

## Aging Scientists and Slowed Advance

**Authors:** Haochuan Cui<sup>1</sup>, Lingfei Wu<sup>2</sup>, James A. Evans<sup>3,4\*</sup>

### Affiliations:

<sup>1</sup> School of Systems Science, Beijing Normal University, Beijing, P.R. China 100875

<sup>2</sup> School of Computing and Information, University of Pittsburgh, Pittsburgh, PA 15260

<sup>3</sup> Department of Sociology, University of Chicago, 1126 E 59th St, Chicago, IL 60637.

<sup>4</sup> Knowledge Lab, University of Chicago, 5735 South Ellis Avenue, Chicago, IL 60637.

\*Corresponding author. E-mail: [jevans@uchicago.edu](mailto:jevans@uchicago.edu) (J.E.)

**With rising life expectancies around the world and an older scientific workforce than ever before<sup>1</sup>, what does aging mean for individual scientists and what do aging scientists mean for scientific progress? Prior research focuses on star scientists<sup>2,3</sup>, their changing dates and rates of breakthrough success across history<sup>4,5</sup>. We examine how all scientists and scholars age in terms of how their stream of ideas and contributions relate to the evolving frontier of knowledge, and how demographically aging fields relate to field-level progress. Analyzing more than 244 million scholars across 241 million articles over the last two centuries, here we show that for all fields, periods, and impact levels, scientists' research ideas and references age over time, their research is less likely to disrupt the state of science and more likely to criticize emerging work. Early success accelerates scientist aging; while changing institutions and fields and collaborating with young scientists slows it. These patterns aggregate within fields such that those with a higher proportion of older scientists experience a lower churn of ideas and more rapid individual aging, suggesting a universal link between aging, activity, and advance.**

What is the relationship between aging and the character of scientific advance? Consider the academic life-course of Albert Einstein. As a 26-year old examiner at the Patent Office in Bern, Switzerland, Einstein wrote four papers that revolutionized science's understanding of space, time, mass, and energy. In the decade that followed, he entered the academy and generalized Special Relativity. From mid-career, he sought unsuccessfully to assemble core physics into a unified field theory<sup>6</sup>, and spent his later years criticizing quantum mechanics' probabilistic interpretation of the universe, posing quantum paradoxes until the end—some of which were observed—ironically entrenching the quantum age<sup>7</sup>. Was the young Einstein unique in promoting new areas with minimal reference to competitors? Was the old Einstein unique in defending past work against emerging ideas inconsistent with or irrelevant to it? Max Planck quipped that science advanced by funerals<sup>8</sup>, a theory generalized in Thomas Kuhn's work on scientific revolutions wrought by young scientists<sup>9</sup>. A recent analysis of the premature demise of nearly 500 star life scientists and a collection of comparable scientists who continued their work showed that a career cut short was associated with a systematic boost of innovation in star scientists' subfields<sup>2</sup>. We generalize this to posit that all scientists tend to embrace new ideas while young but resist them when older. We propose and demonstrate how this aggregates such that the demography of a field's scientists anticipates its churn of ideas and rate of advance.

In this study, we characterize how scientists in all fields and at all levels of productivity age relative to the collective frontier of scientific knowledge. We contribute to prior literature on aging in science by linking scientific advance and obsolescence<sup>10,11</sup> with individual aging<sup>12</sup> and demonstrate that with age comes not only a preference for aging ideas, but active defense against new ones. We further show how this individual proclivity toward defense accumulates within fields and forecasts a reduction in the churn of new ideas conceived and published. By linking the demography of the scientific workforce<sup>1</sup> to the use and production of new ideas, we offer novel insight to concerns about diminishing returns to past scientific investment and efforts<sup>4,5,13,14</sup>.

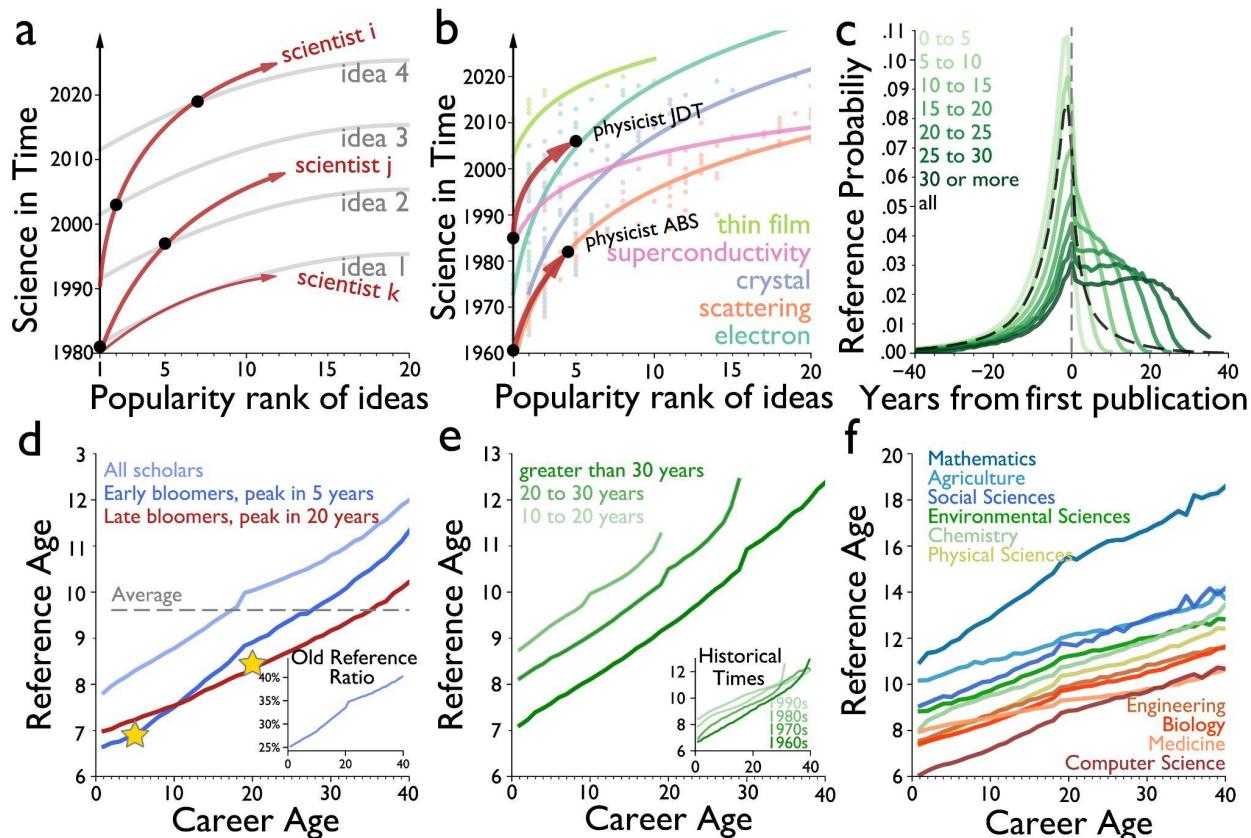
This article makes three contributions to research on the relationship between age and scientific progress. Prior work has emphasized distinct creativity peaks across fields<sup>1,11,12</sup>, and over the life course<sup>3,4</sup>. In contrast, our work focuses on how scientific attention shifts over the life-course, demonstrating a universal pattern across fields and time that aging scientists linearly shift from a focus on present to a focus on past work. From the moment their careers begin, scientists' newly minted ideas move further from the frontier of collective attention. These patterns accelerate with early success but are forestalled by career transitions to new institutions, new topics, and younger collaborators. Second, this article shows for the first time that the process of aging attention is not passive, but active. As scientists age, they not only promote older work, but increasingly police the boundaries of their fields by criticizing the work of younger scholars and contrasting it with their own. Third, we demonstrate that fields aging faster are associated with a lower aggregate churn of ideas. We cannot identify causal mechanisms, but we show that it is much more common for aging to precede slowdown in the churn of ideas within a field than for decreased churn to precede exodus or avoidance by young scientists.

## Research Design

We explore these questions by analyzing the Microsoft Academic Graph (MAG) containing 244,359,707 name-disambiguated authors across 240,874,887 articles published 1800 to 2020

and corroborate them with Clarivate's Web of Science (WOS) database. We apply a neural-network-based NLP model to classify the use of citations—for strength and support or contrast and criticism<sup>15,16</sup>—for 31 million papers and 236 million citations from 1840 to 2020. We focus on a subset of approximately one million scientists and scholars with career lengths of twenty years or longer in which we track shifts in institutional affiliation, changes in topic and field, and citation attention over time. The results we present in the main figures show the behavior of a subset of 1.1 million scholars, with career ages of twenty years or longer who published at least ten papers. In the supplement, we show that these findings are consistent with all scientific and scholarly authors for all papers in MAG, and that these patterns are supported by the WOS database (see details in Methods and SI).

