

Slowed canonical progress in large fields of science

Johan S. G. Chu^{a,1}  and James A. Evans^{b,c,d} 

^aKellogg School of Management, Northwestern University, Evanston, IL, 60208; ^bDepartment of Sociology, University of Chicago, Chicago, IL, 60637; ^cKnowledge Lab, University of Chicago, Chicago, IL, 60637; and ^dSanta Fe Institute, Santa Fe, NM, 87501

Edited by Kenneth W. Wachter, University of California, Berkeley, CA, and approved August 25, 2021 (received for review December 8, 2020)

In many academic fields, the number of papers published each year has increased significantly over time. Policy measures aim to increase the quantity of scientists, research funding, and scientific output, which is measured by the number of papers produced. These quantitative metrics determine the career trajectories of scholars and evaluations of academic departments, institutions, and nations. Whether and how these increases in the numbers of scientists and papers translate into advances in knowledge is unclear, however. Here, we first lay out a theoretical argument for why too many papers published each year in a field can lead to stagnation rather than advance. The deluge of new papers may deprive reviewers and readers the cognitive slack required to fully recognize and understand novel ideas. Competition among many new ideas may prevent the gradual accumulation of focused attention on a promising new idea. Then, we show data supporting the predictions of this theory. When the number of papers published per year in a scientific field grows large, citations flow disproportionately to already well-cited papers; the list of most-cited papers ossifies; new papers are unlikely to ever become highly cited, and when they do, it is not through a gradual, cumulative process of attention gathering; and newly published papers become unlikely to disrupt existing work. These findings suggest that the progress of large scientific fields may be slowed, trapped in existing canon. Policy measures shifting how scientific work is produced, disseminated, consumed, and rewarded may be called for to push fields into new, more fertile areas of study.

scientific progress | durable dominance | entrepreneurial futility | science policy | science of science

A straightforward view of scientific progress would suggest more is better. The more papers published in a field, the greater the rate of scientific progress; the more researchers, the more ground covered. Even if not every article is earth shaking in its impact, each can contribute a metaphorical grain of sand to the sandpile, increasing the probability of an avalanche, wherein the scientific landscape is reconfigured and new paradigms arise to structure inquiry (1, 2). The publication of more papers also increases the probability at least one of them contains an important innovation. A disruptive new idea can destabilize the status quo, siphoning attention from previous work and garnering the lion's share of new citations (3, 4).

Policy reflects this more-is-better view. Scholars are evaluated and rewarded on productivity. Publishing many articles within a set period of time is the surest path to tenure and promotion. Quantity remains the measuring stick at the university (5) and the national levels (6), where comparisons focus on the total number of publications, patents, scientists, and dollars spent.

"Quality" is also predominantly judged quantitatively. Citation counts are used to measure the importance of individuals (7), teams (8), and journals (9) within a field. At the paper level, the assumption is that the best and most valuable papers will attract more attention, shaping the research trajectory of the field (10).

Here, however, we predict that when the number of papers published each year grows very large, the rapid flow of new papers can force scholarly attention to already well-cited papers and limit attention for less-established papers—even those with novel, useful, and potentially transformative ideas. Rather than

causing faster turnover of field paradigms, a deluge of new publications entrenches top-cited papers, precluding new work from rising into the most-cited, commonly known canon of the field.

These arguments, supported by our empirical analysis, suggest that the scientific enterprise's focus on quantity may obstruct fundamental progress. This detrimental effect will intensify as the annual mass of publications in each field continues to grow—which is almost inevitable given the entrenched, interlocking structures motivating publication quantity. Policy measures restructuring the scientific production value chain may be required to allow mass attention to concentrate on promising, novel ideas.

