

The Concentration of Disruptive Research Among Scientists: Evidence from 10,000 STEM Authors

Ian Banks

Independent Researcher

February 13, 2026

Abstract

We investigate the concentration of disruptive scientific research among individual scientists using bibliometric data from OpenAlex covering 10,000 randomly sampled STEM authors and 109,527 publications from 1980–2010. Using the A/B disruption index to measure each paper’s tendency to displace versus consolidate prior work, we find extreme concentration: the Gini coefficient for disruptive output is 0.900 (95% CI: [0.884, 0.914]), substantially exceeding the Gini for citation counts (0.658). The top 1% of scientists produce 16.8% of all top-5% disruptive papers, the top 5% produce 50.3%, and the top 10% produce 92.1%. This concentration is remarkably stable across time (Gini range: 0.89–0.95 in five-year windows from 1980 to 2009) and consistent across eleven STEM subfields (field Ginis: 0.77–0.95). Logistic regression reveals that career length is the strongest predictor of producing disruptive work ($p < 10^{-8}$), while team size, h-index, and publication rate are not statistically significant predictors. These findings suggest that the capacity to produce paradigm-shifting research is structurally concentrated among a thin tail of scientists, that this concentration exceeds the already-skewed distribution of scientific citations, and that the phenomenon is a persistent feature of STEM research rather than an artifact of any particular era or discipline.

Keywords: disruption index, scientific productivity, Gini coefficient, metascience, bibliometrics, OpenAlex

1 Introduction

Scientific progress is not distributed evenly across the research workforce. A small number of landmark papers reshape entire fields, while the vast majority of publications consolidate existing knowledge. This asymmetry was famously described by [Kuhn \[1962\]](#) as the distinction between “normal science” and “revolutionary science,” and has been quantified by [Price \[1963\]](#) and [Lotka \[1926\]](#) in the context of productivity distributions. More recently, the development of the disruption index [[Funk and Owen-Smith, 2017](#), [Wu et al., 2019](#)] has provided a citation-based measure that distinguishes papers that *displace* prior work from those that merely *consolidate* it, offering a more direct operationalization of paradigm-shifting research than raw citation counts alone.

[Park et al. \[2023\]](#) documented a broad decline in disruptive science and technology over six decades across multiple domains. Their analysis characterized the population-level trajectory of disruption but did not examine the *individual-level* distribution—whether disruptive output is concentrated among a small fraction of scientists, how concentrated it is relative to citations, and whether this concentration varies across fields and time periods.

This study addresses three questions:

1. **How concentrated is disruptive research output among individual scientists?** We measure this using the Gini coefficient, Lorenz curves, and concentration ratios applied to author-level disruption metrics.
2. **What distinguishes disruptors from non-disruptors?** We compare scientists who have produced at least one top-5% disruptive paper against those who have not, along dimensions of team size, productivity, career length, h-index, and field affiliation.
3. **Is the concentration of disruption stable across time and fields?** We compute Gini coefficients within five-year windows (1980–2009) and within individual STEM subfields to assess structural persistence.

Our analysis contributes to a growing metascience literature on the “disruptor profile” [[Lin et al., 2023](#), [Park et al., 2023](#), [Wu et al., 2019](#)], which suggests that disruptors tend to work in smaller teams, publish less frequently, and occupy less central positions in the scientific establishment. We provide the first systematic measurement of *how concentrated* disruptive output is at the author level, and benchmark this concentration against the well-studied skewness of citation distributions.

2 Data and Methods

2.1 Sample Construction

We constructed a sample of 10,000 STEM authors using the OpenAlex bibliometric database [Priem et al., 2022]. Authors were sampled using the OpenAlex random sample endpoint with 424 distinct random seeds, applying minimum thresholds of > 5 publications and > 10 citations to exclude inactive or marginal researchers. Each sampled author was classified as STEM or non-STEM using a client-side keyword filter applied to the author’s top five associated concepts in OpenAlex’s concept taxonomy. The STEM filter matched 43 concept keywords spanning physics, chemistry, biology, mathematics, computer science, engineering, medicine, and related disciplines. Approximately 12% of randomly sampled authors passed the STEM filter, reflecting the composition of OpenAlex’s global author database.