As scientists age, we track (A) the age of articles they cite (Fig. 1), (B) the prevalence of keywords that characterize their research<sup>17,18</sup> (Fig. 2), (C) the frequency of indicators inferred from citation context that referenced work is being used constructively or critically<sup>15,16</sup> (Fig. 3), and (D) the disruption of their work—the degree to which it creates new directions by eclipsing citations to the prior work on which it builds<sup>19,20</sup>. We then investigate the link between a field's proportion of aging scientists and field-level patterns of reference aging and keyword turnover, extending analysis from the individual to the field, which recursively influences individual scientists<sup>21</sup>. Finally, we explore behavior that may mitigate intellectual aging, including collaborating with younger coauthors, moving to new institutions and exploring new topics (Fig. 4).



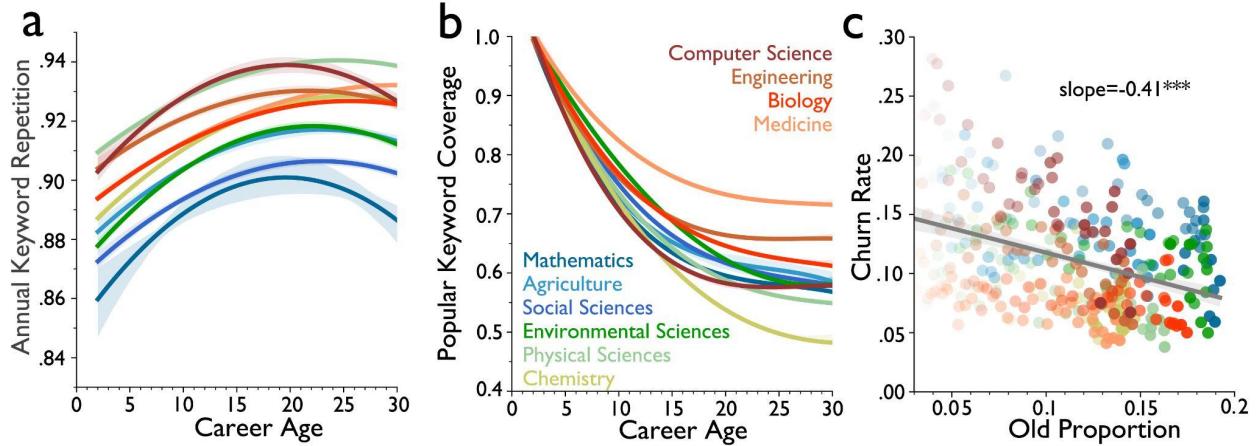
**Fig. 1: Aging scientists tend to cite older literature.** **a**, Conceptual illustration of how scientists age faster than science as a whole, but slower than their first ideas, which we plot empirically with a small sample of scientists and

ideas (e.g., electron, scattering, ion, superconductivity) in **b**. **c**, Demonstrates this pattern for our entire sample of 4,403,830 scientists who began publishing between 1960 and 2015, with the most cited reference papers for all scientists published right before a scientist's first article. Scientists who continue to publish age more slowly. We analyzed 1.1 million scientists who actively published over two decades or longer from 1960 to 2020 and plotted the annual average reference age against career age (see Methods for details). **d**, Average reference age monotonically increases by 50% (from eight to twelve years) over a career of forty years (light blue curve), deviating from the null model that displays the average reference age of all analyzed papers (dotted gray line). The same pattern holds if we calculate the ratio of old references (ten years and older), which increases by 60% from 0.25 to 0.4 over a career of forty years. We compared "early-bloomers" versus "late-bloomers" by analyzing whether the most cited paper (yellow star) of a scientist appears within the first five years or after twenty years, respectively. Both cite older literature over time, but "early-bloomers" experience a faster increase in reference age. Increasing reference ages for aging scientists is robust across career length **e**, historical times (**e**, inset), and academic disciplines (**f**). The mean (intercept) and rate (slope) of increase in reference age slightly vary across disciplines: Mathematics has the highest mean value of 12.7 years while Computer Science has the lowest value of 7.12 years (see Table S1 for OLS regression estimates on a sample of scholars who have published for 10 or more years). In all panels, bootstrapped 95% CIs are displayed as the envelope around each line but are barely invisible.

### Aging Scientists and Fields Cite Older Literature

Across the past half-century of research, scientists collectively favor new findings, tending to cite articles within a decade of publication (average reference age = 9.78, std.= 6.88, Extended Data Fig. 1). Individual scientists, however, commit to older ideas and references as they age (OLS  $\beta = 0.102$ ,  $p < 0.01$ , Fig. 1a-f), a pattern invariant to length of career or time period ( $\beta = [.094-.128]$ ,  $p < 0.01$ , Fig. 1e and Extended Data Fig. 2) and universal across all scientific and scholarly fields we investigate (Fig. 1f). Early career success accelerates this aging process ( $\beta = 0.117$ ,  $p$ -value  $< 0.01$ ) as scientists manage their reputations and follow opportunities to remain focused on topics from acclaimed early work and its influences (Fig. 1d). The average age of references differs by field (see Table S1). Mathematicians reference the oldest literature, published an average of thirteen years prior, and they age at the fastest rate ( $\beta = 0.143$ ,  $p < 0.01$ )—methods of proof change least of any field's techniques over time, preserving the value of older work for older scientists. Computer scientists, by contrast, whose methods change and papers obsolesce quickly, build on work published an average of only seven years before ( $\beta = 0.096$ ,  $p < 0.01$ ).

The observed association between career stage and reference age at the individual level offers insight into the comparable process at the field level. Science increasingly relies on aging scientists to produce new ideas (Extended Data Fig. 3), with the highest fraction from social science (22.7% with a career age of at least 20 years) and the lowest from computer science (13.2% with a career age  $>20$ , Extended Data Fig. 4). Meanwhile, newly published papers ( $<10$  years) are decreasingly likely to be referenced and built upon (29% decrease, Extended Data Fig. 5). Hierarchical regression models allow us to estimate a persistent positive effect of scientist aging on aging citations, controlling for average field age and team size ( $\gamma_{100}=.126$ ,  $p < .01$ , see Methods and Table S2).



**Fig. 2: Aging scientists tend to work on familiar topics.** We track changes in research topics for over 566,369 scientists and the popularity of topics in their fields (see Methods for details). **a**, We quantify paper topic similarity between two successive years for a scientist by calculating the fraction of repeated topics represented in MAG field-of-study keywords<sup>29</sup>. Topic similarity increases and stabilizes with career age, manifesting inertia of focus on familiar research topics. The inset shows the distribution of the most cited reference over scientists' careers, which peaks immediately before publication of the first paper, implying the fixation of key references. **b**, We quantify each scientists' field relevance by calculating the coverage of popular topics in the field within their annual publications. Popular keywords are defined as top 1% MAG field-of-study keywords in the field for a given year revealing that field relevance decays over the career. **c**, Churn rate, defined as the fraction of new topics within a field in one year compared against those of the prior, decreases with the proportion of aging scientists twenty publication-years old or older. This panel displays statistics for ten fields from 1980 to 2020. Data points from recent years are darker. In all panels, bootstrapped 95% CIs are displayed as the envelope around each line.

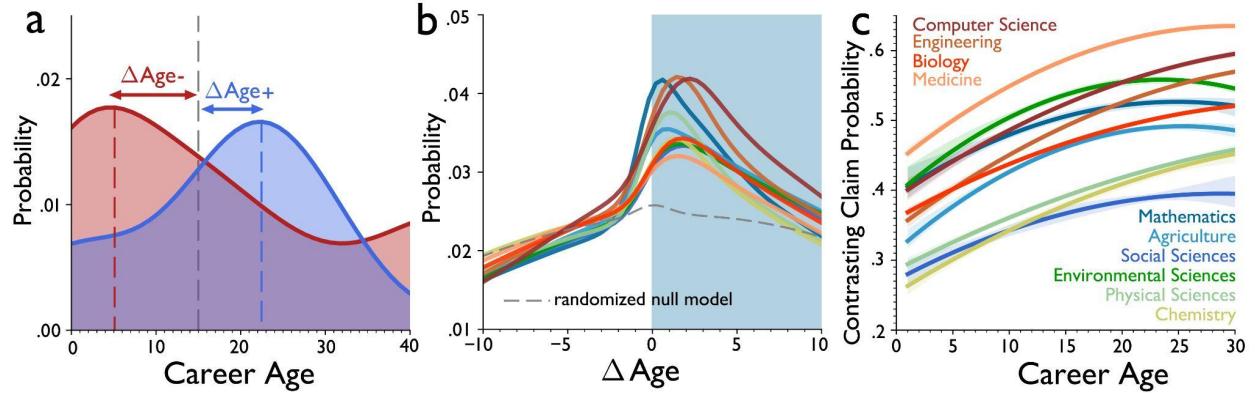
### Aging Scientists and Fields Work on Familiar Topics

While scientists adapt to the evolving frontier of science<sup>22</sup> to increase their chance of priority in discovery<sup>23,24</sup> and respond to urgent demands from society<sup>25</sup>, the tendency to do so decreases with career age. As scientists age, they tend to work on familiar topics (95% bootstrapped CIs manifest a significant ~4% increase for all fields, see Fig. 2a), lag further from the research front (95% bootstrapped CIs manifest a significant ~40% decrease for all fields, Fig. 2b and Extended Data Fig. 6), and produce work that preferentially develops rather than disrupts prior work in their fields (83.3% disruption reduction from 1 to 20 years of career age, Extended Data Fig. 7). Unsurprisingly, fields with a larger proportion of aging scientists experience a slower churn of new topics ( $\beta = -.41, p < .001$ , Fig. 2c and Table S3). Granger causality tests demonstrate that field aging predicts churn (1-year lag  $F$ -statistics = [9.67-.30.53],  $p < 0.001$  for all fields) much more powerfully than churn predicts aging ( $F$ -statistics = [0.10-17.07],  $p > 0.05$  for five fields; and only in biology greater than aging predicts churn), suggesting a possible causal influence (Table S4).

