0.63** (0.26)$^{*}$ | −0.39 (0.24) |
| year=2015 | −0.97 (0.63) | −0.67 (0.59) | −0.82 (0.43) | −0.60 (0.40) | | | **−0.72** (0.26)$^{**}$ | −0.40 (0.24) |
| year=2016 | −1.04 (0.63) | −0.71 (0.59) | **−0.90** (0.43)$^{*}$ | −0.64 (0.40) | | | **−0.84** (0.26)$^{**}$ | **−0.49** (0.24)$^{*}$ |
| year=2017 | −0.94 (0.63) | −0.60 (0.59) | **−0.85** (0.43)$^{*}$ | −0.59 (0.40) | | | **−0.82** (0.26)$^{**}$ | −0.47 (0.24) |
| year=2018 | −1.05 (0.63) | −0.64 (0.59) | | | | | **−0.97** (0.26)$^{***}$ | **−0.55** (0.25)$^{*}$ |
| year=2019 | −0.99 (0.63) | −0.57 (0.59) | | | | | **−0.96** (0.26)$^{***}$ | **−0.51** (0.25)$^{*}$ |
| NREF | | **−0.01** (0.00)$^{***}$ | | **−0.01** (0.00)$^{***}$ | | **−0.02** (0.00)$^{***}$ | | **−0.01** (0.00)$^{***}$ |
| NREF_Diff | | **0.02** (0.00)$^{***}$ | | **0.03** (0.00)$^{***}$ | | **0.03** (0.00)$^{***}$ | | **0.02** (0.00)$^{***}$ |
| n_Authors | | −0.01 (0.01) | | 0.01 (0.01) | | **0.02** (0.01)$^{***}$ | | 0.01 (0.00) |
| ArticleLength | | **−0.01** (0.00)$^{***}$ | | **−0.01** (0.00)$^{***}$ | | **−0.01** (0.00)$^{***}$ | | **−0.01** (0.00)$^{***}$ |
| $R^2$ | 0.07 | 0.18 | 0.04 | 0.17 | 0.02 | 0.15 | 0.18 | 0.27 |
| Adj. $R^2$ | 0.05 | 0.17 | 0.03 | 0.16 | 0.01 | 0.14 | 0.17 | 0.27 |
| Num. obs. | 2952 | 2933 | 3872 | 3847 | 3613 | 3595 | 5988 | 5954 |

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

## 10.4 Code

The R source files for reproduction are available at https://github.com/482262bz/MasterThesis.

The data input files cannot be made public due to WoS policy.