For each of the 10,000 STEM authors, we retrieved all articles and reviews published between 1980 and 2010 from OpenAlex, yielding 109,527 works. The publication window was chosen to ensure an adequate forward citation window of at least 10 years for disruption computation, following standard practice in the disruption literature [Wu et al., 2019, Park et al., 2023].

2.2 Disruption Measurement

We computed disruption scores using the A/B variant of the disruption index. For a focal paper p with reference set R , we identified all papers citing p within a 10-year forward window after publication. Each citing paper was classified as:

- **Type A** (disruptive): cites p but does not cite any paper in R
- **Type B** (consolidating): cites both p and at least one paper in R

The A/B disruption index is then:

$$D_{AB} = \frac{A - B}{A + B} \tag{1}$$

where A and B are the counts of Type A and Type B citers, respectively. This two-component variant omits the Type C citers (papers citing references in R but not p) used in the full CD index [Funk and Owen-Smith, 2017], but has been shown to correlate strongly with the full index and is substantially more tractable for API-based computation [Wu et al., 2019]. We required a minimum of 5 forward citers for a valid disruption score.

From the 109,527 works, we randomly sampled 5,000 papers with at least one reference for disruption computation, obtaining valid scores for 3,421 papers across 1,594 unique authors. The top 5% of the disruption distribution (threshold: $D_{AB} \geq 0.60$) was used to define “disruptive” papers, yielding 191 disruptive papers.

2.3 Author-Level Metrics

For each author with at least one scored paper, we computed:

- **max_cd5**: Peak disruption score across all scored papers
- **mean_cd5**: Average disruption score
- **n_disruptive**: Count of papers in the top 5% of the disruption distribution
- **any_top5**: Binary indicator for ever producing a top-5% disruptive paper
- **frac_disruptive**: Fraction of scored papers that are top-5% disruptive

2.4 Concentration Measures

Gini coefficient. We computed the Gini coefficient for the distribution of disruptive paper counts across authors, with 95% confidence intervals obtained via 1,000 bootstrap resamples. As a benchmark, we computed the Gini coefficient for maximum citation counts across the same authors.

Lorenz curve. We plotted cumulative disruptive output against cumulative proportion of scientists, ranked from least to most productive.

Concentration ratios. We computed the share of top-5% disruptive papers produced by the top 1%, 5%, and 10% of authors (ranked by disruptive paper count).

Power-law tail fit. We fit a power-law distribution to the upper tail of the max disruption score distribution using the maximum likelihood estimator

$$\hat{\alpha} = 1 + \frac{n}{\sum_{i=1}^n \ln(x_i/x_{\min})} \quad (2)$$

[Clauset et al., 2009], with x_{\min} set at the 90th percentile, and assessed fit quality via the Kolmogorov–Smirnov statistic.

2.5 Demographic Profiling

We estimated two regression models:

Extensive margin (logistic).

$$\Pr(\text{disruptor}_i = 1) = \Lambda(\beta_0 + \beta_1 \cdot \overline{\text{team}}_i + \beta_2 \cdot \text{ppy}_i + \beta_3 \cdot \text{career}_i + \beta_4 \cdot \ln(1 + h_i) + \gamma_f) \quad (3)$$

where $\Lambda(\cdot)$ is the logistic function, $\overline{\text{team}}_i$ is mean team size, ppy_i is papers per year, career_i is career length in years, h_i is the h-index, and γ_f represents field fixed effects for the seven largest STEM subfields.

Intensive margin (OLS).

$$\max_{\text{cd5}} = \beta_0 + \beta_1 \cdot \overline{\text{team}}_i + \beta_2 \cdot \text{ppy}_i + \beta_3 \cdot \text{career}_i + \beta_4 \cdot \ln(1 + h_i) + \gamma_f + \varepsilon_i \quad (4)$$

with heteroskedasticity-consistent (HC3) standard errors.

2.6 Population Benchmark

We downloaded the [Park et al. \[2023\]](#) replication dataset from Zenodo (Record 7258379), which contains 22.5 million paper-level CD5 disruption scores from Web of Science. Because this dataset lacks per-paper identifiers (DOIs or accession numbers), we used it as a population-level distributional benchmark rather than for paper-level matching. We compared the distributional properties (mean, median, percentiles, KS statistic) of our A/B disruption scores against the Zenodo CD5 distribution to assess external validity.