### Aging Scientists Critique New Ideas

Aging scientists do not simply ignore new ideas through their choice of references and research topics, but they actively defend against them. With maturity, scientists become more likely to contrast their work with that of younger scientists (2-sample  $K$ - $S$  statistics by field = [0.043 - 0.102],  $p < .001$  for all but medicine  $K$ - $S$  = .031,  $p < .05$ , Fig. 3a-b). Their overall rate of criticism increases over the entire career (see 95% bootstrapped CIs, Fig. 3c, Extended Data Fig.

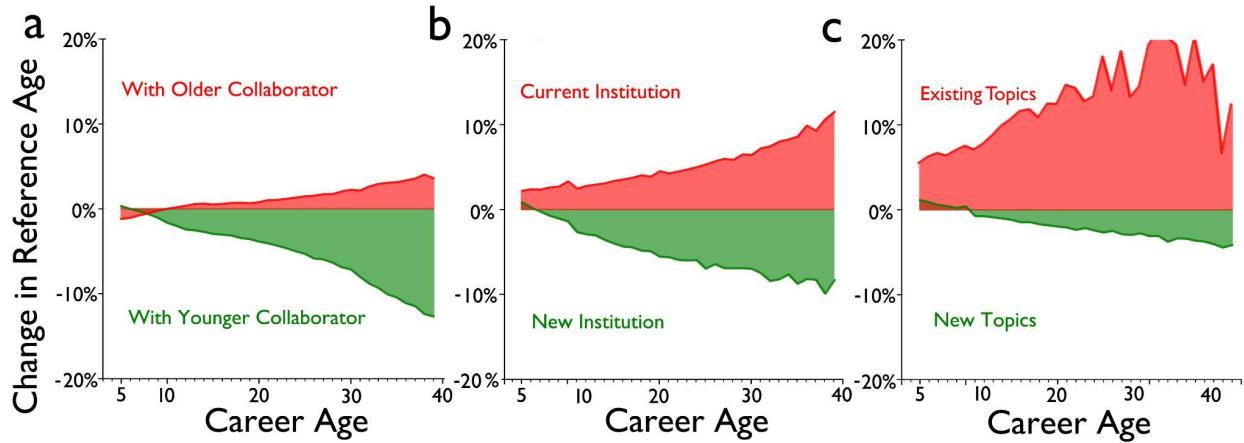
8-9). From hierarchical regression models, we learn that the strongly increasing tendency to criticize ( $\gamma_{100}=.652, p < .01$ ) is mitigated when scientists and scholars work in fields with more aging scientists ( $\gamma_{101}=-.339, p < .01$ ) as there remain fewer young scientists and less disruptive new work to criticize (Table S5). The influence of senior scientist defensiveness is likely amplified by the Matthew Effect<sup>26,27</sup> whereby established scientists attract disproportionate scientific attention.



**Fig. 3: Aging scientists are more likely to criticize young scientist's articles.** We apply a neural-network model<sup>16</sup> to classify citation functions<sup>15</sup> for 31,080,893 papers and 235,598,495 citations from 1840 to 2020 (see Methods for details). We label “contrast or compare” citations as critical and others as constructive. **a**, For exemplary papers by Higgs<sup>30</sup> on the mass-conferring boson, negative citations (blue) are disproportionately from scientists older than the authors, while positive citations (red) are from those younger (2-sample  $K-S$  statistic=.23,  $p$ -value<.01). **b**, Scholars are more likely to challenge articles published by younger scholars. The  $x$ -axis shows the career age difference underlying critical citations, calculated by subtracting the highest career age in the team receiving the critical citation from the team who cited it. The distribution of age difference skews right, deviating from the symmetric null model obtained by randomly shuffling negative citations between papers, manifesting a strong bias against younger scientists. **c**, Fraction of papers with one or more negative references increases by approximately 15% with career age over thirty years across all disciplines. Panels **a** and **b** apply the mixed-Gaussian model to fit the empirical distribution. Panels **c** use polynomial regressions to fit the data, with bootstrapped 95% confidence intervals displayed as the barely visible envelope around each line and manifesting massive, statistically significant increases in criticism as scholars age.

### How Aging Scientists Stay Young

While intellectual aging seems to be universal for individuals, it appears to be mitigated or even reversed by certain scientific behaviors. Scientists and scholars who collaborate with younger coauthors, move to new institutions and explore new topics are more likely to cite new references in comparison with peers who do not ( $\beta_{\text{younger collaborator}} = -0.32$ ,  $\beta_{\text{new institution}} = -0.08$ ,  $\beta_{\text{new topic}} = -0.17, p < .001$ , Fig. 4a-c, Table S6). Indeed, scientists working in fields of large teams such as Medicine are more likely to cite new references than those who work alone as in Mathematics (Extended Data Fig. 10).



**Fig. 4: Collaborating with younger scholars, moving to new institutions, and exploring new topics correlates with decreased reference age.** We analyzed 1,047,637 scholars who contributed to ten or more research articles over two decades or more and investigated how the reference age of their papers changes with different factors. **a**, We separated scholars according to the average career age of their collaborators as lower (green) or higher (red) than group median. **b**, We divided scholars by whether their affiliations changed (green) or remained unchanged (red). **c**, We differentiated scholars by whether their paper keywords changed (green) or remained unchanged (red). To focus on the anti-aging effect of scholars presumably having finished their Ph.D. training, we only display data after their fifth year in career. See Table S6 for OLS regression estimates of reference age against these three variables.

## Discussion

We provide the first quantitative evidence regarding the universal relationship between individual scientist aging, field-level age structure, and scientific advance. Scientists embrace new ideas in rough correspondence with Douglas Adams' cheeky rules of technology acceptance<sup>28</sup>—anything present at intellectual birth is normal; anything invented in one's epistemological adolescence is revolutionary and “you can probably get a career in it”; anything invented after academic maturity is against the natural order of things.

Our study has necessary limitations: we present the effects of scientist aging as an individual process, but science increasingly occurs within teams. We also examine average age within fields, but this only proxies for the rate of interaction between young and old scientists in the production of scientific advance. Moreover, the age of references and keywords, and the prevalence of contrasting citations and disruptively received new work each only measure singular facets of how past ideas are woven into future work. Nevertheless, together they paint a consistent portrait of scientists aging faster than science as a whole, as sketched in our initial Fig. 1a, with exaggerated consequences for fields with concentrated aging. This work demonstrates how tracking the demography of scientists can forecast areas of growth and maturity, and suggests how managing it through policy could help modulate science between crystallization and chaos.

## Methods

### Data sets

We analyze two types of datasets: (i) Publication record for 2,310,301 scholars selected from the Microsoft Academic Graph (MAG). MAG has disambiguated the names of scholars and their

institutions<sup>17</sup>. These scholars published 46,003,252 papers in 48,953 journals across 19 different scientific fields from 1800 to 2020. As data are sparse before the 1950s and after 2015, we focused on papers published 1960-2015. Results presented in the main figures show the behavior of a subset of 1,025,084 scholars, who had a career age of twenty or more years and wrote at least ten papers. This data set is verified using the Web of Science<sup>20</sup>. (ii) Citation graph containing 70,558,203 papers published during 1960-2015, which contains 611,483,153 references from or citations to these papers.

### **Identifying early-successes and late-bloomers**

Recognition arrives at different stages of the career, differentiating “early successes” from “late bloomers”<sup>26</sup> in academia. Here we define a hit paper as those among the top 1% most cited within the same field and year. We then divide scholars who own these field-definitive papers into three groups according to the publication timing of their first hit paper, including “earlier successes” (within the first five years), “late-bloomers” (after twenty years), and the rest (between early and late). Our pattern of results is not sensitive to these thresholds.

### **Mapping the scientific fields of papers and topics studied by scholars**

We used the scientific taxonomy created by the Microsoft Academic Graph team<sup>17</sup> to identify the scientific fields of papers and the topics studied by scholars. The MAG taxonomy has six levels. Level zero comprises 19 coarse-grained fields, level one lists 292 subfields, and level 2-5 contains 543,454 unique keywords (called MAG keywords hereafter). Each MAG paper is labeled by one or more keywords, which permits grouping papers into fields to map change in topics (keywords) for these fields. Using this taxonomy, we also identified the home field for 1,151,907 scholars based on where their productivity is concentrated (if over a half of their papers are published within a field) to explore how scholars change topics at different rates across fields.