This study focuses on the effects of field size: The number of papers published in a field in a given year. Previous studies have found that citation inequality is increasing across a range of disciplines (11), at least partially driven by processes of preferential attachment (12, 13). Papers do not always maintain their citation levels and rankings over the years, however. Disruptive papers can eclipse prior work (4) and natural fluctuations in citation numbers can upset rankings (14). We predict that when fields are large, the dynamics change. The most-cited papers become entrenched, garnering disproportionate shares of future citations. New papers cannot rise into canon by amassing citations through processes of preferential attachment. Newly published papers rarely disrupt established scholarship.

Two mechanisms underlie these predictions (15). First, when many papers are published within a short period of time, scholars are forced to resort to heuristics to make continued sense of the field. Rather than encountering and considering intriguing new ideas each on their own merits, cognitively overloaded reviewers

Significance

The size of scientific fields may impede the rise of new ideas. Examining 1.8 billion citations among 90 million papers across 241 subjects, we find a deluge of papers does not lead to turnover of central ideas in a field, but rather to ossification of canon. Scholars in fields where many papers are published annually face difficulty getting published, read, and cited unless their work references already widely cited articles. New papers containing potentially important contributions cannot garner field-wide attention through gradual processes of diffusion. These findings suggest fundamental progress may be stymied if quantitative growth of scientific endeavors—in number of scientists, institutes, and papers—is not balanced by structures fostering disruptive scholarship and focusing attention on novel ideas.

Author contributions: J.S.G.C. designed research; J.S.G.C. and J.A.E. performed research; J.A.E. secured access to data; J.S.G.C. analyzed data; and J.S.G.C. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

¹To whom correspondence may be addressed. Email: johan.chu@kellogg.northwestern.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2021636118/-DCSupplemental>.

Published October 4, 2021.

and readers process new work only in relationship to existing exemplars (16–18). A novel idea that does not fit within extant schemas will be less likely to be published, read, or cited. Faced with this dynamic, authors are pushed to frame their work firmly in relationship to well-known papers, which serve as “intellectual badges” (19) identifying how the new work is to be understood, and discouraged from working on too-novel ideas that cannot be easily related to existing canon. The probabilities of a breakthrough novel idea being produced, published, and widely read all decline, and indeed, the publication of each new paper adds disproportionately to the citations for the already most-cited papers.

Second, if the arrival rate of new ideas is too fast, competition among new ideas may prevent any of the new ideas from becoming known and accepted field wide. To see why this is so, consider a sandpile model of idea spread in a field. When sand is dropped on a sandpile slowly, one grain at a time, waiting for movement on the sandpile to stop before dropping the next grain, the sandpile over time reaches a scale-free critical state wherein one dropped grain of sand can trigger an avalanche over the whole area of the pile (2). But when sand is dropped at a rapid rate, neighboring miniavalanches interfere with each other, and no individual grain of sand can trigger pile-wide shifts (20). The faster the rate of sand dropping the smaller the domain each new grain of sand can affect. If the arrival rate of papers is too fast, no new paper can rise into canon through localized processes of diffusion and preferential attachment.

The arguments above yield six predictions, two each predicting durable dominance of the most-cited papers, entrepreneurial futurity for newly published papers, and decrease in the disruptiveness (3, 4) of newly published papers. Compared to when a field produces few publications each year, when that field produces many new publications each year: 1) new citations will be more likely to cite the most-cited papers rather than less-cited papers; 2) the list of most-cited papers will change little year to year—the canon ossifies; 3) the probability a new paper eventually becomes canon will drop; 4) new papers that do rise into the ranks of those most cited will not do so through gradual, cumulative processes of diffusion; 5) the proportion of newly published papers developing existing scientific ideas will increase and the proportion disrupting existing ideas will decrease; and 6) the probability of a new paper becoming highly disruptive will decline.

Results

Each of these predictions is borne out in citation patterns across the Web of Science dataset, as shown in Figs. 1–4. As fields get larger, the most-cited papers become durably dominant, entrenched atop the citation distribution. New papers, in contrast, suffer

diminished probability of ever becoming very highly cited and cannot gradually accumulate attention over time. Published papers tend to develop existing ideas more than disrupt them, and rarely launch disruptive new streams of research.

