

Macher, Jeffrey T.; Rutzer, Christian; Weder, Rolf

Working Paper

The illusive slump of disruptive patents

WWZ Working Paper, No. 2023/06

Provided in Cooperation with:

Center of Business and Economics (WWZ), University of Basel

Suggested Citation: Macher, Jeffrey T.; Rutzer, Christian; Weder, Rolf (2023) : The illusive slump of disruptive patents, WWZ Working Paper, No. 2023/06, University of Basel, Center of Business and Economics (WWZ), Basel,
<https://doi.org/10.5451/unibas-ep95069>

This Version is available at:

<https://hdl.handle.net/10419/278568>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



University
of Basel

Faculty of
Business and Economics



June 2023

The Illusive Slump of Disruptive Patents

WWZ Working Paper 2023/06

Jeffrey Macher, Christian Rutzer, Rolf Weder

A publication of the Center of Business and Economics (WWZ), University of Basel.

© WWZ 2023 and the authors. Reproduction for other purposes than the personal use needs the permission of the authors.

Universität Basel
Peter Merian-Weg 6
4052 Basel, Switzerland
wwz.unibas.ch

Corresponding Author:
Dr. Christian Rutzer
Tel.: +41 61 207 33 00
Mail: christian.rutzer@unibas.ch

The Illusive Slump of Disruptive Patents

Jeffrey T. Macher*, Christian Rutzer† and Rolf Weder‡

June 19, 2023

Abstract

Despite tremendous growth in the volume of new scientific and technological knowledge, the popular press has recently raised concerns that disruptive innovative activity is slowing. These dire prognoses were mainly driven by Park et al. (2023), a *Nature* publication that uses decades of data and millions of observations coupled with a novel quantitative metric (the CD index) that characterizes innovation in science and technology as either consolidating or disruptive. We challenge the Park et al. (2023) methodology and findings, principally around concerns of truncation bias and exclusion bias. We show that 88 percent of the decrease in disruptive patents over 1980-2010 reported by the authors can be explained by their truncation of all backward citations before 1976. We also show that this truncation bias varies by technology class. We update the analysis to 2016 and account for a change in U.S. patent law that allows for citations to patent applications in addition to patent grants, which is ignored by the authors in their analysis. We show that the number of highly disruptive patents has increased since 1980—particularly in IT technologies. Our results suggest caution in using the Park et al. (2023) methodology as a basis for research and decision making in public policy, industry restructuring or firm reorganization aimed at altering the current innovation landscape.

Keywords: Disruptive Innovation, Truncation Bias, Exclusion Bias, U.S. Patent Law Change

JEL: O30, O32, O33

1 Introduction

Continued innovation in science and technology is considered a bedrock for and driving force of growth and prosperity in most economies. A paper co-authored by Park, Leahey and Funk entitled, “Papers and Patents are Becoming Less Disruptive Over Time” was recently published in *Nature* (2023). Given its provocative title and findings, as well as the topic examined, the paper attracted significant and global media attention. Park et al. (2023) has been featured in hundreds of international newspapers and magazines (Altmetric, 2023): for example, *The Economist* (2023) emphasized in a report on the changing

*Corresponding author. Georgetown University; jeffrey.macher@georgetown.edu.

†University of Basel; christian.rutzer@unibas.ch.

‡University of Basel; rolf.weder@unibas.ch.

nature of science that "Papers and Patents are Becoming Less Disruptive", indicating in a subtitle that "why that is, is a mystery"; The New York Times (2023) further stated: "What Happened to All of Science's Big Breakthroughs?"; and the Financial Times (2023) noted: "Science is Losing its Ability to Disrupt".

One reason for this attention is that the Park et al. (2023) results show marked declines in disruptive innovation over time. The authors state in the abstract that their "results suggest that slowing rates of disruption may reflect a fundamental shift in the nature of science and technology" (p. 138), and note in the conclusion that "this trend is unlikely to be driven by changes in citation practices or the quality of published work. Rather, the decline represents a substantive shift in science and technology, one that reinforces concerns about slowing innovative activity. We attribute this trend in part to scientists' and inventors' reliance on a narrower set of existing knowledge" (p. 142).

Another reason for this attention is that the Park et al. (2023) results have implications regarding the organization of the entire science and technology innovation process—from government research labs and universities to private and public enterprises. The authors' methodology and findings unsurprisingly piqued the interests of researchers, commentators, and journalists on a global scale who offered myriad explanations, such as increased pressures on scientists to apply for large (interdisciplinary) projects, rising administrative burdens, declining basic research funding, increasing risk-aversion among scientists, and pressures to publish rapidly (Rust, 2023; Yanai and Lercher, 2023). Still other researchers raised concerns around the Park et al. (2023) methodology as a measure of disruptive innovation (Ruan et al., 2021; Bornmann et al., 2020a,b).

An important question arises, however, as to whether the argument made by Park et al. (2023) is accurate. We have major concerns around their methodological approach regarding truncation and exclusion bias related to measurement. We show using patent data that the worrisome result suggested by the authors is instead mainly a consequence of the omission of citations to older innovations: i.e., the authors artificially truncate all backward citations of patents before 1976. We also show using patent data that substantive differences in the results occur from not considering patent law changes: i.e., the authors neglect backward citations of patents published after 2000 to patent applications. Both of these biases have large measurement effects on their proposed CD index and direct implications on the accuracy of their findings and conclusions.

In what follows, we first compare the Park et al. (2023) methodology with truncation to our own methodology with no truncation. We explain why truncation bias can have such a significant impact, and then assess the effects of truncation bias using patent data. We then update the patent data over 2011-2016 and take into account a change in patent law that allows for citations to patent applications as well as to patent grants. We show that the number of highly disruptive patents has increased since 2005. We conclude with a discussion of our main findings and the dangers of measurement bias in influencing public policy, industry organization, and firm organization decisions and interventions.

2 Measuring Patent Disruptiveness

The CD index was developed by Funk and Owen-Smith (2017) and used by Park et al. (2023) to characterize whether a patent or scientific publication is considered more consolidating (i.e., building upon previous research and reinforcing the status quo) or more

disruptive (i.e., obsolescing previous research and pushing into new directions).¹ It is based on the idea that a patent or scientific publication is disruptive if "the subsequent work that cites it is less likely to also cite its predecessors" (Park et al., 2023, p. 139). The CD index ranges from -1 (consolidating) to 1 (disruptive). The measure uses five-year post-publication windows in its construction, referred to as CD_5 . For scientific papers, it starts in 1945; for patents, it starts in 1980.

Fig. 1 plots two average CD_5 indices for approximately 3.66 million patents over 1980-2010 from the United States Patent and Trademark Office (USPTO): one based on the Park et al. (2023) methodology; the other based on our methodology.² As is readily apparent, the evolution of each index differs markedly: the Park et al. (2023) methodology produces an average annual CD index that starts at 0.39 in 1980 (the first observation year) but declines rapidly over time; our methodology instead indicates an average annual CD index that starts at 0.09 in 1980 and declines more gradually over time. Fig. 1 nonetheless shows marked convergence in the two indices in the later years of the sample—particularly since 2000.