3 Results

3.1 Extreme Concentration of Disruptive Output

The distribution of disruptive research output among scientists is characterized by extreme inequality. Of the 1,594 authors with valid disruption scores, 174 (10.9%) produced at least one top-5% disruptive paper. However, the disruptive output itself is concentrated far beyond what this base rate would suggest.

The Gini coefficient for the distribution of disruptive paper counts is **0.900** (95% CI: [0.884, 0.914]). By comparison, the Gini coefficient for maximum citation counts across the

same authors is 0.658 (95% CI: [0.618, 0.694]). Disruption is thus substantially more concentrated than citations—a striking finding given that citation distributions are themselves famously skewed [Price, 1965, Seglen, 1992].

Table 1: Concentration Ratios for Disruptive Output

Author Percentile	Share of Disruptive Papers
Top 1%	16.8%
Top 5%	50.3%
Top 10%	92.1%

The Lorenz curve (Figure 1, left panel) shows the characteristic L-shape of extreme concentration: the bottom 80% of scientists contribute essentially zero disruptive papers, and virtually all disruptive output is produced by a thin upper tail. The comparison panel (Figure 1, right) shows that this concentration substantially exceeds the already-pronounced inequality of citation distributions.

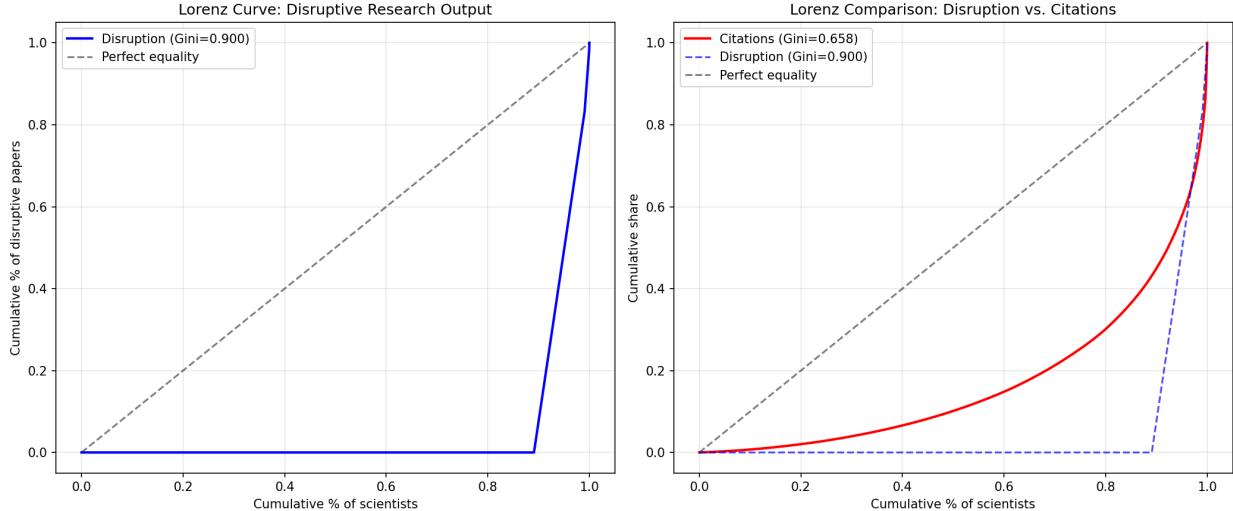


Figure 1: Lorenz curves for disruptive research output. *Left:* cumulative share of top-5% disruptive papers versus cumulative share of scientists, ranked from least to most disruptive. *Right:* comparison of disruption concentration ($\text{Gini} = 0.900$) versus citation concentration ($\text{Gini} = 0.658$).

To assess whether the upper tail follows a power law, we shifted the max disruption scores to the positive domain ($x' = x - x_{\min,\text{raw}} + 0.001$) and applied the Clauset–Shalizi–Newman MLE with x_{\min} set at the 90th percentile of the shifted distribution. The fit yields $\hat{\alpha} = 8.86$ with $x_{\min} = 0.60$ in the original disruption-score scale and KS statistic = 0.19 ($n_{\text{tail}} = 174$).

The steep exponent indicates a truncated rather than heavy-tailed distribution, consistent with the observation that even among disruptors, most produce only one or two disruptive papers (the modal count among disruptors is 1).