The most-cited papers garner disproportionately higher shares of citations in larger fields. The largest fields have a Gini coefficient of citation shares of around 0.5 (Fig. 1A), which is as large as income inequality in the most unequal countries—only China and South Africa have Gini coefficients higher than 0.5 (21). Disproportionate numbers of citations to top-cited papers drive this increase in unequal attention. For example, when the field of Electrical and Electronic Engineering published ~10,000 papers a year, the top 0.1% most-cited papers collected 1.5% and the top 1% most-cited collected 8.6% of total citations. When the field grew to 50,000 published papers a year, the top 0.1% captured 3.5% of citations, and the top 1% captured 11.9%. When the field was larger still with 100,000 published papers per year, the top 0.1% received 5.7% of citations within the field and the top 1% received 16.7%. The bottom 50% least-cited papers in contrast decreased in share as the field grew larger, dropping from garnering 43.7% of citations at 10,000 papers to slightly above 20% at both 50,000 and 100,000 papers per year.

Canons crystallize as fields grow large. Churn in the identity and ordering of the most-cited papers decreases with larger field size. The pattern holds consistent when looking at data across all fields and at individual large fields across time: When the number of papers published per year is larger, the rank correlation between the top-50 most-cited papers in the focal year and the next increases (Fig. 1B). The predicted Spearman rank correlation of the top-50 most-cited list in a field between subsequent years increases from 0.25 when 1,000 papers are published in the focal year to 0.74 when 100,000 papers are published yearly.

This crystallization of canon happens because the most-cited papers maintain their number of citations year over year when fields are large, while all other papers' citation counts decay. Fig. 2 displays the predicted ratio of current year to previous year citations for papers at various percentiles of citation-share ranking. In years where few papers are published, the ratio for the most-cited papers is significantly below 1 and not much different from less-cited papers. When the number of papers published grows large, however, the ratio for the most-cited papers is close to 1, significantly higher than that of less-cited papers. In very large field-years, with about 100,000 papers published, the most-cited papers on average see no decline in their numbers of citations received year over year. Papers just outside the top 1% most cited in the field-year, in contrast, lose on average about 17% of their citation counts each year, and those at the fifth percentile and

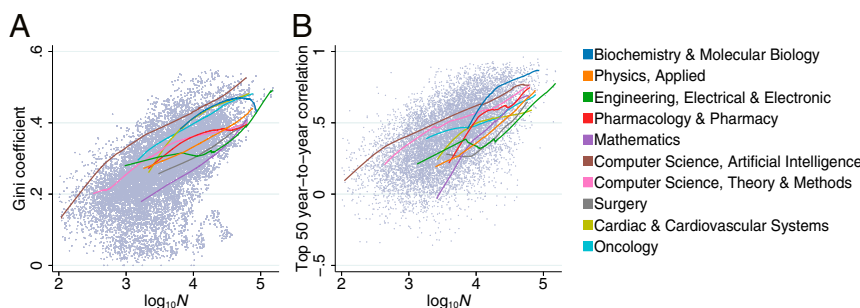


Fig. 1. Changes in citation dynamics by size of field: Durable dominance of canon. X axes are logged (base 10) number of papers published in the subject-year (N). Each dot represents a subject-year in the Web of Science. Lines are lowest trendlines for the 10 largest nonmultidisciplinary subjects, listed in order of total number of papers published over all years in the dataset in the legend. (A) Gini coefficient of citation-share inequality by subject-year. The most-cited papers garner a larger proportion of new citations in years when more papers are published in a field. (B) Spearman rank correlation of the top-50 most-cited list between adjacent years by subject-year. The ordering of most-cited papers is more stable when more papers are published in a field.