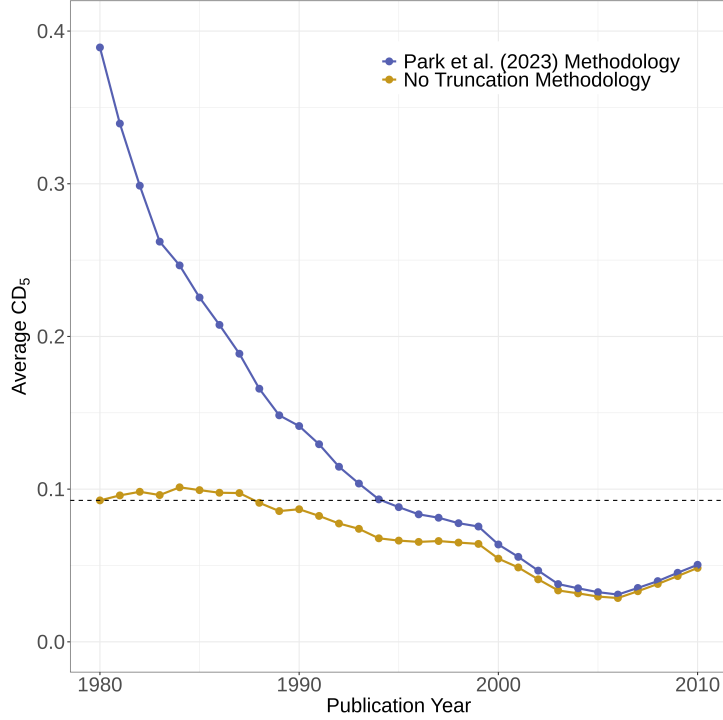
The difference in the average CD indices arises from Park et al. (2023) truncating all backward citations before 1976. To the best of our understanding, the authors do not explain or acknowledge this truncation explicitly despite it having important consequences. The Park et al. (2023) truncated CD index has much larger values compared to our untruncated CD index for those patents published closer to the truncation year. The reason is simple: patents published closer to the truncation year (e.g., 1980) frequently cite patents published in proximate years (e.g., pre-1976), while patents published well-past the truncation (e.g., 2010) infrequently cite those patents. In the earlier years, many backward citations are truncated; in the later years, only a few are. As a result, the CD indices diverge in the early years and converge in the latter years.

This implies that the Park et al. (2023) results are significantly biased due to the truncation of backward patent citations, with the bias strongest for those patents published closer to the truncation year. Hence, their findings of a sharp decline in the average patent disruptiveness can mostly be attributed to a reduction in measurement bias; not to a decline in disruptive innovations. After correcting for truncation bias, our methodology shows that the CD index has a much lower starting point and more modest decline over time. Specifically, 88 percent of the total decline in the average disruptiveness of patents found by Park et al. (2023) can be explained by the authors artificially truncating all backward citations to patents published before 1976.

¹Since its introduction by Funk and Owen-Smith (2017) the index has been used in a wide variety of analyses to capture the disruptive content of patents (e.g. Kaltenberg et al., 2023; Wu et al., 2019) and of scientific papers (e.g. Bornmann et al., 2020a; Wu et al., 2019).

²As in Park et al. (2023), we consider only utility patents and draw our data from PatentsView (version February 21, 2023). The database contains all US patents that have been published between 1976 and September 29, 2022. While Park et al. (2023) focus on patents of the aggregate NBER technology fields "Chemical", "Computer and Communications", "Pharmaceutical and Medical", "Electrical and Electronic", and "Mechanical", which encompasses a total of 3,046,672 granted utility patents over 1980-2010, our main analysis utilizes all granted utility patents published by the USPTO during the same period, resulting in a total of 3,662,051 patents. For more details, see Appendix A.

Figure 1: The CD index: Park et al. (2023) Methodology Versus Our Methodology



The figure shows the average CD_5 Index calculated using the Park et al. (2023) methodology and our methodology. The Park et al. (2023) methodology (in blue) omits backward citations to granted USPTO patents published before 1976; our methodology (in yellow) includes all backward citations to granted USPTO patents.

3 Why Truncation Matters

An example helps illustrate our argument: We compare two patents that both appear in the Park et al. (2023) dataset:³ US 4181011 was issued in 1980; US 6511791 was issued in 2003. US 4181011 makes 13 citations to patents published between 1958 and 1974; US 6511791 makes 11 citations to patents published between 1987 and 2001. Using the Park et al. (2023) methodology and replication data (*patentsview_analytical_df.csv*), the number of backward citations for patent US 4181011 is zero—given the exclusion of backward citations to patents published before 1976. In contrast, the number of backward citations for patent US 6511791 is 11—the same number as listed in the patent document.

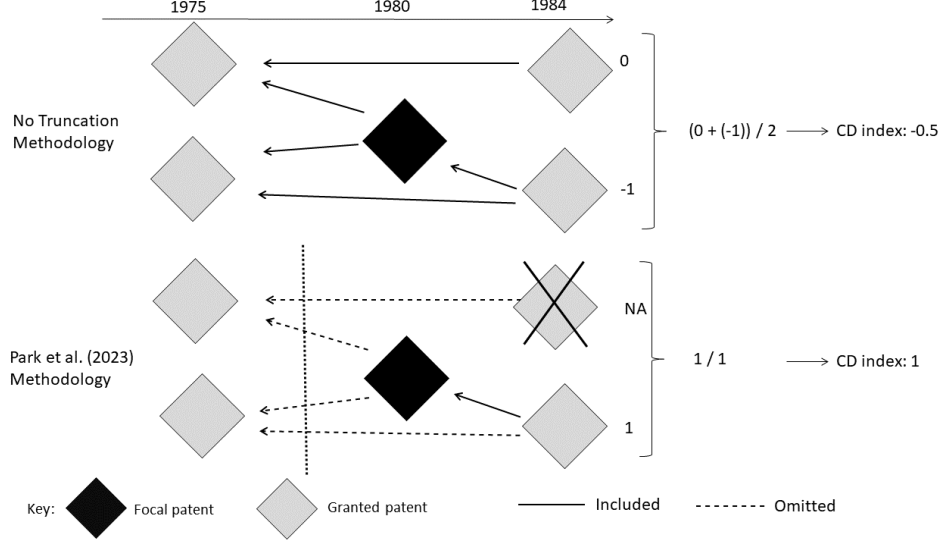
Fig. 2 provides intuition as to how the truncation of backward citations at a given point in time can bias the CD index: without truncation, the focal patent has a CD index value of -0.5 and is considered consolidating; with truncation (i.e., excluding pre-1976 backward citations), the focal patent has a CD index value of 1 and is considered disruptive.

Formally, this can be shown as follows. Park et al. (2023) define the CD index for a focal patent as:

$$CD_t = \frac{1}{N} \sum_{N=i}^N (-2f_{it}b_{it} + f_{it}), \text{ with} \quad (1)$$

³Fig. B.3 in the Appendix shows the first page from each patent document listing the backward citations.

Figure 2: How Truncation of Backward Citations Affects the CD Index



The illustration shows how the CD index of a patent is biased upwards if backward citations are truncated.

$$f_{it} = \begin{cases} 1, & \text{if } i \text{ cites the focal patent within } t \text{ years of post-publication of the focal patent} \\ 0, & \text{otherwise} \end{cases}$$

and

$$b_{it} = \begin{cases} 1, & \text{if } i \text{ cites parts of the predecessors of the focal patent within } t \text{ years of post-publication of the focal patent} \\ 0, & \text{otherwise} \end{cases}$$

where N is the sum of the number of patents citing: the focal patent only (N_F), the focal patent and parts of its predecessors (N_B), and parts of its predecessors only (N_R) within t years after the focal patent publication. A forward citation increases the consolidating (disruptive) nature of the focal patent if the forward cited patent does (does not) cite the predecessor patents (i.e., those patents also cited by the focal patent).

We extend the CD index by directly considering truncation. Truncation affects the CD index in two ways. First, left-truncation (i.e., with respect to the earlier years) reduces the number of counted forward citations N if the truncated backward citations are cited within the t post-publication years of the focal patent:

$$N^{tr} < N^{non-tr}, \quad (2)$$

where tr stands for truncated and $non-tr$ for non-truncated backward citations.