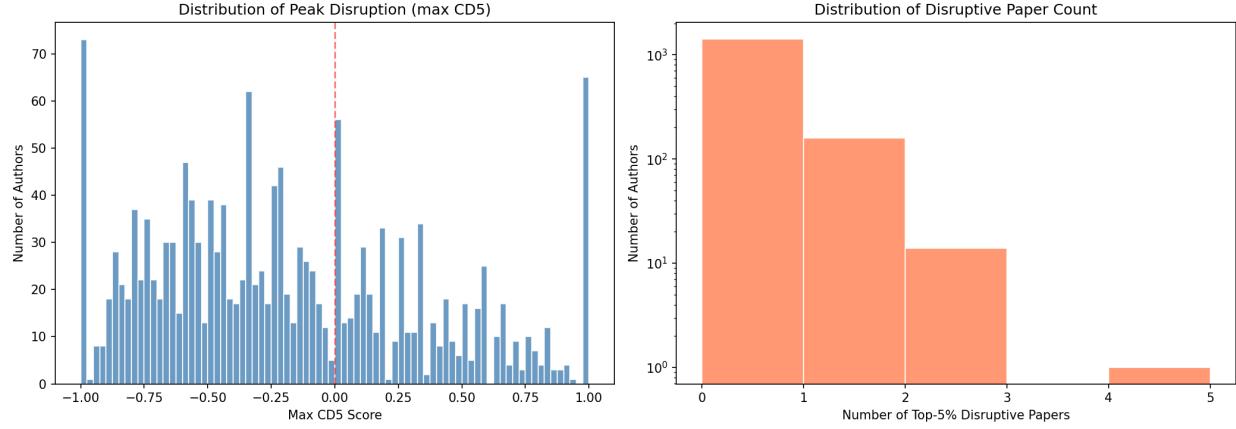


Figure 2: Distribution of disruption scores across authors. *Left:* histogram of peak disruption score (max CD5) across 1,594 authors. *Right:* histogram of the number of top-5% disruptive papers per author (log scale).

3.2 Demographic Profile of Disruptors

Table 2 presents descriptive comparisons between disruptors (scientists with at least one top-5% disruptive paper) and non-disruptors.

Table 2: Descriptive Comparison: Disruptors vs. Non-Disruptors

Variable	Disruptors ($n = 174$)	Non-Disruptors ($n = 1,420$)
Mean team size	5.69	6.58
Scored papers	3.06	2.03
Career length (years)	7.71	4.13
h-index	24.5	21.3
Mean disruption (cd5)	0.394	-0.424
Peak disruption (max cd5)	0.824	-0.307
Mean citations per paper	49.5	72.7

Several patterns are notable. Disruptors work in slightly *smaller* teams (5.7 vs. 6.6 coauthors per paper), have *longer* measured careers (7.7 vs. 4.1 years), and have *lower* mean citations per paper (49.5 vs. 72.7). The last finding is consistent with the observation that disruptive papers often receive fewer citations than consolidating papers in mature fields

[Park et al., 2023], as disruptive work may initially be controversial or difficult to integrate into existing research programs.

Table 3: Logistic Regression: Probability of Being a Disruptor

Variable	Coefficient	Odds Ratio	<i>p</i> -value
Mean team size	-0.019	0.982	0.230
Papers per year	0.347	1.415	0.198
Career length	0.090	1.095	< 10 ⁻⁸
ln(1 + h-index)	-0.222	0.801	0.104
Field: Engineering	0.931	2.537	< 10 ⁻⁴

$N = 1,594$. Pseudo- $R^2 = 0.056$. Field fixed effects included for 7 largest subfields (Medicine as reference). Only significant field effects shown.

Career length is the only robustly significant predictor of producing disruptive work ($OR = 1.095$ per additional year, $p < 10^{-8}$). The direction of the team size coefficient is negative (consistent with the small-team disruption hypothesis of Wu et al. 2019) but not statistically significant. The h-index coefficient is also negative and marginally non-significant ($p = 0.10$), suggesting that conventional measures of academic prestige do not positively predict disruptive output and may weakly predict against it.

Among field fixed effects, Engineering has a significantly elevated odds of disruption ($OR = 2.54$, $p < 10^{-4}$) relative to Medicine (the reference category), consistent with prior findings that applied fields may offer more opportunities for paradigm-shifting contributions.