### **Quantifying keyword coverage and repetition for scholars and the churn rate of fields**

To calculate the annual popular keyword coverage of a scholar, we identify the top 1% popular MAG keywords from his/her home field in the given year, and calculate the fraction of field keywords covered by the scholar’s publications. We also compare paper keywords between successive years for each scholar’s publication history to calculate the rate of keyword repetition. We calculate field “churn rate” by applying the same keyword repetition analysis to fields.

### **Detecting citation functions and calculating claim contrasting probability**

We use a deep learning method<sup>16</sup> developed based on the pioneering work by Jurgens et al.<sup>15</sup> to infer the function or role of citations based on surrounding words. This NLP model successfully classifies citations into six groups, including “Background,” “Extends,” “Uses,” “Motivation,” “Compare/Contrast,” and “Future work.” We identified the functions of 235,598,500 citations of 12,177,040 papers published during 1895 to 2020 based on their discourse context provided in the MAG data. 8.16% of total citations are labeled as “Compare/Contrast” by our model. 43% of the analyzed papers had at least one outbound “Compare/Contrast” citation and were identified as “claim contrasting” papers. We calculate the annual fraction of claim-contrasting papers for each scholar and analyze how it changes over the career. In Figure 2a, we calculated the 2-sample K-S statistic between 100 samples each from the distribution of citations from older and younger scholars. (Across 20 such sampling estimations, the statistic was always

significant). In Figure 2b, we calculated it between the randomized null model and the age distribution of critical citations for each field.

### **Analyzing anti-aging contingent on collaborators, institutions, and topics**

We select 2,310,301 scholars who contributed to ten or more research articles over two decades or longer as the focal group and analyze their configuration of collaborators, institutions, and topics throughout their career. To identify collaborator preference, we calculate the median career age of collaborators for all scholars from the focal group at a given career year, and separate papers as either “collaborating with seniors” or “collaborating with juniors” based on whether the mean career age of collaborators within that paper is higher or lower than the group average. In identifying institution preference, we keep track of all institutions for a target scholar from the focal group, and analyze whether a paper was written before (“old institutions”) or after (“new institutions”) a new institution was reported. To identify preferences in topic, we track the expanding vocabulary of MAG keywords for each scholar from the focal group and label a paper as either “old topics” or “new topics” according to whether it introduces keywords new to the researcher.

### **Estimating relationships between focal variables**

We initially estimated OLS models to explore the relationship between reference and career age. We selected 1,047,637 scholars who published over twenty years and contributed to ten or more research articles then estimated the effect for each field. The average coefficient across all fields is approximately 0.09, i.e., a one month increase of reference age for each year across the publishing career.

We then estimated Hierarchical Linear Models (HLMs) to evaluate the same relationship, but accounting for mean age in the field (Table S2). We analyze the annual statistics of reference age and career age for all 4,677 scientists who belong to the 1 million focal scientists under analysis with trackable records in the citation context datasets. We assign a scholar’s “home” if half or more papers authored by that scholar over the entire career were published within one of the eleven scientific fields under study. For each scholar in a year, we calculated the average career age of their “home” field in the same year. The HLM comprises three levels including annual statistics of a scientist (level-1), scientists (level-2), and fields (level-3). Annual reference age (*RefAge*) is the dependent variable for level-1, career age (*Age*) and average team size are the independent variables for level-1, and mean career age in field (*MeanAge*) and mean teamsizes in field (*MeanTS*) are the independent variables for level-3 as shown below. Number of subscripts designate a variable at that level of the model (e.g.,  $r_{100}$  at level-3 of the models has 3 subscripts).

Level-1

$$RefAge = \pi_0 + \pi_1 Age + \pi_2 TS + e$$

Level-2

$$\begin{aligned}\pi_0 &= \beta_{00} + r_0 \\ \pi_1 &= \beta_{10} + r_1 \\ \pi_2 &= \beta_{20} + r_2\end{aligned}$$

Level-3

$$\beta_{00} = r_{000} + r_{001}(MeanAge) + r_{002}(MeanTS) + u_{00}$$

$$\begin{aligned}\beta_{10} &= r_{100} + r_{101}(MeanAge) + r_{102}(MeanTS) + u_{10} \\ \beta_{20} &= r_{200} + u_{20}\end{aligned}$$

Our findings from this estimated model (see Table S2) confirm that holding constant mean age within field and mean team size, increased scientist age continues to manifest a statistically significant and substantial positive relationship with reference age, consistent with the OLS model above. Mean age posts an additional positive relationship, suggesting that aging scientists in older fields cite the oldest work.

We then estimated HLMs to explore the relationship between career age and critical rate, accounting for the mean age of scholars in the field (Table S3). As above, we analyzed these data for all 4,677 scientists who belong to the focal 1 million scientists under analysis with trackable records in the citation context datasets, assigning home discipline in the same way. This HLM involves three levels including annual statistics of a scientist (level-1), that scientist over the career (level-2), and that scientist's field (level-3). Annual reference age (*RefAge*) is the dependent variable for level-1, career age (*Age*) and avg. team size are the independent variables for level-1, and mean career age in field (*MeanAge*) and mean teamsize in field (*MeanTS*) are the independent variable for level-3 as shown below.

Level-1

$$CriRate = \pi_0 + \pi_1 Age + \pi_2 TS + e$$

Level-2

$$\begin{aligned}\pi_0 &= \beta_{00} + r_0 \\ \pi_1 &= \beta_{10} + r_1 \\ \pi_2 &= \beta_{20} + r_2\end{aligned}$$

Level-3

$$\begin{aligned}\beta_{00} &= r_{000} + r_{001}(MeanAge) + r_{002}(MeanTS) + u_{00} \\ \beta_{10} &= r_{100} + r_{101}(MeanAge) + r_{102}(MeanTS) + u_{10} \\ \beta_{20} &= r_{200} + u_{20}\end{aligned}$$

Our findings from this estimated model (see Table S3) confirm that holding constant mean age within field and team size, increased scientist age continues to manifest a substantial and statistically significant negative relationship with critical rate. As discussed in the main manuscript, mean age for the field posts a negative relationship with critical rate because in older, slow-moving fields there are less new discoveries to criticize. Moreover, as mean team size grows, critical rate decreases with larger teams focus on constructive advances.

We then performed Granger Causality tests to explore whether average age in field more likely anticipated churn rate, as we suspect, or churn rate better predicts average scholar age in field. For all 11 fields in the Microsoft Academic Graph, we calculate churn rate in the field and old proportion (percentage of scholars whose career year > 20 in each field) year by year, then predict each other with one and two year lags in both directions. Next, we applied Granger causality tests on subfield churn rate and subfield old proportion, as detailed Table S4. When we alternatively weighted scholars according to their number of annual papers, the pattern of findings remained constant.

Finally, we estimated OLS models tracing the relationship between churn rate on old proportion across 271 subfields and eleven fields (see Table S5). Churn rate is negatively associated with old scientist proportion, both for unweighted (model-1) and weighted data (model-2 in which scientists are weighted according to their number of published papers.)

**Table S1. OLS models regressing reference age on career age.** Ordinary Least Square (OLS) regression coefficients for the data underlying Figure 1 for the 1,047,637 scholars who published over twenty years and contributed to ten or more research articles. The average coefficient across all fields is approximately 0.09, i.e., a one month increase of reference age for each year across the publishing career.

| Field                  | Intercept | Coefficient | R <sup>2</sup> |
|------------------------|-----------|-------------|----------------|
| Mathematics            | 12.699*** | 0.143***    | 0.986          |
| Agriculture            | 10.771*** | 0.075***    | 0.952          |
| Social Sciences        | 9.938***  | 0.102***    | 0.971          |
| Business & Management  | 9.974***  | 0.107***    | 0.983          |
| Environmental Sciences | 9.797***  | 0.078***    | 0.973          |
| Chemistry              | 8.778***  | 0.110***    | 0.978          |
| Physical Sciences      | 8.762***  | 0.092***    | 0.979          |
| Engineering            | 8.310***  | 0.089***    | 0.976          |
| Biology                | 8.443***  | 0.072***    | 0.957          |
| Medicine               | 8.554***  | 0.045***    | 0.895          |
| Computer Science       | 7.119***  | 0.096***    | 0.957          |

\* p-value < 0.05; \*\* p-value < 0.01; \*\*\* p-value < 1.0 × 10<sup>-40</sup>

**Table S2. HLM model regressing reference age on career age.** HLM comprising three levels for the annual statistics of a scientist (level-1), that scientist over the lifecourse (level-2), and the scientist's field (level-3). Annual reference age (*RefAge*) is the dependent variable, with career age (Age) and team size (TS) independent variables in level-1. Mean career age in the field (MEAN\_AGE) and mean team size in the field (MEAN\_TS) are the independent variables in level-3 as shown below, revealing that as mean age goes up, references get older, but as mean team size increases, references become more recent. Controlling for these higher-level factors, scientist career age retains by the most significant and substantial relationship with reference age.