Second, left-truncation reduces the b_{it} value if a forward citation of the truncated predecessor is also a forward citation of the focal patent within t years of publication. In this case:

$$-f_{it}b_{it}^{tr} = 0 \text{ and } -f_{it}b_{it}^{non-tr} = -1 \rightarrow -f_{it}b_{it}^{tr} > -f_{it}b_{it}^{non-tr}. \quad (3)$$

Considering equations (2) and (3), we distinguish among four distinct cases: (i) no backward citations are truncated; (ii) backward citations are truncated but none are cited within the t post-publication years of the focal patent; (iii) truncated backward citations are cited within the t post-publication years of the focal patent but none cite the focal

patent; and (iv) the truncated backward citations are cited within the t post-publication years of the focal patent and at least one of them cite the focal patent. In case (i) and case (ii), equations (2) and (3) hold as equalities and truncation has no effect on the CD index:

$$CD_t^{tr} = CD_t^{non-tr}. \quad (4)$$

In case (iii), the inequality of equation (2) holds but equation (3) is strictly an equality. In case (iv), the inequalities of equations (2) and (3) hold. In case (iii) and case (iv), the truncated CD index is biased upward as long as the summation part of equation (1) of the truncated CD index is positive, because each of the two effects artificially increases its value. Moreover, *ceteris paribus*, the bias is stronger in case (iv). Both results are easily seen by adding the inequalities of equations (2) and (3) to the CD index provided in equation (1). As a result:

$$CD_t^{tr} > CD_t^{non-tr}. \quad (5)$$

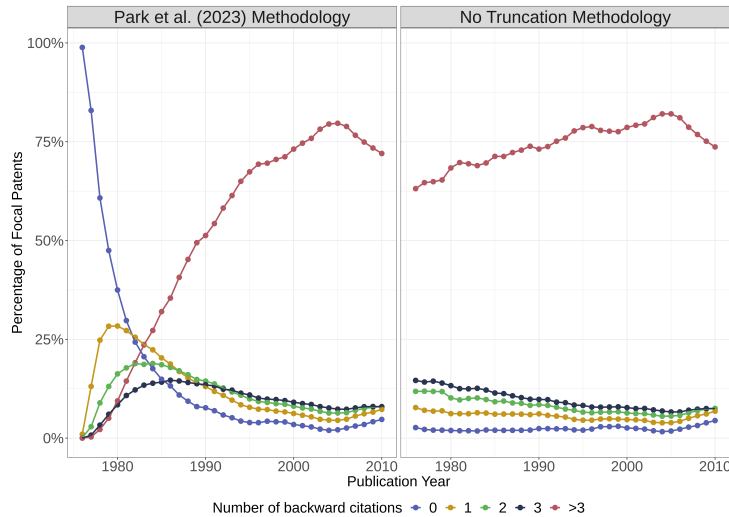
Finally, truncation may lead to an upward or downward biased CD index if the summation of equation (1) is negative and the inequalities of either equation (2) or equations (2) and (3) hold. In this case:

$$CD_t^{tr} \geqslant CD_t^{non-tr}. \quad (6)$$

4 Assessing Truncation Effects

Fig. 3 categorizes focal patents by the number of backward citations to demonstrate the impact of truncation. The left panel uses the Park et al. (2023) methodology; the right panel uses our methodology.

Figure 3: Patent Percentage by Backward Citation Categories

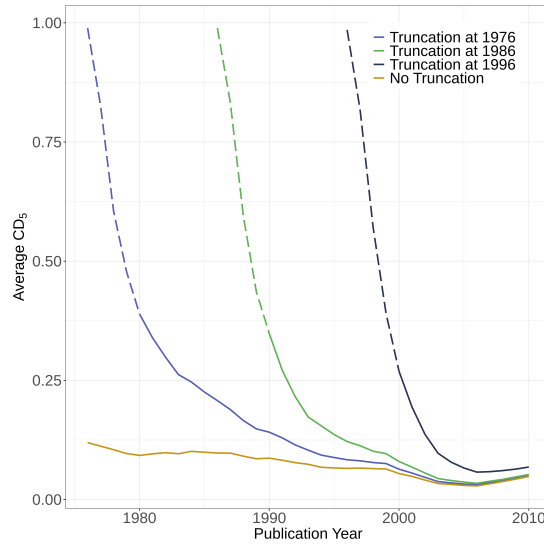


Notes: The figure divides patents into five categories based on the number of backward citations. The vertical-axis shows the percentage of focal patents in each category. The left panel (Park et al. (2023) Methodology) excludes backward citations to patent grants published before 1976; the right panel (No Truncation Methodology) includes backward citations to patent grants regardless of publication year.

Both panels show that the majority of patents have backward citations. However, notable differences are found depending upon whether truncation is or is not present. The

left (truncation) panel indicates that 38 percent of patents had no backward citations in 1980 and 99 percent of patents had no backward citations in 1976 (the truncation year), given that the only backward citations included are those patents also published in 1976. The right (no truncation) panel indicates that less than two percent of patents had no backward citations in these years. Fig. 3 also indicates that patents published further from the truncation year have lower proportions of truncated backward citations. The reason is again straightforward: More recent patents tend to cite more recent patents.

Figure 4: Truncation Time is Not Important

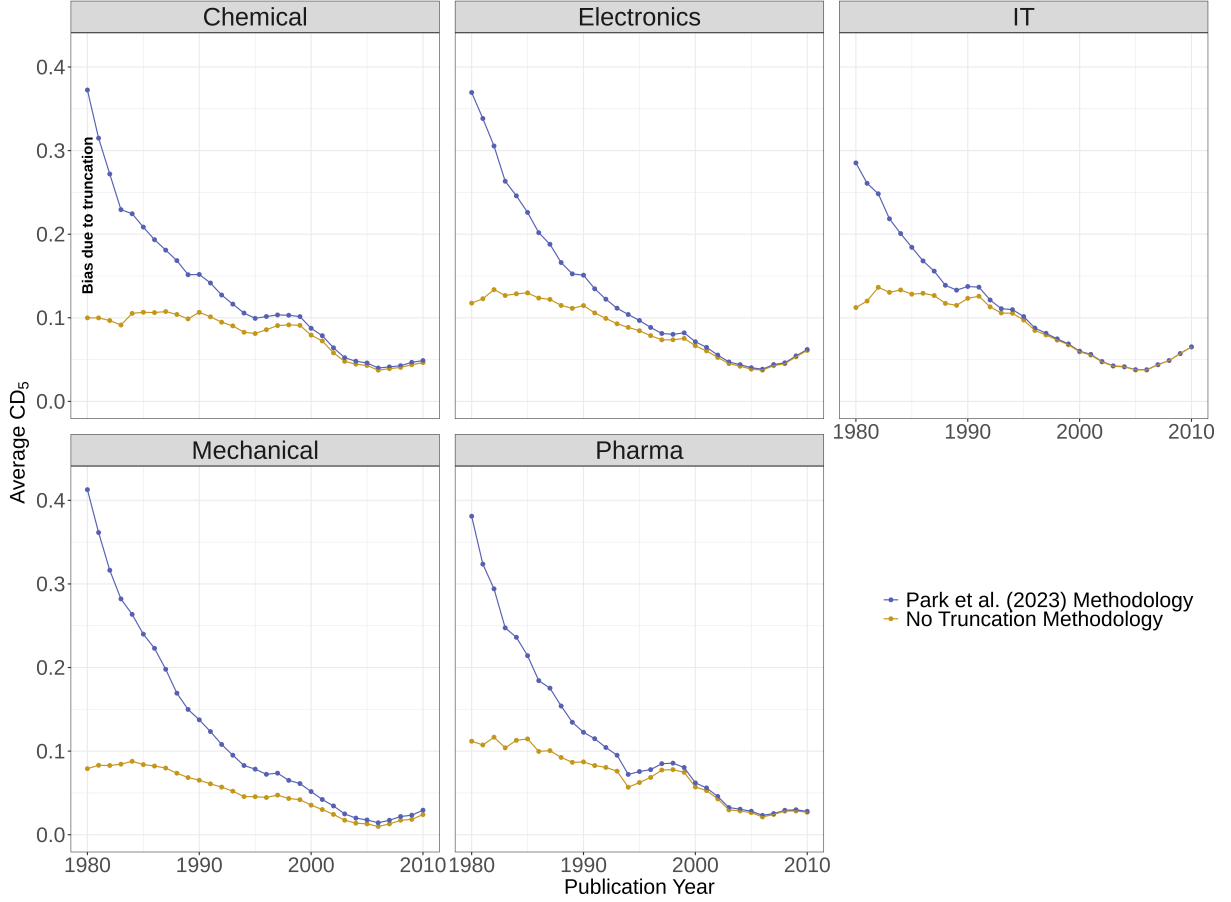


The figure shows average CD_5 indices excluding backward citations to patents published before 1976 (blue line) as in Park et al. (2023), before 1986 (green line), and before 1996 (black line). Park et al. (2023) use a four-year window from the truncation year for the start of their analysis. The dashed lines show how the truncated CD_5 indices change post-truncation year and within this four-year window for illustrative purposes. The solid lines show how the CD_5 indices change post-truncation year and beyond this four-year window (as examined by Park et al., 2023). The average CD_5 index with no truncation is included for comparison purposes.