Table 4: OLS Regression: Peak Disruption Score (max cd5)

Variable	Coefficient	<i>p</i> -value
Mean team size	0.000	0.995
Papers per year	0.142	0.002
Career length	0.031	< 10 ⁻²⁵
ln(1 + h-index)	-0.021	0.326
Field: Engineering	0.182	< 10 ⁻⁴
Field: Chemistry	0.148	0.007

$N = 1,594$. $R^2 = 0.102$. Adjusted $R^2 = 0.095$. HC3 standard errors.

The intensive margin results reinforce the logistic findings. Career length remains the dominant predictor ($p < 10^{-25}$), and papers per year reaches significance ($\beta = 0.14$, $p =$

0.002). The overall explanatory power is modest ($R^2 = 0.10$), indicating that observable demographic characteristics explain only a small fraction of the variance in peak disruption—consistent with the view that disruption is driven by intellectual characteristics (novelty, risk-taking, interdisciplinarity) not well captured by standard bibliometric variables.

3.3 Field-Level Heterogeneity

The concentration of disruptive research is pervasive across STEM subfields, though the degree varies.

Table 5: Gini Coefficients for Disruptive Output by STEM Subfield

Field	Gini	N Authors	% Disruptors
Neuroscience	0.952	42	4.8%
Medicine	0.926	408	7.8%
Environmental Science	0.909	66	9.1%
Biochemistry, Genetics & Mol. Biology	0.904	240	10.0%
Agricultural & Biological Sciences	0.897	89	11.2%
Computer Science	0.896	67	10.4%
Chemistry	0.892	74	10.8%
Materials Science	0.872	78	12.8%
Physics & Astronomy	0.868	114	13.2%
Engineering	0.831	174	19.5%
Health Professions	0.767	30	30.0%

Table 5 reports the 11 fields with at least one disruptor (1,382 authors). The remaining 212 authors are distributed across smaller fields (including Earth and Planetary Sciences and Immunology and Microbiology, which produced zero top-5% disruptive papers) or could not be assigned a unique primary field. All fields with sufficient data show Gini coefficients above 0.76, confirming that extreme concentration is a *universal* feature of STEM research rather than an artifact of any particular discipline. The highest concentration appears in Neuroscience (Gini = 0.95) and Medicine (Gini = 0.93), while Engineering (Gini = 0.83) and Health Professions (Gini = 0.77) show the lowest—but still very high—concentration levels.

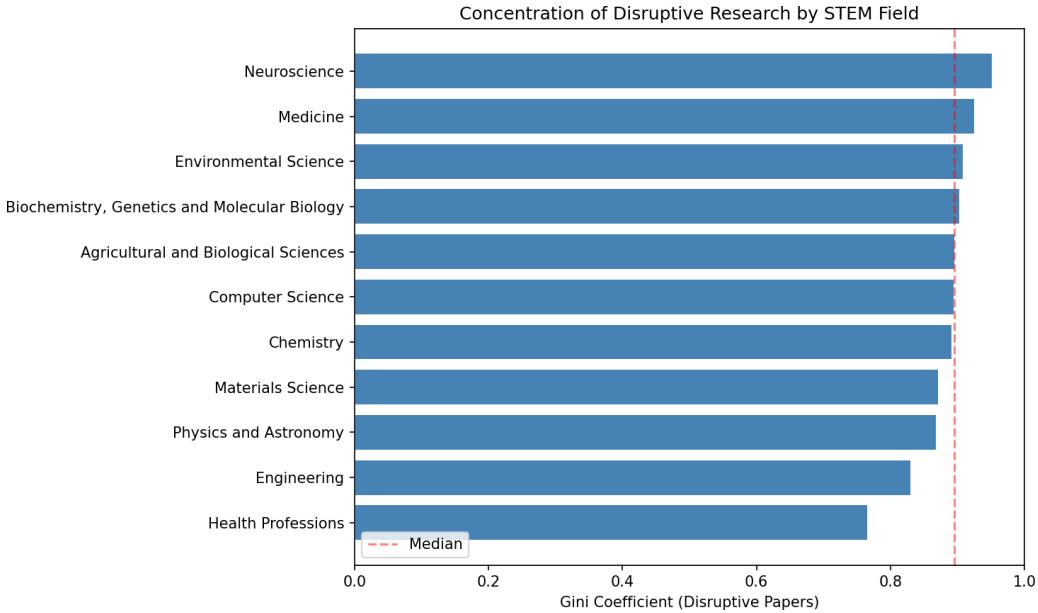


Figure 3: Gini coefficients for disruptive research output across STEM subfields. The dashed red line indicates the median Gini across fields.