**Final estimation of fixed effects (with robust standard errors)**

| Fixed Effect               | Coefficient   | Standard error | T-ratio | Approx. d.f. | p-value |
|----------------------------|---------------|----------------|---------|--------------|---------|
| For INTRCPT1, $\pi_0$      |               |                |         |              |         |
| For INTRCPT2, $\beta_{00}$ |               |                |         |              |         |
| INTRCPT3, $\gamma_{000}$   | 9.515         | 0.425          | 22.404  | 9            | <0.001  |
| MEAN_AGE, $\gamma_{001}$   | 0.401         | 0.390          | 0.997   | 9            | 0.211   |
| MEAN_TS, $\gamma_{002}$    | 0.389         | 0.390          | 0.997   | 9            | 0.176   |
| For AGE slope, $\pi_1$     |               |                |         |              |         |
| For INTRCPT2, $\beta_{01}$ |               |                |         |              |         |
| INTRCPT3, $\gamma_{100}$   | <b>0.126</b>  | 0.017          | 7.630   | 9            | <0.001  |
| MEAN_AGE, $\gamma_{101}$   | <b>0.016</b>  | 0.010          | 1.488   | 9            | 0.171   |
| MEAN_TS, $\gamma_{102}$    | <b>-0.002</b> | 0.009          | -0.246  | 9            | 0.811   |
| For TS slope, $\pi_2$      |               |                |         |              |         |
| For INTRCPT2, $\beta_{02}$ |               |                |         |              |         |
| INYRCPT3, $\gamma_{200}$   | -0.0389       | 0.011          | -3.505  | 11           | 0.006   |

**Final estimation of level-1 and level-2 variance components**

| Random Effect    | Standard Deviation | Variance Component | d.f. | $\chi^2$  | p-value |
|------------------|--------------------|--------------------|------|-----------|---------|
| INTRCPT1, $r_0$  | 2.022              | 4.088              | 4642 | 44192.832 | <0.001  |
| Age slope, $r_1$ | 0.191              | 0.036              | 4642 | 9625.648  | <0.001  |
| TS slope, $r_2$  | 0.030              | 0.000              | 4642 | 5169.017  | <0.001  |
| level-1, e       | 2.427              | 5.891              |      |           |         |

Note: The chi-square statistics reported above are based on only 4654 of 4677 units that had sufficient data for computation. Fixed effects and variance components are based on all the data

**Final estimation of level-3 variance components**

| Random Effect               | Standard Deviation | Variance Component | d.f. | $\chi^2$ | p-value |
|-----------------------------|--------------------|--------------------|------|----------|---------|
| INTRCPT1/INTRCPT2, $u_{00}$ | 1.427              | 2.066              | 9    | 1741.678 | <0.001  |
| AGE/INTRCPT2, $u_{10}$      | 0.049              | 0.049              | 9    | 111.190  | <0.001  |
| TS/INTRCPT2, $u_{20}$       | 0.033              | 0.033              | 11   | 26.128   | <0.001  |

**Table S3. OLS Models regressing Churn Rate on Old Proportion.** Churn Rate relationship with old scientist proportion for unweighted (model-1) and weighed data (model-2; scientists are weighted according to their number of published papers.)

| Independent Variables    | Model 1: unweighted |             | Model 2: weighted by paper productivity |             |
|--------------------------|---------------------|-------------|---|-------------|
|                          | Intercept           | Coefficient | Intercept                               | Coefficient |
| <b>Mathematics</b>       | 0.184***            | -0.350***   | 0.186***                                | -0.282***   |
| <b>Geology</b>           | 0.197***            | -0.359***   | 0.194***                                | -0.266***   |
| <b>Economics</b>         | 0.136***            | -0.188***   | 0.135***                                | -0.155***   |
| <b>Psychology</b>        | 0.167***            | -0.263***   | 0.164***                                | -0.190***   |
| <b>Materials Science</b> | 0.165***            | -0.390***   | 0.164***                                | -0.282***   |
| <b>Physics</b>           | 0.120***            | -0.157***   | 0.117***                                | -0.098***   |
| <b>Chemistry</b>         | 0.128***            | -0.247***   | 0.133***                                | -0.187***   |
| <b>Biology</b>           | 0.165***            | -0.344***   | 0.167***                                | -0.254***   |
| <b>Medicine</b>          | 0.164***            | -0.409***   | 0.156***                                | -0.243***   |
| <b>Engineering</b>       | 0.179***            | -0.407***   | 0.176***                                | -0.311***   |
| <b>Computer Science</b>  | 0.202***            | -0.509***   | 0.202***                                | -0.363***   |

\* p-value < 0.05; \*\* p-value < 0.01; \*\*\* p-value < 0.001

**Table S4. Granger Causality Test on Churn Rate and Average Age in Field.** Predictive relationships between churn rate in the field and old proportion (percentage of scholars whose career year > 20 in each field) year by year with one and two year lags in both directions (e.g., Churn Rate predicting Old Proportion; Old Proportion predicting Churn Rate). Granger causality tests on subfield churn rate and subfield old proportion suggest that old proportion is much more likely to predict churn rate than the converse.

| Independent Variables    | Old Proportion Predicts Churn Rate |             | Churn Rate Predicts Old Proportion |             |
|--------------------------|------------------------------------|-------------|------------------------------------|-------------|
|                          | SSR F-test                         |             | SSR F-test                         |             |
|                          | lags=1 year                        | lags=2 year | lags=1 year                        | lags=2 year |
| <b>Mathematics</b>       | 9.67***                            | 4.76**      | 8.17***                            | 5.55***     |
| <b>Geology</b>           | 16.76***                           | 7.74***     | 0.36                               | 4.69**      |
| <b>Economics</b>         | 11.45***                           | 9.16***     | 0.61                               | 1.59        |
| <b>Psychology</b>        | 18.66***                           | 6.65***     | 0.32                               | 0.96        |
| <b>Materials science</b> | 14.29***                           | 4.34**      | 5.49**                             | 3.65**      |
| <b>Physics</b>           | 9.08***                            | 2.19        | 4.05**                             | 2.27        |
| <b>Chemistry</b>         | 11.09***                           | 2.45*       | 1.29                               | 0.52        |
| <b>Biology</b>           | 13.82***                           | 4.01**      | 17.07***                           | 8.31***     |
| <b>Medicine</b>          | 17.27***                           | 5.07***     | 0.1                                | 0.57        |
| <b>Engineering</b>       | 16.38***                           | 4.98**      | 2.84*                              | 1.8         |
| <b>Computer science</b>  | 30.53***                           | 13.33***    | 4.19**                             | 0.75        |

\* p-value < 0.05; \*\* p-value < 0.01; \*\*\* p-value < 0.001

**Table S5. HLM model regressing critical rate on career age.** HLM modeling three levels including annual statistics of a scientist (level-1), that scientist over the lifecourse (level-2), and the scientist's field (level-3). Annual critical rate (*CriRate*) is the dependent variable with career age (Age) and team size (TS) independent variables in level-1. Mean career age in the field (MEAN\_AGE) and mean team size in the field (MEAN\_TS) are the independent variables in level-3 as shown below, revealing that as mean age and team size go up, the critical rate goes down, with less new work to criticize and teams collaborating around constructive new work. Controlling for these higher-level factors, scientist career age retains a significant and substantial relationship with critical rate.

**Final estimation of fixed effects (with robust standard errors)**

| Fixed Effect               | Coefficient   | Standard error | T-ratio | Approx. d.f. | p-value |
|----------------------------|---------------|----------------|---------|--------------|---------|
| For INTRCPT1, $\pi_0$      |               |                |         |              |         |
| For INTRCPT2, $\beta_{00}$ |               |                |         |              |         |
| INTRCPT3, $\gamma_{000}$   | 50.247        | 1.470          | 34.177  | 9            | <0.001  |
| MEAN_AGE, $\gamma_{001}$   | -3.259        | 1.251          | -2.606  | 9            | 0.029   |
| MEAN_TS, $\gamma_{002}$    | -1.810        | 0.763          | -2.373  | 9            | 0.042   |
| For AGE slope, $\pi_1$     |               |                |         |              |         |
| For INTRCPT2, $\beta_{01}$ |               |                |         |              |         |
| INTRCPT3, $\gamma_{100}$   | <b>0.652</b>  | 0.061          | 10.704  | 9            | <0.001  |
| MEAN_AGE, $\gamma_{101}$   | <b>-0.338</b> | 0.060          | -5.610  | 9            | <0.001  |
| MEAN_TS, $\gamma_{102}$    | <b>-0.142</b> | 0.044          | -3.215  | 9            | 0.011   |
| For TS slope, $\pi_2$      |               |                |         |              |         |
| For INTRCPT2, $\beta_{02}$ |               |                |         |              |         |
| INYRCPT3, $\gamma_{200}$   | 0.1617        | 0.088          | 1.836   | 11           | 0.093   |

**Final estimation of level-1 and level-2 variance components**

| Random Effect    | Standard Deviation | Variance Component | d.f. | $\chi^2$  | p-value |
|------------------|--------------------|--------------------|------|-----------|---------|
| INTRCPT1, $r_0$  | 12.790             | 163.592            | 4642 | 12024.477 | <0.001  |
| Age slope, $r_1$ | 1.342              | 1.802              | 4642 | 6073.891  | <0.001  |
| TS slope, $r_2$  | 1.323              | 1.752              | 4642 | 5118.783  | <0.001  |
| level-1, e       | 235.576            | 1265.681           |      |           |         |