It is readily apparent that the number of years considered for backward citations matters. The analysis of Park et al. (2023) begins in 1980 and includes backward citations to 1976 (the first year that USPTO patents are available). Hence, patents published in 1980 include four years of backward citations. Fig. 4 shows that had their analysis started in 1976 (i.e., the year of truncation), the bias would be much larger: nearly all patents from 1976 would be classified as disruptive. Fig. 4 also shows that the starting year decision does not alleviate truncation bias and measurement concerns: e.g., had their analysis started in 1990 or in 2000, a four-year window for backward citations would produce similarly sharp declines in the CD index in the earliest years after the truncation year which then stabilizes and converges to the unbiased index over time.

Fig. 5 indicates that this bias is not identical across the technologies examined by Park et al. (2023). The bias is most pronounced in the aggregate technology "Mechanical" and least pronounced in "IT". The difference is again most likely explained by truncation: IT patents have developed more recently and rapidly in comparison to Mechanical patents, suggesting older patents are relatively more often cited in the latter category than in the former category (see Appendix Fig. B.4). As a result, fewer backward citations are truncated in IT patents and more backward citations are truncated in Mechanical patents.

Figure 5: Size of Truncation Bias Differs by Technologies



The figure shows the average CD_5 index for various aggregate WIPO technologies, as determined in Table B.3. The blue line shows the CD index when backward citations to granted U.S. patents published before 1976 are excluded, per the Park et al. (2023) methodology. The yellow line shows the CD_5 index when all backward citations are included, per our methodology.

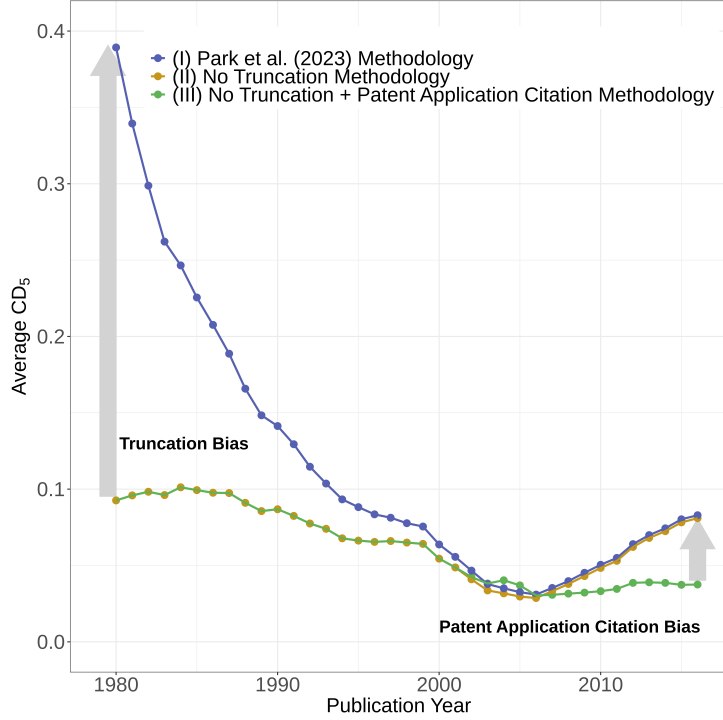
5 Updating Years and Patent Law Change

We examine two factors that could affect the CD index post-2010. First, we update the analysis window using patent data to 2016. Second, we consider a major change to U.S. patent law: in particular, the Inventor Protection Act of 1999 requires that U.S. patent applications be published 18 months after the initial application filing, effectively shifting citation patterns from strictly patent grants to both patent grants and patent applications (Johnson and Popp, 2003).⁴ Fig. 6 shows these post-2010 changes in three average CD_5 indices: the Park et al. (2023) methodology (I); our methodology including all backward citations to strictly patent grants (II); and our methodology including all backward citations to patent grants and patent applications (III). Methodology (III) considers only those citations to patent applications that receive patent grants by the end of 2021. This approach best achieves consistency between the pre- and post-Inventor Protection Act periods, as only patent grants were allowable citations.⁵

⁴See Appendix A for more details on the additional data used, and Appendix B for a more detailed discussion of how excluding citations to patent applications affects the CD index of patents.

⁵Fig. B.5 in the Appendix shows similar results when citations to patent applications that have not (yet) received patent grants are included.

Figure 6: Patent Law Change Creates Exclusion Bias in Park et al. (2023) Methodology



The figure shows the average CD_5 index under different methodologies: the Park et al. (2023) methodology that excludes backward citations to patent grants before 1976 (I); our methodology that includes all backward citations to strictly patent grants (II); and our methodology that includes all backward citations to patent grants and patent applications (III).

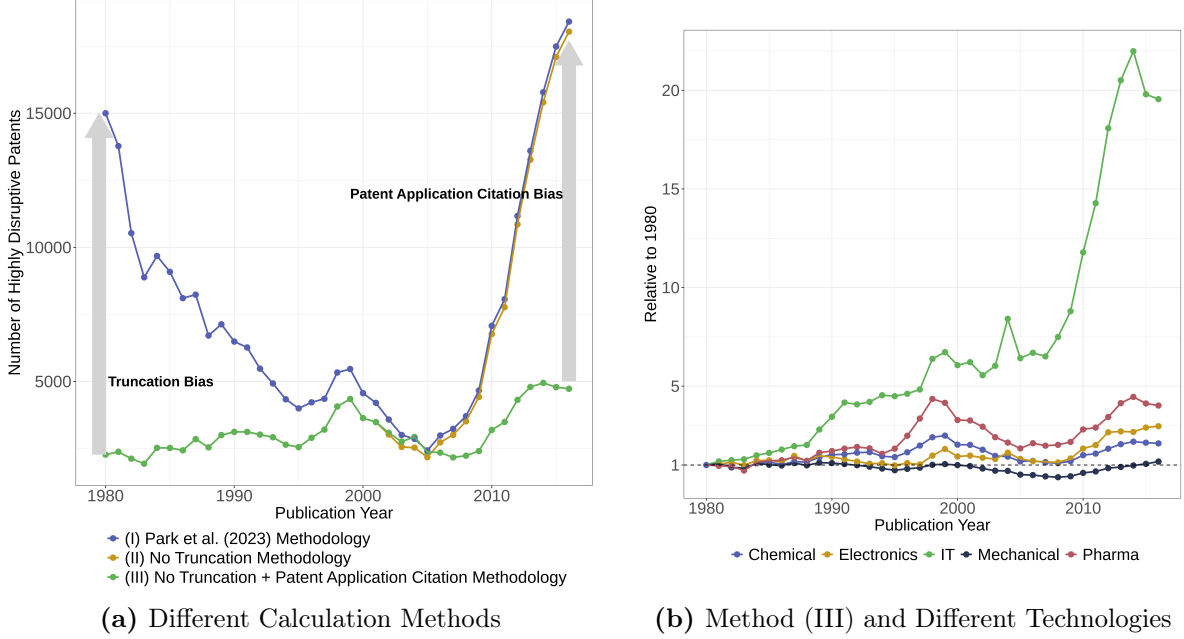
Fig. 6 is identical to Fig. 1 up to 2001 because—prior to the Inventor Protection Act’s passage—patents only cited patent grants. The inclusion of citations before 1976 eliminates truncation bias and subsequently decreases the CD index over 1980-2005—as seen by comparing (I) to (II) and (III). The inclusion of citations to patent applications eliminates exclusion bias and subsequently decreases the CD index over 2005-2016—as seen by comparing (I) and (II) to (III). Correcting for both biases indicates that the average CD index decreases slightly over 1980-2016 and is relatively stable since 2005—as shown in methodology (III).