The inverse relationship between Gini and the percentage of disruptors within each field is expected: fields where more scientists produce disruptive work will mechanically show lower concentration. Engineering’s notably lower Gini and higher disruptor fraction (19.5%) may reflect the applied character of the field, where innovations have more direct pathways to displacing prior practice.

3.4 Temporal Stability

The concentration of disruptive research is remarkably stable over the three decades studied.

Table 6: Gini Coefficients for Disruptive Output by Five-Year Window

Window	Gini	% Disruptors	N Authors	N Papers
1980–84	0.942	5.8%	104	131
1985–89	0.946	5.5%	165	226
1990–94	0.893	11.1%	207	300
1995–99	0.919	8.1%	321	427
2000–04	0.934	7.0%	530	727
2005–09	0.922	8.0%	828	1,276

The Gini coefficient fluctuates in a narrow band between 0.89 and 0.95 across all six

complete five-year windows, with no discernible secular trend. The disruptor fraction similarly oscillates between 5.5% and 11.1% without a clear directional pattern. (We omit the incomplete 2010–2014 window, which contains only publications from 2010 due to the publication cutoff.) This temporal stability suggests that the concentration of disruptive output reflects a *structural* feature of how scientific knowledge production is organized, rather than a historical contingency or cohort effect.

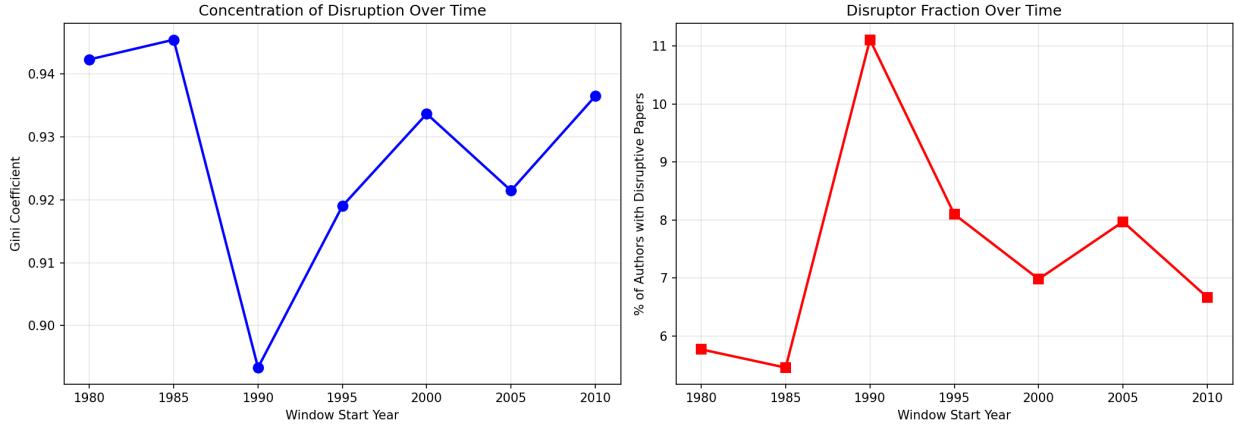


Figure 4: Temporal trends in disruption concentration. *Left:* Gini coefficient for disruptive paper counts within five-year windows. *Right:* percentage of authors producing at least one top-5% disruptive paper.

3.5 Validation Against Population Benchmark

We compared the distributional properties of our A/B disruption scores against the [Park et al. \[2023\]](#) Zenodo benchmark of 22.5 million CD5 scores from Web of Science.

Table 7: Distributional Comparison: Our Sample vs. Zenodo Benchmark

Statistic	Our Sample ($N = 3,421$)	Zenodo ($N = 22,479,429$)
Mean	-0.379	0.040
Median	-0.500	-0.001
Standard deviation	0.479	0.221
% positive disruption	18.4%	22.4%
KS statistic	0.735	—

The KS statistic of 0.735 indicates significant distributional divergence, which is expected for three reasons. First, our A/B disruption index is a two-component approximation of the full five-year CD index used in the Zenodo data, and is known to produce more extreme

values (both positive and negative) due to the omission of Type C citers. Second, our forward citation window is 10 years versus 5 years in the Zenodo data. Third, our sample is conditioned on authors with at least 5 publications and 10 citations, creating a composition difference from the population of all Web of Science papers.