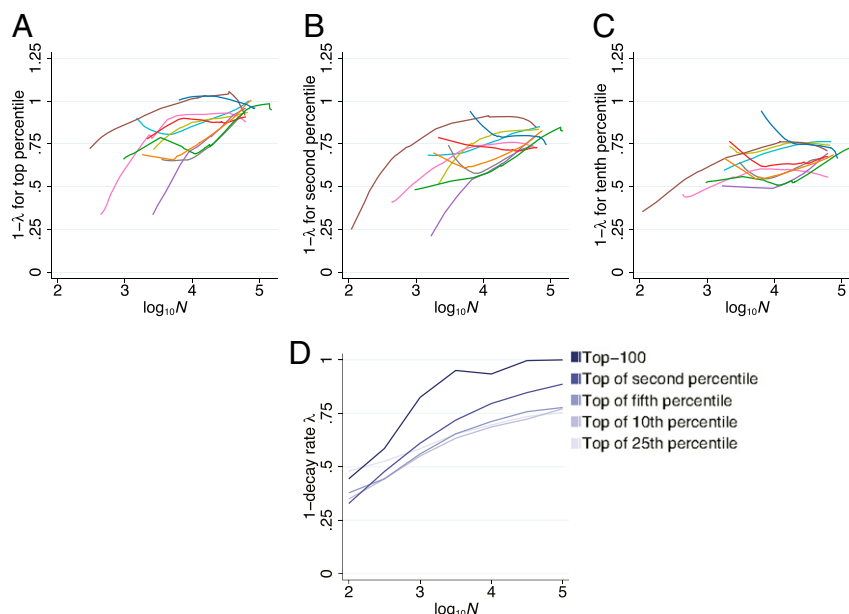


Fig. 2. Citation decay rates. *X* axes are logged (base 10) number of papers published in the subject-year (N). (A–C) 1-decay rate (λ) for top-percentile (A), 2nd-percentile (B), and 10th-percentile (C) most-cited papers in the 10 largest fields (see Fig. 1 for legend). 1- λ trends toward 1 for the top-percentile most-cited papers in the largest fields; the most-cited papers on average maintain their number of citations year over year when many new papers are published in the field. 1- λ trends toward values <1 for the second and lower percentiles; all but the most-cited papers receive diminishing numbers of citations on average year over year in large fields. (D) 1- λ across all subject-years. The pattern is consistent with A–C. Across all large fields, only the top-cited papers maintain their level of citations on average year to year; all other papers on average receive fewer citations year over year.

below trend toward losing a quarter of their citations year over year.

The probability of a paper ever reaching (even for 1 y) the top 0.1% most cited in its field shrinks when it is published in the same year as many others. This holds true cross-sectionally across fields in the same year, and across years in individual fields (Fig. 3A). When papers in large fields do become most cited, it is rarely through a process of local diffusion and preferential attachment. Fig. 3B presents the median time in years for an article to break into the field's canon, conditional on the paper ever becoming one of the top cited in its field. When a field is small, papers rise slowly over time into the top 0.1% most cited, consistent with a process of cumulative attention gathering. A linear regression across all subjects for the year 1980 predicts a median time of 9 y for a successful paper to reach the 0.1% most cited in its field when published in the same year as 1,000 other papers in the field. Papers entering the canon in the largest fields, by contrast, shoot quickly to the top, inconsistent with a cumulative process where scholars discover new work by reading references cited in others' work. The same regression predicts a median of less than a year for papers to reach the top 0.1% in large fields with 100,000 papers published each year.

Most papers published in the same year as many others build on, rather than disrupt, existing literature (Fig. 4A). A logistic fit predicts 49% of papers have disruption measure (3, 4) $D > 0$ (and conversely 51% $D < 0$) when 1,000 papers are published in the field-year. The predicted proportion of disruptive papers drops to 27% when 10,000 papers are published and 13% at 100,000 papers. Even when $D > 0$, the disruptive impact of a newly published paper is muted in larger fields. Fig. 4B presents the proportion of new papers by field-year that rank in the top-5 percentile of disruption measure. Lowess estimates show the proportion of new papers with top-5 percentile disruption measure shrinks from 8.8% at 1,000 papers published in the field-year to 3.6% at 10,000 papers per year and 0.6% at 100,000 papers.