6 Increasing Number of Highly Disruptive Patents

With both biases considered and accounted for, our analysis documents a notable increase in the number of highly disruptive patents (i.e., patents with $CD_5 \in (0.75, 1]$) over 1980-2016 and particularly since 2010: Panel (a) of Fig. 7 shows that the number of highly disruptive patents has more than doubled since 1980 (Methodology III). It also reveals the major distortions in the CD index if truncation bias (via pre-1976 backward citations) and exclusion bias (via backward citations to patent applications) are not considered (Methodology I and II).

Using 1980 as a benchmark, Panel (b) of Fig. 7 shows that disruption is heterogeneous across technologies: the number of highly-disruptive IT patents increases nearly twenty-fold; the number of highly-disruptive Pharma patents increases more than four-fold; and the number of highly-disruptive Mechanical patents is stable. These results challenge the

Figure 7: Investigation of Highly Disruptive Patents



The figure illustrates the number of and change in highly disruptive patents (i.e., $CD_5 \in (0.75, 1]$) over time. Panel (a) shows the number of highly disruptive patents under different methodologies: the Park et al. (2023) methodology that excludes backward citations to patent grants before 1976 (I); our methodology that includes all backward citations to strictly patent grants (II); and our methodology that includes all backward citations to patent grants and patent applications (III). The total number of highly disruptive patents differs from that reported in Fig. 4b (p. 141) of Park et al. (2023) because they only consider a subset of technologies. (See Appendix A for more detailed discussion.) Panel (b) shows the evolution of highly-disruptive patents across aggregated WIPO technology categories using methodology (III), normalized using 1980 as a baseline.

Park et al. (2023) argument that "despite large increases in scientific productivity, the number of papers and patents with CD_5 values in the far right tail of the distribution remains nearly constant over time" (p. 140). Our findings are, however, consistent with a recent scientific publication that indicates "a paper published in Nature (Park et al., 2023) showed that research is becoming less disruptive. This does not seem to be the case for the next generation of cancer chemotherapy based on the methionine dependence of cancer" (p. 277, Kubota et al., 2023).

Thus far, we have taken an aggregated methodological approach, but this masks two disparities inherent at the individual patent level. First, the truncation of backward citations misclassifies a large proportion of patents as highly disruptive. Appendix Table B.1 illustrates how CD_5 categories for patents published in 1980 calculated using the Park et al. (2023) (truncation) methodology are distributed across CD_5 categories using our (no truncation) methodology. As an example, 25 percent of the USPTO patents in 1980 that are considered highly disruptive under the truncation methodology are actually consolidating under the no truncation methodology.

Second, the exclusion of backward citations to patent applications misclassifies a large proportion of recent patents as highly disruptive. Appendix Table B.2 shows how patents in 2016 that belong to a particular CD_5 category using the Park et al. (2023) (truncation) methodology are distributed across CD_5 categories using our (no truncation) methodology and accounting for backward citations to patent applications. As an example, 19 percent

of the USPTO patents in 2016 that are considered highly disruptive using the Park et al. (2023) methodology are actually consolidating when accounting for both truncation and backward citations to patent applications. Determining the disruptive potential of individual patents thus requires careful consideration to ensure against bias and misclassification. Such bias and misclassification is the case when the data provided by Park et al. (2023) or Funk and Owen-Smith (2017) are used for analysis and decision-making.

7 Discussion

Our examination of the Park et al. (2023) CD index methodology was prompted by the stark decrease in the share of disruptive innovation in science and technology the authors find in patents over 1980-2010 and in scientific publications over 1945-2010. At least for patents, our analysis instead suggests that the decrease in the CD index is mainly due to the truncation of backward citations, while a modest increase in the CD index post-2005 is mainly driven by the exclusion of backward citations to patent applications brought on by a change in patent law.

Variations of the index proposed in Park et al. (2023) and based upon other research (Leydesdorff et al., 2021; Bornmann et al., 2020a) are also affected by truncation bias and exclusion bias when backward citations to granted patents and to patent applications are not appropriately accounted for. Our analysis is important for the research community that may overlook these fundamental issues and work with biased data – either via the CD index or some other measure.

Our analysis highlights, in particular, the dangers of truncation bias from backward citations and exclusion bias from citations to patent applications. Our correction and methodology shows—in comparison to Park et al. (2023)—that: (1) the CD index starting level is significantly lower, modestly declines up to 2005, and is mostly stable since; (2) the starting truncation "date" does not matter—any backward citation "cutoff" in any chosen year can produce similar CD index patterns; (3) truncation bias differences are heterogeneous across technology classes; and (4) the number of highly disruptive patents has increased—particularly since 2010 and in IT.

Our analysis suggests, in general, caution against overt government policy interventions, research policy changes, or substantive industry and firm reorganizations that seek to improve or disrupt the innovation status-quo. If it was accurate that breakthroughs in science and technology have substantially decreased over time, the current organization of the entire science and technology innovation process could be called into question and should be reevaluated. In particular, reforms that seek to improve the creation of disruptive innovation or restart innovation within R&D centers and research labs—e.g., in government agencies, universities and firms—would become paramount.

But our results suggest—at least in the patent data—that the so-called recent slump in disruptive innovation is illusory. An unbiased CD index is instead markedly stable over the past several decades, and highly disruptive patents have increased overall—especially in IT and Pharma. Thus, an interesting question for future research could investigate why these industries have seen *increases* in disruptive innovation, relative to others. Given our findings, we should finally emphasize that premature changes in the current innovation processes of science and technology might bring with them unintended consequences.

8 Credit Authorship Contribution Statement

Jeffrey T. Macher: Conceptualization, Writing – original draft, Writing – review & editing. Christian Rutzer: Calculation, Conceptualization, Investigation, Visualization, Writing – original draft, Writing – review & editing. Rolf Weder: Conceptualization, Writing – original draft, Writing – review & editing.

9 Declaration of Competing Interest

The authors have received no external funding and declare no competing interests.

10 Data Availability

The code and data needed to reproduce this study are available at Zenodo <https://zenodo.org/deposit/8020004>. The raw data used in this research was obtained from PatentsView, a platform that provides unrestricted access to detailed information on all granted USPTO patents published since 1976. All raw data can be downloaded at <https://patentsview.org/download/data-download-tables>.

We used R version 4.1.2 to perform the computations, mainly using the tidyverse and data.table libraries for data processing and ggplot2 for visualization. The computations were performed using SLURM on the sciCORE cluster at the University of Basel, Switzerland.

A Data

The study uses data from the February 21, 2023 version of PatentsView, which includes all U.S. patents published between 1976 and September 29, 2022. As in previous research Park et al. (2023); Funk and Owen-Smith (2017), we focus only on USPTO patents. We include PatentsView data that assigns each USPTO patent to technology categories as defined by the World Intellectual Property Organization (WIPO) Schmoch (2008). The specific WIPO technologies used to create our aggregated technology groups are provided in the Table B.3. It is important to note that our technology classification differs from that of Park et al. (2023), who categorize patents based on NBER technology fields that are no longer available.

In Sections 2-4, we use all granted utility patents over 1980-2010. Park et al. (2023) examine granted utility patents over this timeframe and in the NBER technology fields of "Chemical," "Computers and Communications," "Drugs and Medical," "Electrical and Electronic," and "Mechanical." Their data includes 3,046,672 patents and 29,777,375 backward citations (see replication data file *patentsview_analytical_df.csv* of Park et al. (2023)). We examine granted utility patents over the same timeframe, but cover all technology fields. Our data thus includes 3,662,051 patents and 42,617,016 backward citations in the case of no truncation (i.e., as they appear in the original patent documents) and 35,192,186 backward citations in the case of truncation.