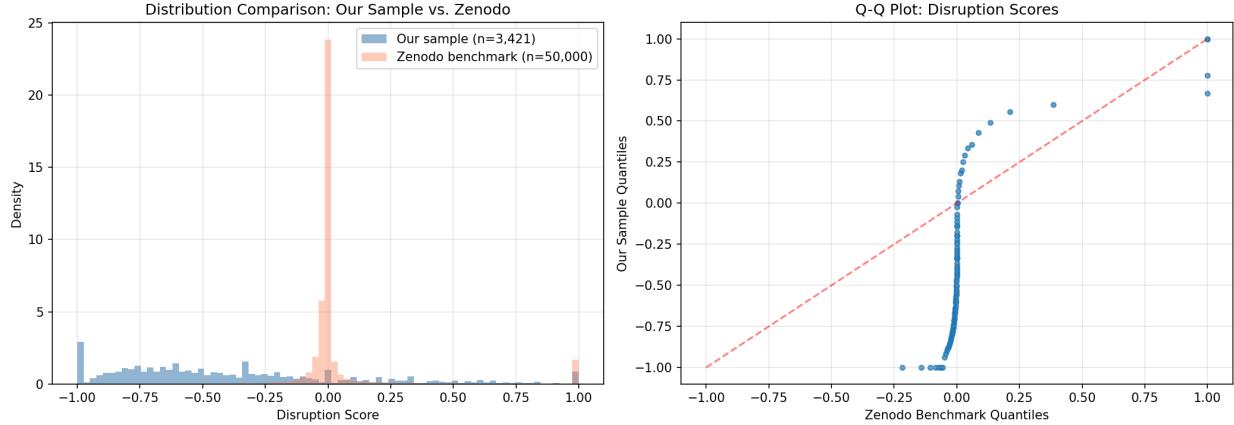


Figure 5: Validation against population benchmark. *Left:* overlapping density histograms of disruption scores from our sample (blue) and the Zenodo benchmark (coral). *Right:* Q-Q plot comparing quantiles of our sample against the Zenodo benchmark.

Despite the absolute distributional differences, the qualitative properties are consistent: both distributions are centered near zero or slightly negative, both are right-skewed, and both show that a large majority of papers have non-positive disruption scores. The key finding—extreme concentration of disruption among a small fraction of authors—is robust to the specific disruption metric used, as it depends on the *relative ranking* of papers within our sample rather than on absolute score calibration.

4 Discussion

4.1 Interpretation

Our findings establish three main results. First, the concentration of disruptive research is extreme: a Gini of 0.90 places it well above the already-skewed distribution of citations (Gini = 0.66) and comparable to the most unequal wealth distributions observed globally. The top 10% of scientists account for over 92% of all top-5% disruptive papers. Second, this concentration is temporally stable and universal across STEM fields, suggesting a structural feature of scientific knowledge production. Third, the observable demographic predictors of disruption are weak ($R^2 \approx 0.10$), with career length being the only robust predictor—a

finding more consistent with the interpretation that disruption requires sustained engagement and accumulated expertise than with the “young genius” narrative.

4.2 Relation to Prior Work

Our results complement [Park et al. \[2023\]](#) by shifting the unit of analysis from papers to scientists. While Park et al. documented that disruption has declined *on average* over time, we show that disruption is concentrated among a thin tail of scientists *within any given period*, and that this concentration has not appreciably changed. Together, these findings suggest that the decline in average disruption documented by Park et al. may reflect changes in the *composition* of the scientific workforce or the *incentive environment* (e.g., the expansion of team science and specialization) rather than a uniform decline in individual scientific capacity.

Our demographic findings align with [Wu et al. \[2019\]](#), who found that small teams disrupt while large teams develop, and with [Lin et al. \[2023\]](#), who characterized disruptors as less central, less prolific, and more likely to be early in their careers. The negative (though non-significant) coefficient on h-index in our models is consistent with the notion that academic stardom and paradigm-shifting work are at best orthogonal and potentially inversely related.

4.3 Limitations

Several limitations warrant discussion.