Note: The chi-square statistics reported above are based on only 4654 of 4677 units that had sufficient data for computation. Fixed effects and variance components are based on all the data

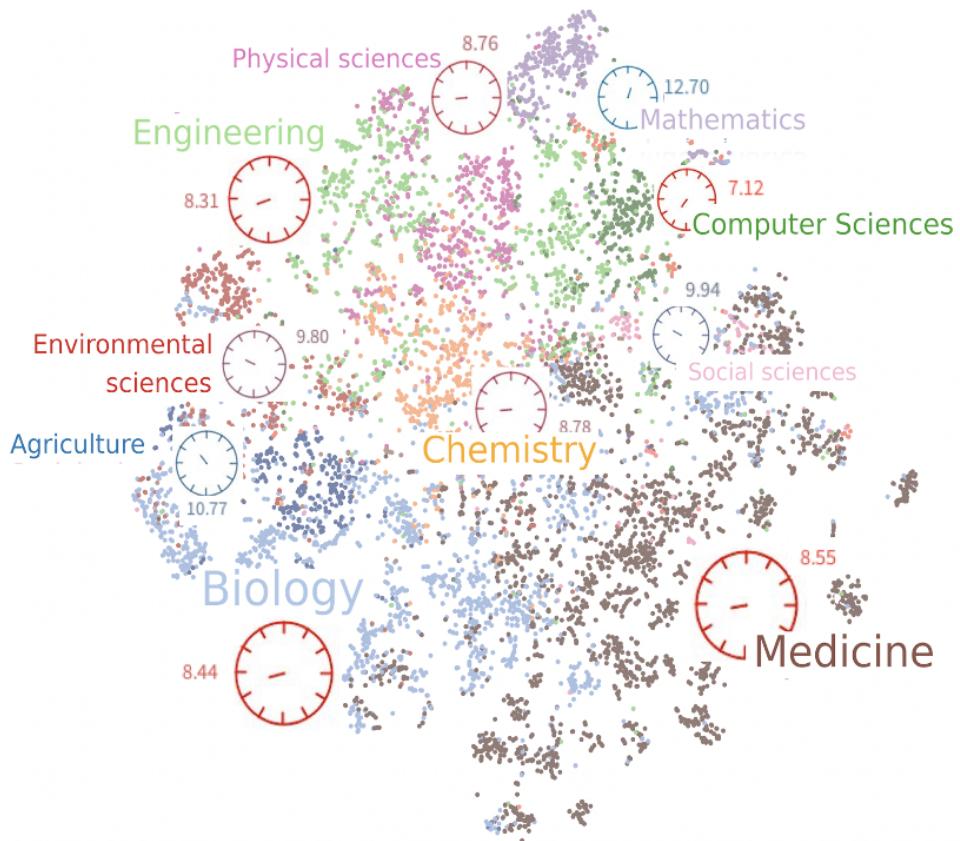
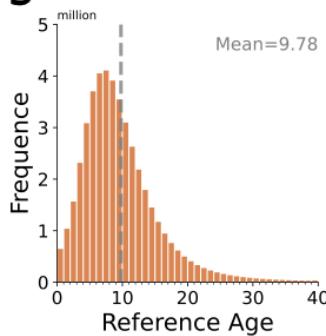
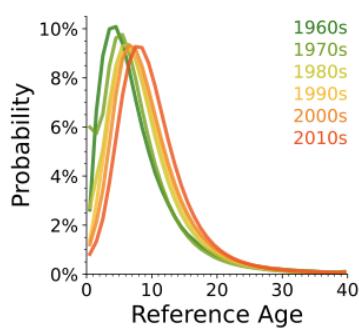
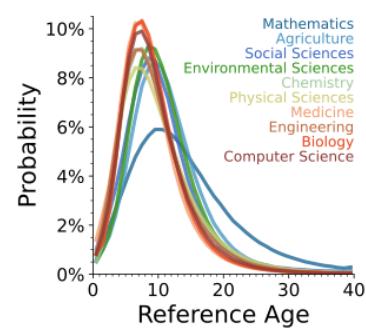
**Final estimation of level-3 variance components**

| Random Effect               | Standard Deviation | Variance Component | d.f. | $\chi^2$ | p-value |
|-----------------------------|--------------------|--------------------|------|----------|---------|
| INTRCPT1/INTRCPT2, $u_{00}$ | 4.779              | 22.842             | 9    | 427.209  | <0.001  |
| AGE/INTRCPT2, $u_{10}$      | 0.131              | 0.0171             | 9    | 23.638   | 0.005   |
| TS/INTRCPT2, $u_{20}$       | 0.137              | 0.019              | 11   | 19.727   | 0.049   |

**Table S6. OLS models regressing reference age on career age, controlling for moderating influences.** We selected 1,047,637 scholars who published over twenty years and contributed to ten or more research articles (same as Table S1). We then construct a sample of 5,426,422 data points covering each scholar in each publishing year. For a scholar-publishing year, we calculate the average reference age and career age of that scholar, and constructed three dummy variables to describe their behavior, including whether they joined a new institution, published on a new topic, or collaborated with scholars younger than themselves. We regressed reference age on career age and these dummy variables using Ordinary Least Square (OLS) regression and reported the estimated coefficients and significances below.

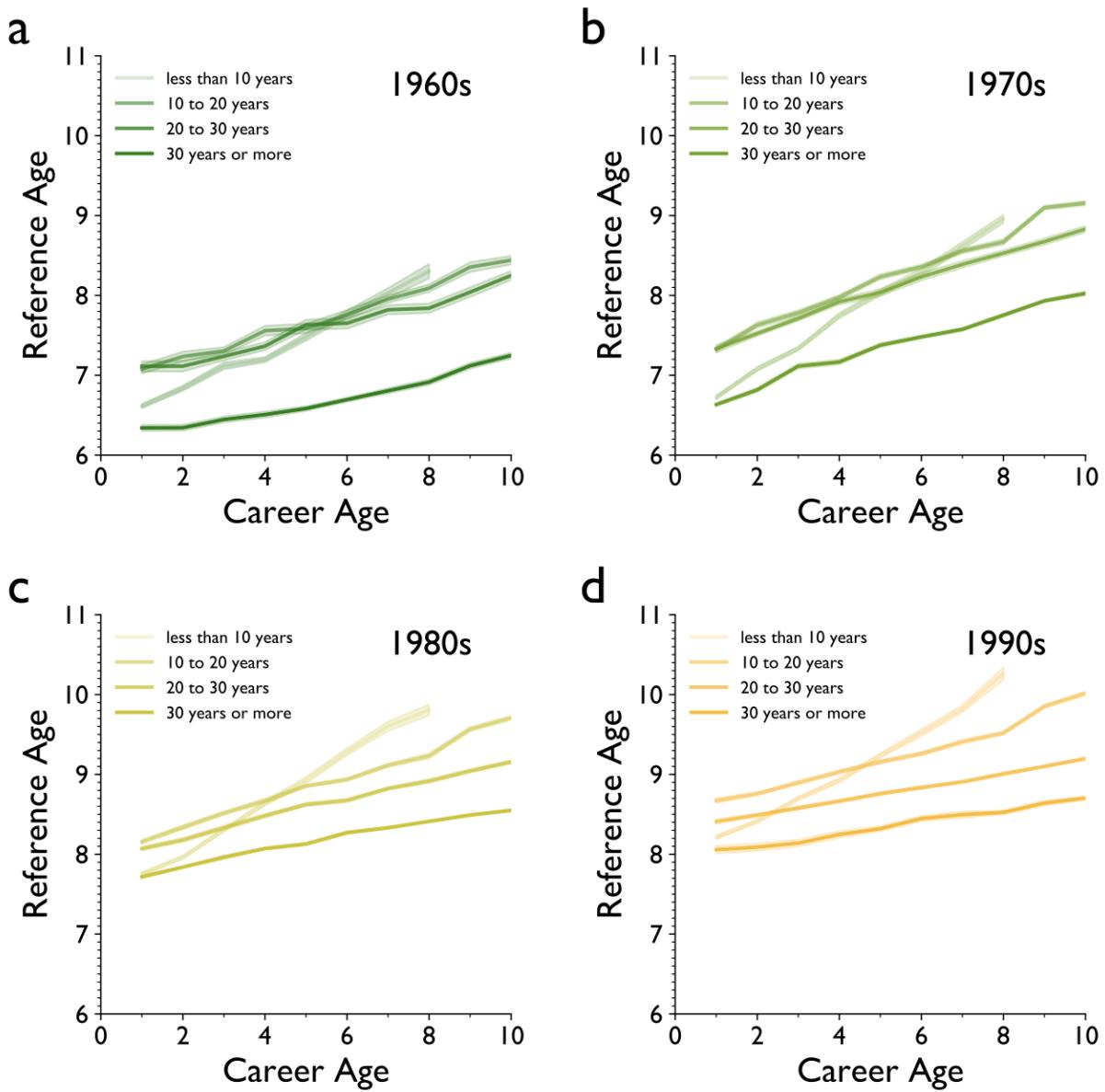
| Independent Variables               | Dependent Variable: Reference Age |                |
|-------------------------------------|-----------------------------------|----------------|
|                                     | Coefficient                       | Standard Error |
| Intercept                           | 8.6596***                         | 0.004          |
| Career Age                          | 0.0511***                         | <0.001         |
| Joining in new institution          | -0.0823***                        | 0.01           |
| Expanding to new topics             | -0.1699***                        | 0.012          |
| Collaborating with younger scholars | -0.3249***                        | 0.004          |