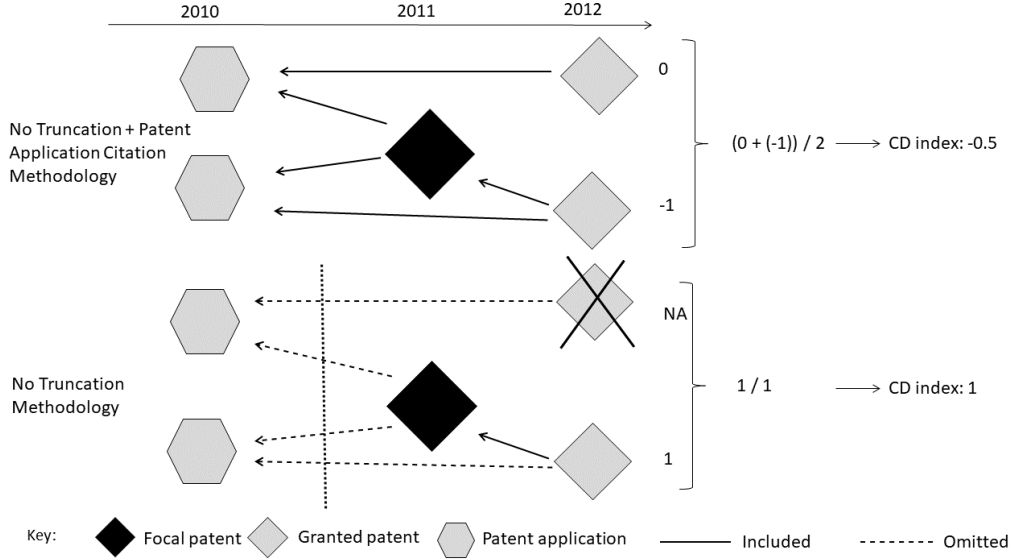
In Sections 5-6, we expand the dataset over 1980-2016, which results in 5,324,224 granted utility patents and 74,943,850 backward citations. With the passage of the Inventor Protection Act of 1999, patents beginning in November 2000 include backward citations to patent applications and to patent grants. Thus, in our methodology (III), we include all backward citations to patent grants and to patent applications that were eventually granted using supplemental PatentsView data. We replace all citations to patent applications with the corresponding patent grant information: i.e., the grant number and publication date. This replacement can result in some backward citations with publication dates later than the citing patent, but the backward citation publication date is not relevant to the CD_5 index methodology. What matters are publication dates of forward citations. We do not include patent applications published within five years after the publication year of a focal patent but with corresponding patent grants published later in the analysis. Finally, we do not include citations to patent applications that have not been granted by the end of 2021 (the latest date available) in the analysis. This approach results in 87,880,452 backward citations.

B CD Index and Patent Law Change

The Inventor Protection Act of 1999 might bias the CD index if citations to patent applications are not considered—especially since patent applications are now increasingly cited *vis-a-vis* patent grants. As Kuhn et al. (2020) note, citations to patent applications accounted for 25 percent of all citations made by USPTO patents in 2015.

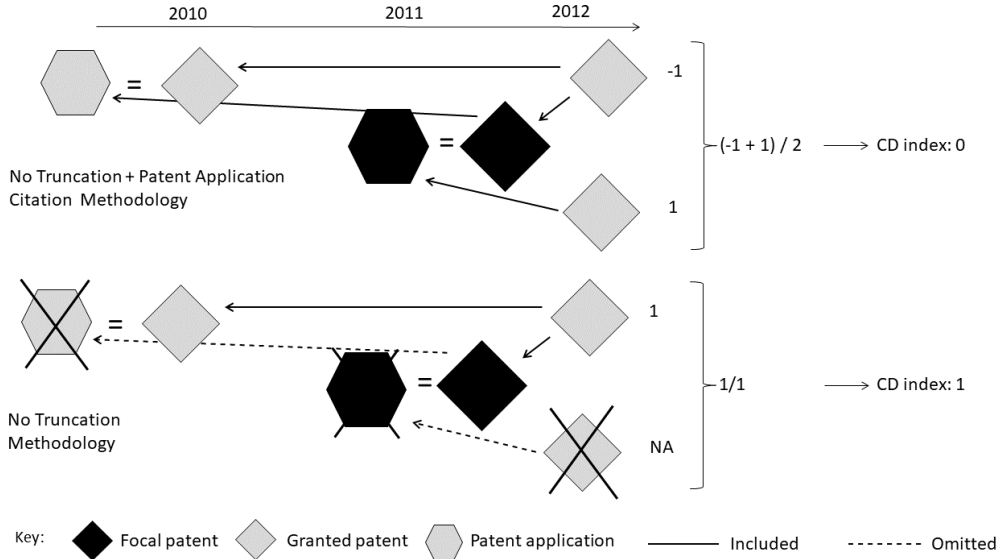
First, as Fig. B.1 illustrates, ignoring patent applications can create bias by falsely declaring a patent as completely disruptive. The channels are the same as when backwards citations are truncated in time. Formally, it affects the CD index via b_{it} and N in equation (1), but the driving force is now missing citations to patent applications. As shown by equations (2)-(6), this can create a bias in the CD index similar to left-truncated backward citations.

Figure B.1: Excluding Backward Citations to Patent Applications Affects the CD Index



The illustration shows how excluding citations to patent applications can bias the CD index of a patent, similar to the truncation of backward citations shown in Figure 2.

Figure B.2: Backward Citations to Patent Application and Corresponding Patent Grant Affects the CD Index



The illustration shows backward citations made to a patent application and the corresponding patent grant can bias the CD index. The patent can be mis-classified as fully disruptive due to the exclusion of citations to the patent application. Specifically, the analysis misses citation to a predecessor application document as well as citation to the application document of the focal patent.

Second, ignoring patent applications can create bias if a patent application is cited in one instance and the subsequent patent grant is cited in another instance. This can occur from citations to predecessor patents or to the focal patent—i.e., applications and grants. Fig. B.2 illustrates this possibility in more detail. Formally, it affects the CD index in equation (1) as before in b_{it} or N but now also in f_{it} . This may result in an upward or

downward bias of the CD index. An upward bias occurs, for example, when the focal patent cites a patent application and a successor patent cites the focal patent and the patent grant (instead of the patent application). In such a case, the citation to the patent application is missed. A downward bias occurs, for example, if a successor patent only cites the patent application of the focal patent and no predecessor patents (applications or grants) of the focal patent. As this would not be considered a forward citation to the focal patent, it thereby loses some of the disruptive value of the focal patent. The Inventor Protection Act of 1999 may therefore bias the CD index if citations to patent applications for patents published after November 29, 2000, are not properly taken into account.