These empirical results are aligned with our theory's predictions. Our current analyses cannot, however, rule out other causal explanations. The *SI Appendix* considers the most salient alternative explanation—that the changes observed are driven by the passage of time and maturing of fields rather than field size. While the number of papers published in a field tends to increase over time, this increase is not lockstep. Analysis shows significant effects of field size over and above effects of time (*SI Appendix*, Table S1 and Fig. S1). The *SI Appendix* also examines the

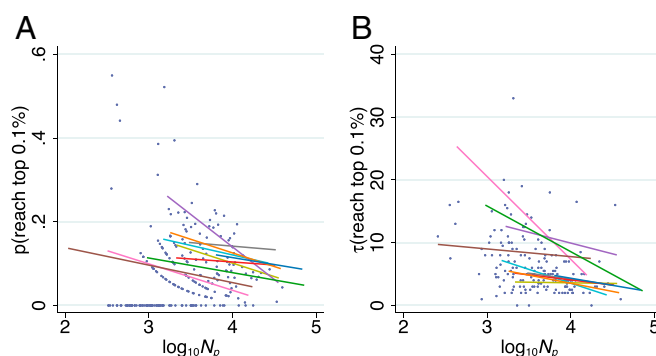


Fig. 3. New papers rising into canon. *X* axes are logged (base 10) number of papers published in the subject-year (N_p). Each dot represents a subject in the year 1980. Lines are linear trendlines for the 10 largest non-multidisciplinary subjects (see Fig. 1 for legend) for papers published in 2000 and earlier. (A) Probability (p , in %) of a paper ever reaching the top 0.1% of most-cited articles. The probability of a newly published paper ever reaching the top 0.1% most cited in its field decreases when it is published in the same year as more papers in its field. (B) Median number of years (τ) for a paper to reach the top 0.1% of most-cited articles, conditional on reaching the top 0.1%. Papers published in the same year as many others in their subject do not gradually and cumulatively build up citations to reach the top 0.1%.

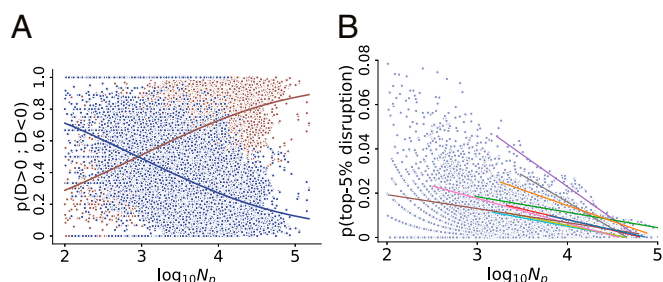


Fig. 4. Propensity for disruption. *X* axes are logged (base 10) number of papers published in the subject-year (N_p). Each dot represents a subject-year. (A) Proportion (p) of papers published in the subject-year with disruption measure, D , from Wu et al. (4) greater than 0 ($D > 0$; blue dots and line) and lesser than 0 ($D < 0$; red dots and line). Lines are logistic fits to the data. Papers published in years with many others in their subject tend to develop existing ideas rather than introduce disruptive new ones. (B) Proportion (p) of papers published in the subject-year with top 5-percentile disruption measure ($D \geq 0.0256$). Lines are linear trendlines for the 10 largest non-multidisciplinary subjects (see Fig. 1 for legend). Papers published in the same year as many others in their subject are unlikely to be highly disruptive.

mechanisms of change. We find that veteran scholars change their citation patterns as a field grows. While field size at the time a scholar entered the field does influence their propensity to reference the most-cited articles, the field's size when an article is published has a much stronger effect (*SI Appendix, Tables S2 and S3*). Even well-established, veteran scholars come to cite canonical articles much more often when many other papers are also being published.