\*  $p < 0.05$ ; \*\*  $p$ -value  $< 0.01$ ; \*\*\*  $p$ -value  $< 0.001$ ,  $N = 5,426,422$

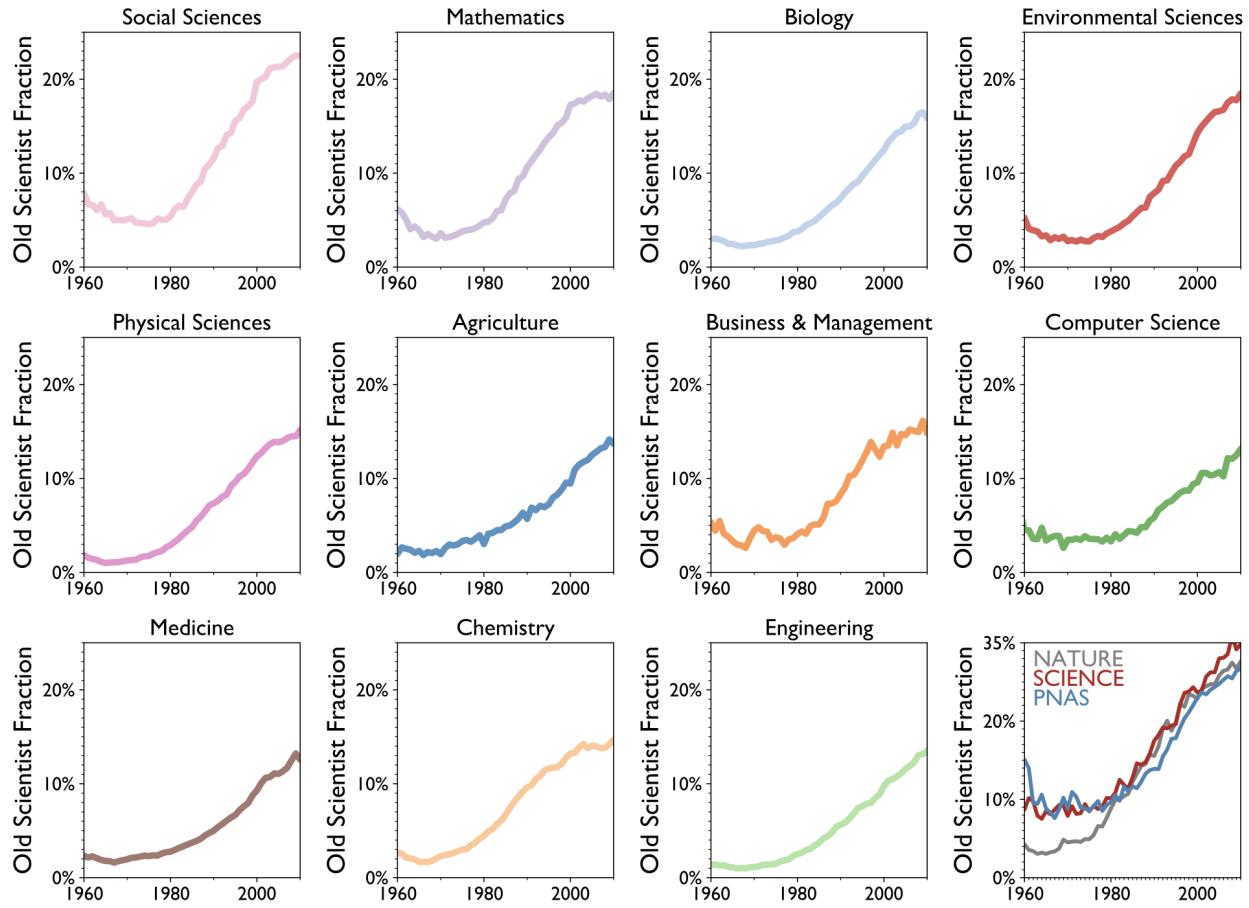
**a****b****c****d**

**Extended Data Fig. 1. The innovation clocks of science and distributions of reference age.** **a**, Dots in the background show 10,237 Web of Science journals, colored by scientific field. Journals frequently cited together are placed near each other. This is achieved by learning the vector representation of journals from their co-citation in paper references (Lin et al. 2021). Here we selected 1,962,583 papers published in 2015 to construct this embedding space of journals. The clock within each field shows the average reference age of papers using the “hour hand”, ranging from 12.7 years (std. 9.52) for Mathematics (the most rapidly aging with lowest pressure for innovation) to 7.12 years (std. 5.57) for computer science (the most slowly aging, or highest pressure for innovation). Reference ages across fields are calculated from 79 million papers published over a century. Clock

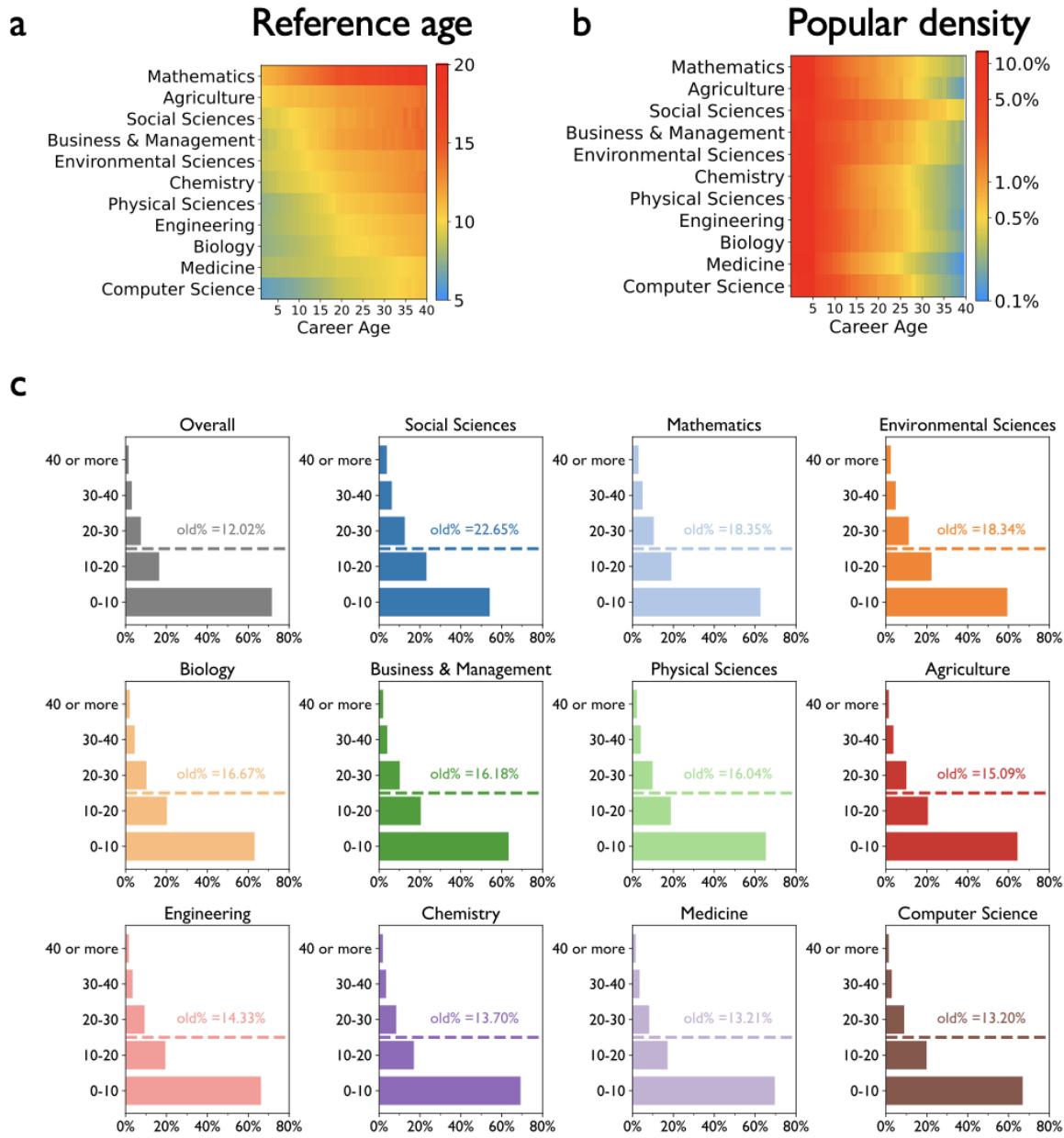
color reflects the aging pattern with size proportional to field productivity. **b**, Distribution of reference age across the 79 million analyzed MAG papers, for which the average value is 9.78 years (std. 6.88). **c**, Distribution of reference age across six decades from the 1960s to 2010s. **d**, Distribution of reference age across eleven scientific fields.