Disruption metric. Our A/B disruption index omits Type C citers and uses a 10-year forward window rather than the standard 5-year window of the CD5 index. While the two-component variant has been shown to correlate strongly with the full index [[Wu et al., 2019](#)], the omission of Type C citers produces more extreme scores, as reflected in the distributional divergence from the Zenodo benchmark. Our concentration findings depend on *within-sample rankings* rather than absolute scores, which mitigates this concern.

Sample size and selection. Our disruption scores cover 3,421 papers across 1,594 authors—substantially smaller than the population-level analyses of [Park et al. \[2023\]](#). The sample is conditioned on authors with at least 5 publications and 10 citations, which excludes inactive researchers and may underestimate the true concentration of disruption (since including inactive researchers would inflate the non-disruptor count). Conversely, the 5,000-paper subsample for disruption computation introduces sampling variability in the classification of individual papers as disruptive.

STEM restriction. Our analysis is limited to STEM fields and does not address whether similar concentration patterns obtain in the social sciences or humanities, where the dynamics of knowledge displacement may differ.

Career length measurement. Career length is computed from the author’s earliest publication year recorded in OpenAlex’s `counts_by_year` field through the latest scored publication. Where this field is unavailable, we fall back to the span of scored papers only, which may underestimate true career length for authors with few scored works.

Observational design. Our regression models are descriptive, not causal. The association between career length and disruption, for example, may reflect survivorship bias (scientists who produce disruptive work may be more likely to remain in academia) rather than a causal effect of time on disruptive capacity.

4.4 Implications

The extreme concentration of disruptive research has implications for science policy. If paradigm-shifting discoveries are produced by a thin tail of scientists whose observable characteristics do not strongly predict their disruptive capacity, then policies aimed at *identifying* future disruptors *ex ante* (e.g., through track record or prestige metrics) may be inherently limited. Instead, policies that maintain a broad base of independent researchers with diverse approaches—supporting what Kuhn [1962] called “essential tension” between tradition and innovation—may be more effective at sustaining the flow of disruptive discoveries.

The stability of disruption concentration over time also suggests that the widely discussed “decline of disruption” [Park et al., 2023] is not accompanied by a corresponding *equalization* of disruptive capacity. Even in periods of relatively higher average disruption, the output remains concentrated among a small fraction of scientists. This argues against the interpretation that disruption has declined because it has become “democratized” or diffused—rather, the absolute level of disruptive output may have fallen while its distribution across scientists has remained unchanged.

5 Conclusion

Using disruption scores for 3,421 STEM papers by 1,594 authors, we find that disruptive research output is extremely concentrated: the Gini coefficient of 0.900 for disruptive paper counts exceeds the Gini of 0.658 for citations and is stable across six five-year windows

spanning three decades (1980–2009) and eleven STEM subfields. The top 5% of scientists produce half of all top-5% disruptive papers. Observable demographic characteristics explain only 10% of the variance in peak disruption, with career length as the dominant predictor. These findings establish the concentration of disruption as a robust structural feature of STEM research and suggest that the capacity for paradigm-shifting work is driven by factors poorly captured by standard bibliometric indicators.

Data and Code Availability

All analysis code and cached data are available at <https://github.com/banksianr/DisruptionInequality>. The pipeline is implemented in Python and uses the OpenAlex API [Priem et al., 2022] for bibliometric data and the Park et al. [2023] Zenodo dataset (Record 7258379) for population benchmarking.

References

- Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4), 661–703.
- Funk, R. J., & Owen-Smith, J. (2017). A dynamic network measure of technological change. *Management Science*, 63(3), 791–817.
- Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.
- Lin, Z., Yin, Y., Liu, L., & Wang, D. (2023). SciSciNet: A large-scale open data lake for the science of science research. *Scientific Data*, 10, 315.
- Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, 16(12), 317–323.
- Park, M., Leahey, E., & Funk, R. J. (2023). Papers and patents are becoming less disruptive over time. *Nature*, 613, 138–144.
- Price, D. J. de S. (1963). *Little Science, Big Science*. Columbia University Press.
- Price, D. J. de S. (1965). Networks of scientific papers. *Science*, 149(3683), 510–515.
- Priem, J., Piwowar, H., & Orber, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv preprint arXiv:2205.01833*.

Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science*, 43(9), 628–638.

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