Discussion

These findings suggest troubling implications for the current direction of science. If too many papers are published in short order, new ideas cannot be carefully considered against old, and processes of cumulative advantage cannot work to select valuable innovations. The more-is-better, quantity metric-driven nature of today's scientific enterprise may ironically retard fundamental progress in the largest scientific fields. Proliferation of journals and the blurring of journal hierarchies due to online article-level access can exacerbate this problem.

Reducing quantity may be impossible. Proscribing the number of annual publications, shuttering journals, closing research institutions, and reducing the number of scientists are hard-to-swallow policy prescriptions. Even if a scientist wholeheartedly agreed with the implications of our study, curtailing their output would be impractical given the damage to their career prospects and those of their colleagues and students, for example. Limiting article quantity without altering other incentives risks deterring the publication of novel, important new ideas in favor of low-risk, canon-centric work.

Still, some changes in how scholarship is conducted, disseminated, consumed, and rewarded may help accelerate fundamental progress in large fields of science. A clearer hierarchy of journals with the most-prestigious, highly attended outlets devoting pages to less canonically rooted work could foster disruptive scholarship and focus attention on novel ideas. Reward and promotion systems, especially at the most prestigious institutions, that eschew quantity measures and value fewer, deeper, more novel contributions could reduce the deluge of papers

competing for a field's attention while inspiring less canon-centric, more innovative work. A widely adopted measure of novelty vis a vis the canon could provide a helpful guide for evaluations of papers, grant applications, and scholars. Revamped graduate training could push future researchers to better appreciate the uncomfortable novelty of ideas less rooted in established canon. These measures, while not easy to implement across large fields, may help push scholarship off the local attractor of existing canon and toward more novel frontiers.

The current study is at the level of fields and large subfields, and one could argue that progress now occurs at lower sub-disciplinary levels. To examine lower levels at scale requires more precise methods for classifying papers—perhaps using temporal citation network community detection—than are currently available. But note that the fields and subfields identified in the Web of Science correspond closely to real-world self-classifications of journals and departments. Established scholars transmit their cognitive view of the world to their students via field-centric reading lists, syllabi, and course sequences, and field boundaries are enforced through career-shaping patterns of promotion and reward.

It may be that progress still occurs, even though the most-cited articles remain constant. While the most-cited article in molecular biology (22) was published in 1976 and has been the most-cited article every year since 1982, one would be hard pressed to say that the field has been stagnant, for example. But recent evidence (23) suggests that much more research effort and money are now required to produce similar scientific gains—productivity is declining precipitously. Could we be missing fertile new paradigms because we are locked into overworked areas of study?

Materials and Methods

We utilize the Web of Science dataset, analyzing papers published between 1960 and 2014 inclusive. The resulting dataset contains 90,637,277 papers and 1,821,810,360 citations. The Web of Science classifies academic fields, or in some cases, large subfields, into what it terms subjects. There are 241 subjects in the classification, and we use these as the basis for our field-level analyses. The annual count of citations received by a focal paper from newly published papers in the same subject constitutes our main variable of interest.

To calculate 1–decay rate (λ) for the 10 largest nonmultidisciplinary subjects (Fig. 2 A–C), for each subject, we binned years by the base 10 log of number of publications (cutpoints at 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, and 5.5), and paper years by percentile most cited in the field-year (cutpoints at 1, 2, 3, ..., 100). For each (logged number of publications) \times (citation percentile) bin, we regressed the number of citations to a paper the subsequent year on number of citations to a paper in the focal year. The coefficient of this regression yields 1– λ .

To calculate 1– λ across all subjects (Fig. 2D), we selected the top 100 most-cited papers from each subject-year in the 1st, 2nd, 5th, 10th, and 25th percentiles. We binned subject-years by the base 10 log of number of publications (cutpoints at 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, and 5.5). For each bin \times selected percentile, we regressed the number of citations to a paper the subsequent year on number of citations to a paper in the focal year. The coefficient of this regression yields 1– λ .

Data Availability. Web of Science data are available from Clarivate Analytics (<https://clarivate.libguides.com/rawdata>). Disruption score data (4) are from Lingfei Wu, with replication data at <https://doi.org/10.7910/DVN/JPWNK>. Aggregate data used to produce figures are available at <https://github.com/advrk/PNAS2021>.