B.1 Figures

Figure B.3: Examples of Backward Citations in Patent Documents

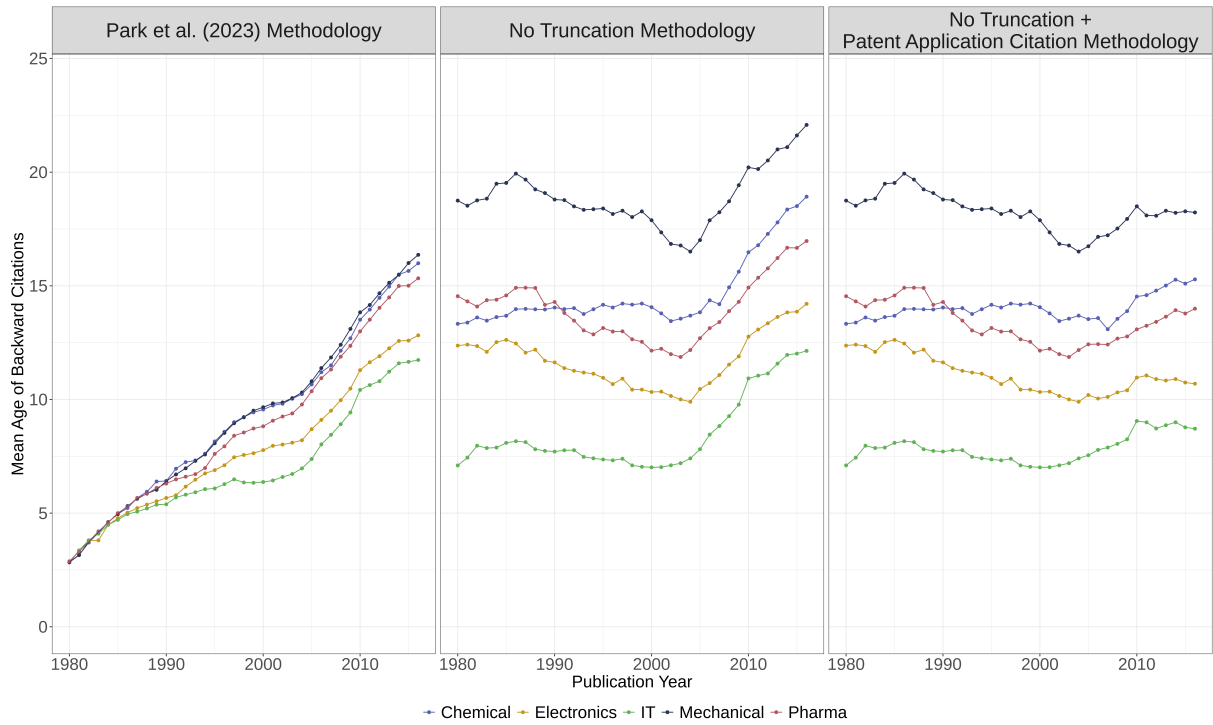
United States Patent [19] 4,181,011 Brendel [45] Jan. 1, 1980		(12) United States Patent Bukofsky et al.		(10) Patent No.: US 6,511,791 B1 (45) Date of Patent: Jan. 28, 2003	
<p>[54] LOAD CELL 3,554,025 1/1971 Andersson et al. .</p> <p>[75] Inventor: Albert E. Brendel, Lake Orion, Mich. 3,576,128 4/1971 Lockery .</p> <p>[73] Assignee: Maatschappij van Berkel's Patent 3,577,779 5/1971 Laimins .</p> <p>N.V., Rotterdam, Netherlands 3,805,604 4/1974 Ormond .</p> <p>[21] Appl. No.: 904,978 3,858,143 12/1974 Schaberg 73/141 A X</p> <p>[22] Filed: May 11, 1978</p> <p>[51] Int. Cl.: C01L 1/22</p> <p>[52] U.S. Cl.: 73/141 A</p> <p>[58] Field of Search 73/141 A, 720, 726, 338/5</p> <p>[56] References Cited</p> <p>U.S. PATENT DOCUMENTS</p> <p>2,866,059 12/1958 Laimins .</p> <p>3,004,231 10/1961 Laimins .</p> <p>3,184,964 5/1965 Hedrick et al. 73/141 A</p> <p>3,341,796 9/1967 Eisele 73/141 A X</p> <p>3,363,456 1/1968 Laimins .</p> <p>3,439,761 4/1969 Laimins .</p> <p>3,486,369 12/1969 Korzilius 73/141 A X</p> <p>3,512,595 5/1970 Laimins .</p> <p>13 Claims, 10 Drawing Figures</p>		<p>(54) MULTIPLE EXPOSURE PROCESS FOR FORMATION OF DENSE RECTANGULAR ARRAYS 5,972,567 * 10/1999 Hu et al. 430/312</p> <p>(75) Inventors: Scott J. Bukofsky, Hopewell Junction, NY (US); Gerhard Kunkel, Radebeul (DE); Richard Wise, New Windsor, NY (US); Alfred K. Wong, Beacon, NY (US) 6,042,998 * 3/2000 Brueck et al. 430/316</p> <p>6,136,517 * 10/2000 Fletcher 430/394</p> <p>6,261,727 * 7/2001 Wong 430/5</p> <p>6,278,123 * 8/2001 Hu et al. 250/492.2</p> <p>* cited by examiner</p> <p>(73) Assignees: International Business machines Corporation, Armonk, NY (US); Infineon Technologies North American Corp., San Jose, CA (US)</p> <p>(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.</p> <p>(21) Appl. No.: 09/561,469</p> <p>(22) Filed: Apr. 28, 2000</p> <p>(51) Int. Cl.: G03F 7/00</p> <p>(52) U.S. Cl.: 430/311; 430/312; 430/394; 430/396</p> <p>(58) Field of Search 430/311, 312, 430/394, 396</p> <p>(56) References Cited</p> <p>U.S. PATENT DOCUMENTS</p> <p>4,702,592 * 10/1987 Geiger et al. 355/53</p> <p>5,308,741 5/1994 Kemp 430/312</p> <p>5,563,012 10/1996 Neisser 430/22</p> <p>5,780,188 * 7/1998 Robson 430/30</p> <p>5,815,245 9/1998 Thomas et al. 355/43</p> <p>5,905,020 * 5/1999 Hu et al. 430/394</p> <p>15 Claims, 8 Drawing Sheets</p>		<p>ABSTRACT</p> <p>A method for exposing a workpiece in a dual exposure step-and-repeat process starts by forming a design for a reticle mask. Deconstruct the design for the reticle mask by removing a set(s) of the features that are juxtaposed to form hollow polygonally-shaped clusters with a gap in the center. Form unexposed resist on the workpiece. Load the workpiece and the reticle mask into the stepper. Expose the workpiece through the reticle mask. Reposition the workpiece by a nanostep. Then expose the workpiece through the reticle mask after the repositioning. Test whether the plural exposure process is finished. If the result of the test is NO the process loops back to repeat some of the above steps. Otherwise the process has been completed. An overlay mark is produced by plural exposures of a single mark. A dead zone is provided surrounding an array region in which printing occurs subsequent to exposure in an original exposure. Alternatively, the workpiece can be fully exposed first by stepping a series of full steps, then going back to the starting position, making a nanostep to reset the starting position and re-exposing from the reset starting position in the same way with full steps from the nanosteped position. The clusters may be in the shape of a hexagon or a diamond.</p>	

(a) First Page of US Patent 4181011

(b) First Page of US Patent 6511791

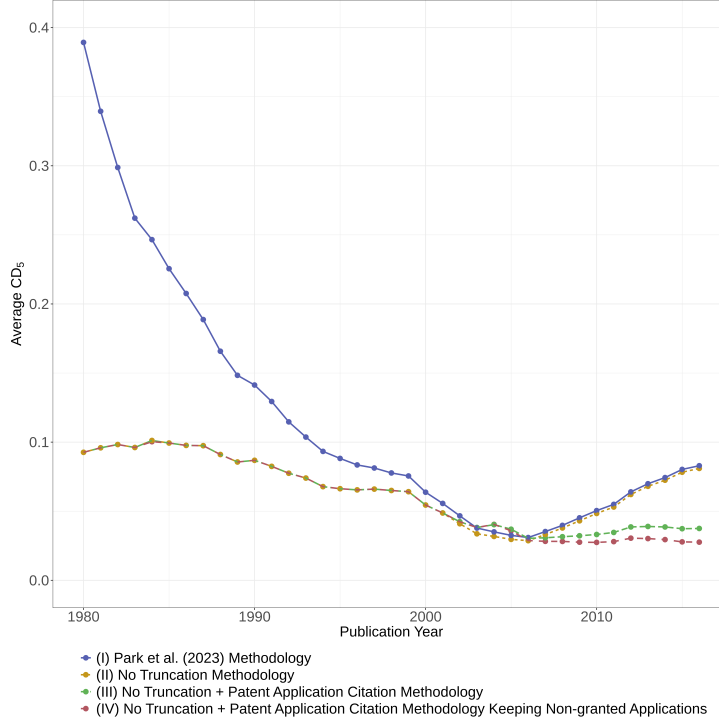
The left panel shows US Patent 4181011 makes 13 backward citations to granted patents, all of which were published before 1976. The right panel shows US Patent 6511791 makes 11 backward citations to other granted patents, all of which were published after 1976.

Figure B.4: Average Backward Citation Age



The figure shows the average age of backward citations for various aggregate WIPO technologies, as determined in Table B.3. The left panel (Truncation Methodology) shows that the average backward citation ages increase in each technology because citations before 1976 are not considered. The middle panel (No Truncation Methodology) shows average backward citation ages are relatively stable in each technology up to 2005 and then increase markedly post-2005 because citations to patent applications are not considered. The right panel (No Truncation + Patent Application Citation Methodology) shows average backward citation ages are relatively stable in each technology and over time.