ACKNOWLEDGMENTS. Thanks to seminar participants at Hong Kong University of Science and Technology and Northwestern University for comments, to Clarivate Analytics for supplying Web of Science data, PNAS reviewers for suggestions on citation dynamics literature and disruption analysis, and Lingfei Wu for help with disruption measures.

1. T. S. Kuhn, *The Structure of Scientific Revolutions* (University of Chicago Press, Chicago, ed. 2, 1970).
2. P. Bak, C. Tang, K. Wiesenfeld, Self-organized criticality: An explanation of the 1/f noise. *Phys. Rev. Lett.* **59**, 381–384 (1987).
3. R. J. Funk, J. Owen-Smith, A dynamic network measure of technological change. *Manage. Sci.* **63**, 791–817 (2017).

4. L. Wu, D. Wang, J. A. Evans, Large teams develop and small teams disrupt science and technology. *Nature* **566**, 378–382 (2019).
5. C. Baden-Fuller, F. Ravazzolo, T. Schweizer, Making and measuring reputations: The research ranking of European business schools. *Long Range Plann.* **33**, 621–650 (2000).
6. Scientific American, "The world's best countries in science" (2017). <https://www.scientificamerican.com/article/the-worlds-best-countries-science/>. Accessed 31 August 2020.

7. S. Alonso, F. J. Cabrerizo, E. Herrera-Viedma, F. Herrera, *hg-index*: A new index to characterize the scientific output of researchers based on the *h*- and *g*-indices. *Scientometrics* **82**, 391–400 (2010).
8. B. F. Jones, S. Wuchty, B. Uzzi, Multi-university research teams: Shifting impact, geography, and stratification in science. *Science* **322**, 1259–1262 (2008).
9. G. F. Davis, Editorial essay: Why do we still have journals? *Adm. Sci. Q.* **59**, 193–201 (2014).
10. J. G. Foster, A. Rzhetsky, J. A. Evans, Tradition and innovation in scientists' research strategies. *Am. Sociol. Rev.* **80**, 875–908 (2015).
11. M. W. Nielsen, J. P. Andersen, Global citation inequality is on the rise. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2012208118 (2021).
12. H. Jeong, Z. Neda, A.-L. Barabási, Measuring preferential attachment in evolving networks. *Europhys. Lett.* **61**, 567–572 (2003).
13. D. Wang, C. Song, A.-L. Barabási, Quantifying long-term scientific impact. *Science* **342**, 127–132 (2013).
14. N. Blumm *et al.*, Dynamics of ranking processes in complex systems. *Phys. Rev. Lett.* **109**, 128701 (2012).
15. J. S. G. Chu, A theory of durable dominance. *Stat. Sci.* **3**, 498–512 (2018).
16. A. Tversky, D. Kahneman, Judgment under uncertainty: Heuristics and biases. *Science* **185**, 1124–1131 (1974).
17. B. Schwartz, *The Paradox of Choice: Why More is Less* (Harper Collins, New York, 2004).
18. E. W. Zuckerman, The categorical imperative: Securities analysts and the illegitimacy discount. *Am. J. Sociol.* **104**, 1398–1438 (1999).
19. A. L. Stinchcombe, Should sociologists forget their mothers and fathers. *Am. Sociol.* **17**, 2–11 (1982).
20. C. Adami, J. Chu, Critical and near-critical branching processes. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **66**, 011907 (2002).
21. Organisation for Economic Co-Operation and Development (OECD), Key indicators on the distribution of household disposable income and poverty (2017). <https://www.oecd.org/social/soc/IDD-Key-Indicators.xlsx>. Accessed 31 August 2020.
22. M. M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).
23. N. Bloom, C. I. Jones, J. Van Reenen, M. Webb, Are ideas getting harder to find? *Am. Econ. Rev.* **110**, 1104–1144 (2020).