Figure B.5: Average CD Index By Various Methods



The figure shows the average CD index calculated using four different methodologies. Methodology (I) via Park et al. (2023) excludes backward citations to patent grants before 1976 and backward citations to patent applications. Methodology (II) includes backward citations to patent grants before 1976, but excludes backward citations to patent applications. Methodology (III) includes backward citations to patent grants before 1976 and to patent applications that have received patent grants by the end of 2021. This methodology replaces the application number with the corresponding patent grant number and publication date of the application with the publication date of the patent grant. Finally, methodology (IV) includes backward citations to patent grants before 1976, to patent applications that have received patent grants by the end of 2021, and to patent applications that have not received patent grants by the end of 2021. This methodology replaces the application numbers with their corresponding granted patent numbers while keeping the original application dates, which is necessary to ensure that all citations are taken into account as illustrated by Fig. B.2. Note methodology (IV) lies slightly below methodology (III) after 2005 due to the inclusion of backward citations to patent applications that did not receive patent grants and thus would not appear prior to the Act's passage. Citations to patent applications can increase the number of citations to predecessor patents, which in turn may lead to a decrease in the CD index as the value of N in equation (1) is larger. In this case, however, the decline in the CD index does not necessarily suggest a decline in the number of disruptive innovations. Instead, it could be due to what Petersen et al. (2019) terms "citation inflation".

B.2 Tables

Table B.1: Truncated vs. Non-Truncated Distribution of Patents' CD_5 Index for 1980

No Truncation	Truncation				
	-1, 0	0, 0.25	0.25, 0.5	0.5, 0.75	0.75, 1
-1, 0	1.00	0.18	0.19	0.19	0.25
0, 0.25	0.00	0.82	0.52	0.34	0.33
0.25, 0.5	0.00	0.00	0.29	0.32	0.22
0.5, 0.75	0.00	0.00	0.00	0.16	0.05
0.75, 1	0.00	0.00	0.00	0.00	0.15

The table shows how the CD_5 values of patents of a particular group calculated using the Park et al. (2023) methodology are distributed across the CD_5 values using the correct methodology without truncation. Reading example: The last row of the last column shows that only 15 percent of all patents with a CD_5 index between 0.75 and 1 and calculated by the Park et al. (2023) methodology actually belong to the same group when all backward citations are taken into account.

Table B.2: Truncated vs. Non-Truncated Distribution of Patents' CD_5 Index for 2016

No Truncation and Patent application	Truncation				
	-1, 0	0, 0.25	0.25, 0.5	0.5, 0.75	0.75, 1
-1, 0	0.69	0.26	0.18	0.18	0.19
0, 0.25	0.29	0.71	0.63	0.52	0.56
0.25, 0.5	0.01	0.02	0.15	0.18	0.10
0.5, 0.75	0.00	0.00	0.03	0.10	0.04
0.75, 1	0.00	0.00	0.01	0.02	0.12

The table shows how patents belonging to a particular group of CD_5 values calculated with the Park et al. (2023) methodology are distributed across the CD_5 values using the correct methodology without truncation and considering citations to patent applications. Reading example: The last row of the last column shows that only 12 percent of all patents with a CD_5 index between 0.75 and 1 calculated by the Park et al. (2023) methodology belong to the same group when all backward citations and citations to patent applications are taken into account.

Table B.3: Aggregation of Technologies

WIPO Field ID	WIPO Field Name	Aggregate Technology
1	Electrical machinery, apparatus, energy	Electronics
2	Audio-visual technology	Electronics
3	Telecommunications	Electronics
4	Digital communication	IT
5	Basic communication processes	Electronics
6	Computer technology	IT
7	IT methods for management	IT
8	Semiconductors	IT
9	Optics	Instruments
10	Measurement	Instruments
11	Analysis of biological materials	Chemical
12	Control	Instruments
13	Medical technology	Pharma
14	Organic fine chemistry	Chemical
15	Biotechnology	Chemical
16	Pharmaceuticals	Pharma
17	Macromolecular chemistry, polymers	Chemical
18	Food chemistry	Chemical
19	Basic materials chemistry	Chemical
20	Materials, metallurgy	Chemical
21	Surface technology, coating	Chemical
22	Micro-structural and nano-technology	Chemical
23	Chemical engineering	Chemical
24	Environmental technology	Chemical
25	Handling	Mechanical
26	Machine tools	Mechanical
27	Engines, pumps, turbines	Mechanical
28	Textile and paper machines	Mechanical
29	Other special machines	Mechanical
30	Thermal processes and apparatus	Mechanical
31	Mechanical elements	Mechanical
32	Transport	Mechanical
33	Furniture, games	other
34	Other consumer goods	other
35	Civil engineering	other

The WIPO technologies are defined in Schmoch (2008). The USPTO assigns each patent to one or more of these technologies. The assignment can be found in the data available in PatentsView.

References

- Altmetric, 2023. Overview of attention for article *Papers and patents are becoming less disruptive over time* published in nature, january 2023. <https://nature.altmetric.com/details/140824206/citations>. Accessed: 2023-04-26.
- Bornmann, L., Devarakonda, S., Tekles, A., Chacko, G., 2020a. Are disruption index indicators convergently valid? the comparison of several indicator variants with assessments by peers. *Quantitative Science Studies* 1, 1242–1259.
- Bornmann, L., Devarakonda, S., Tekles, A., Chacko, G., 2020b. Disruptive papers published in scientometrics: meaningful results by using an improved variant of the disruption index originally proposed by wu, wang, and evans (2019). *Scientometrics* 123, 1149–1155.
- Financial Times, 2023. Science is losing its ability to disrupt (2023). <https://www.ft.com/content/c8bfd3da-bf9d-4f9b-ab98-e9677f109e6d>.
- Funk, R.J., Owen-Smith, J., 2017. A dynamic network measure of technological change. *Management science* 63, 791–817.
- Johnson, D.K., Popp, D., 2003. Forced out of the closet: The impact of the american inventors protection act on the timing of patent disclosure. *RAND Journal of Economics* 34, 96–112.
- Kaltenberg, M., Jaffe, A.B., Lachman, M.E., 2023. Invention and the life course: Age differences in patenting. *Research Policy* 52, 104629.
- Kubota, Y., Han, Q., Aoki, Y., Masaki, N., Obara, K., Hamada, K., Hozumi, C., Wong, A.C., Bouvet, M., Tsunoda, T., et al., 2023. Synergy of combining methionine restriction and chemotherapy: The disruptive next generation of cancer treatment. *Cancer Diagnosis & Prognosis* 3, 272.
- Kuhn, J., Younge, K., Marco, A., 2020. Patent citations reexamined. *The RAND Journal of Economics* 51, 109–132.
- Leydesdorff, L., Tekles, A., Bornmann, L., 2021. A proposal to revise the disruption index. *Profesional de la información (EPI)* 30.
- Park, M., Leahey, E., Funk, R.J., 2023. Papers and patents are becoming less disruptive over time. *Nature* 613, 138–144.
- Petersen, A.M., Pan, R.K., Pammolli, F., Fortunato, S., 2019. Methods to account for citation inflation in research evaluation. *Research Policy* 48, 1855–1865.
- 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.
- Rust, R., 2023. Rebooting disruptive science: Exploring the challenges and potential solutions. *European Journal of Clinical Investigation* .

Schmoch, U., 2008. Concept of a technology classification for country comparisons. Technical Report. WIPO.

The Economist, 2023. Papers and patents are becoming less disruptive (2023). <https://www.economist.com/science-and-technology/2023/01/04/papers-and-patents-are-becoming-less-disruptive>.

The New York Times, 2023. What happened to all of science's big breakthroughs? (2023). <https://www.nytimes.com/2023/01/17/science/science-breakthroughs-disruption.html>.

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

Yanai, I., Lercher, M.J., 2023. Make science disruptive again. *Nature Biotechnology* 41, 1–2.