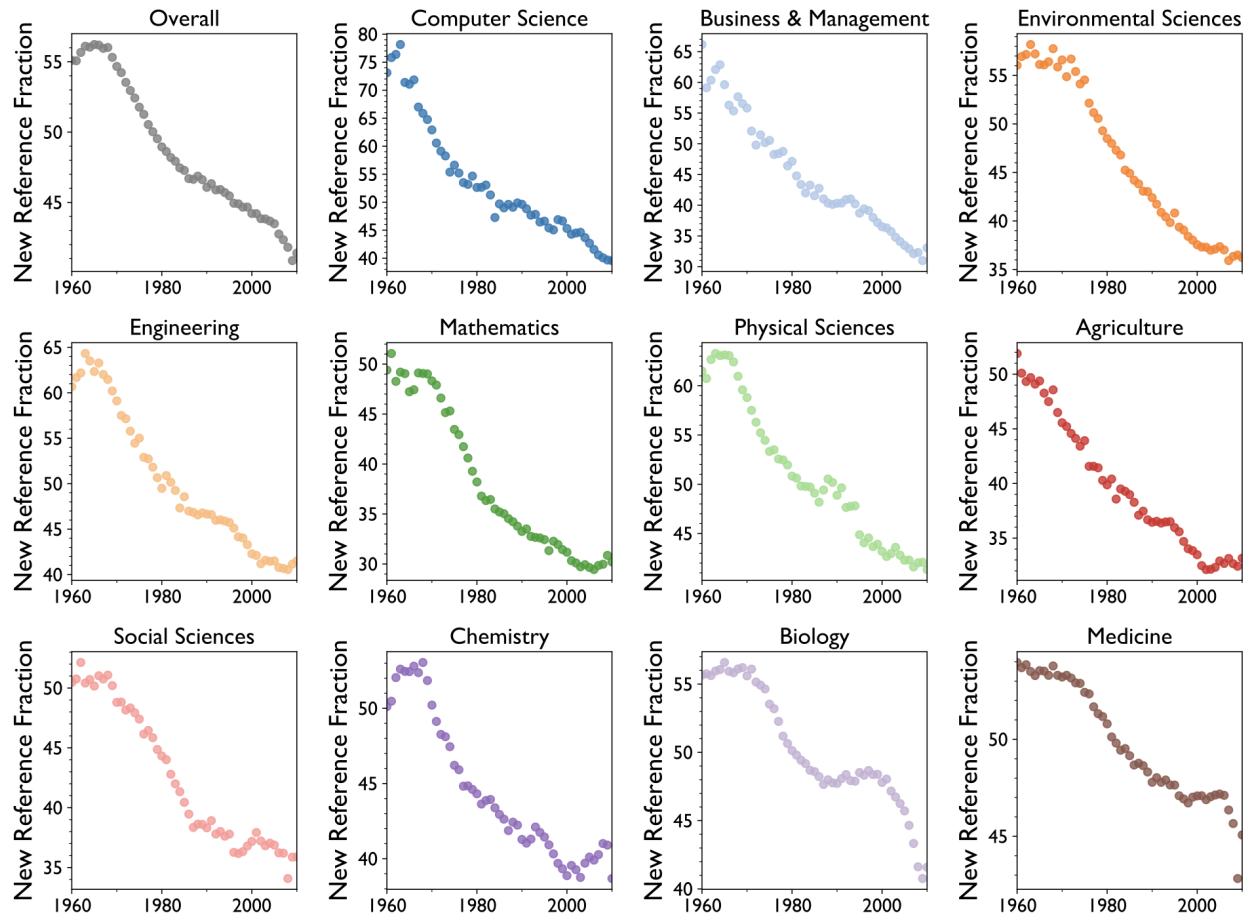
**Extended Data Fig. 2. Increase of reference age over individual careers across scientists of distinct career length in different historical times.** Increase of reference age with career age is consistent for scientists who started publishing at different historical times, including the 1960s (a), 1970s (b), 1980s (c), and 1990s (d). For each panel, we divide scientists into four groups by career length, including less than ten years, ten to twenty years, twenty to thirty years, and thirty years or longer. In all panels, bootstrapped 95% confidence intervals are displayed as the envelope around each line but are barely invisible. These demonstrate the substantial and statistical significance of decreasing disruption with age.



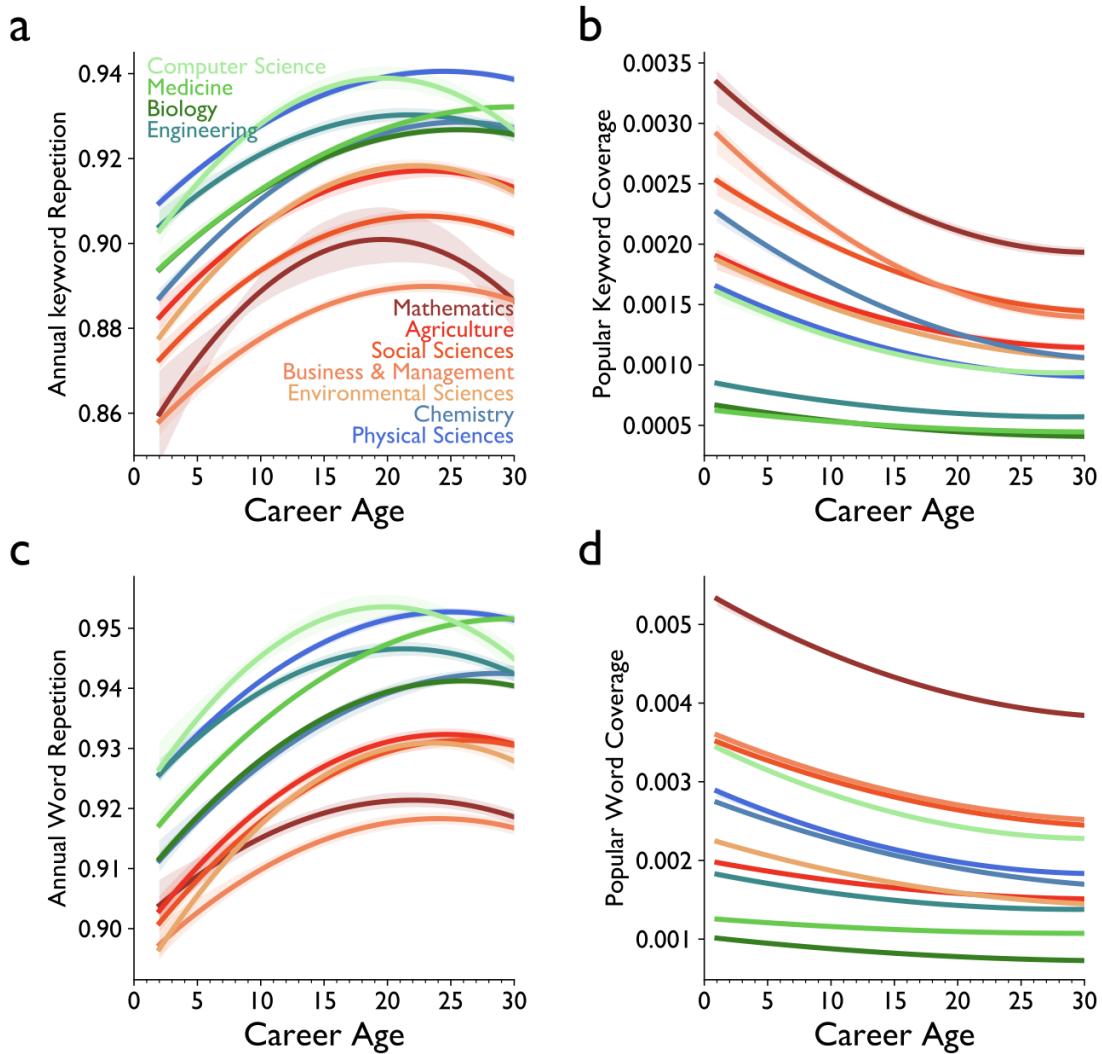
**Extended Data Fig. 3. Increasing fraction of old scientists across fields.** We calculate the annual fraction of old scientists with a career age twenty years or longer and plot it over historical time. This fraction increases over time across fields. Social Sciences have seen the fastest increase in the old scientist fraction from below 10% to over 20%. In comparison, Engineering has the slowest increase from 2% to 15%. This finding verifies previous observations regarding the aging scientific workforce<sup>1</sup> but at a much larger scale of the scientific population and over a longer time span. In the bottom right we calculate the annual fraction of old scientists who publish in three prestigious journals (*Nature*, *Science*, and *PNAS*), and plot it against historical time. This fraction increases dramatically over time across all three journals more rapidly than for science as a whole.



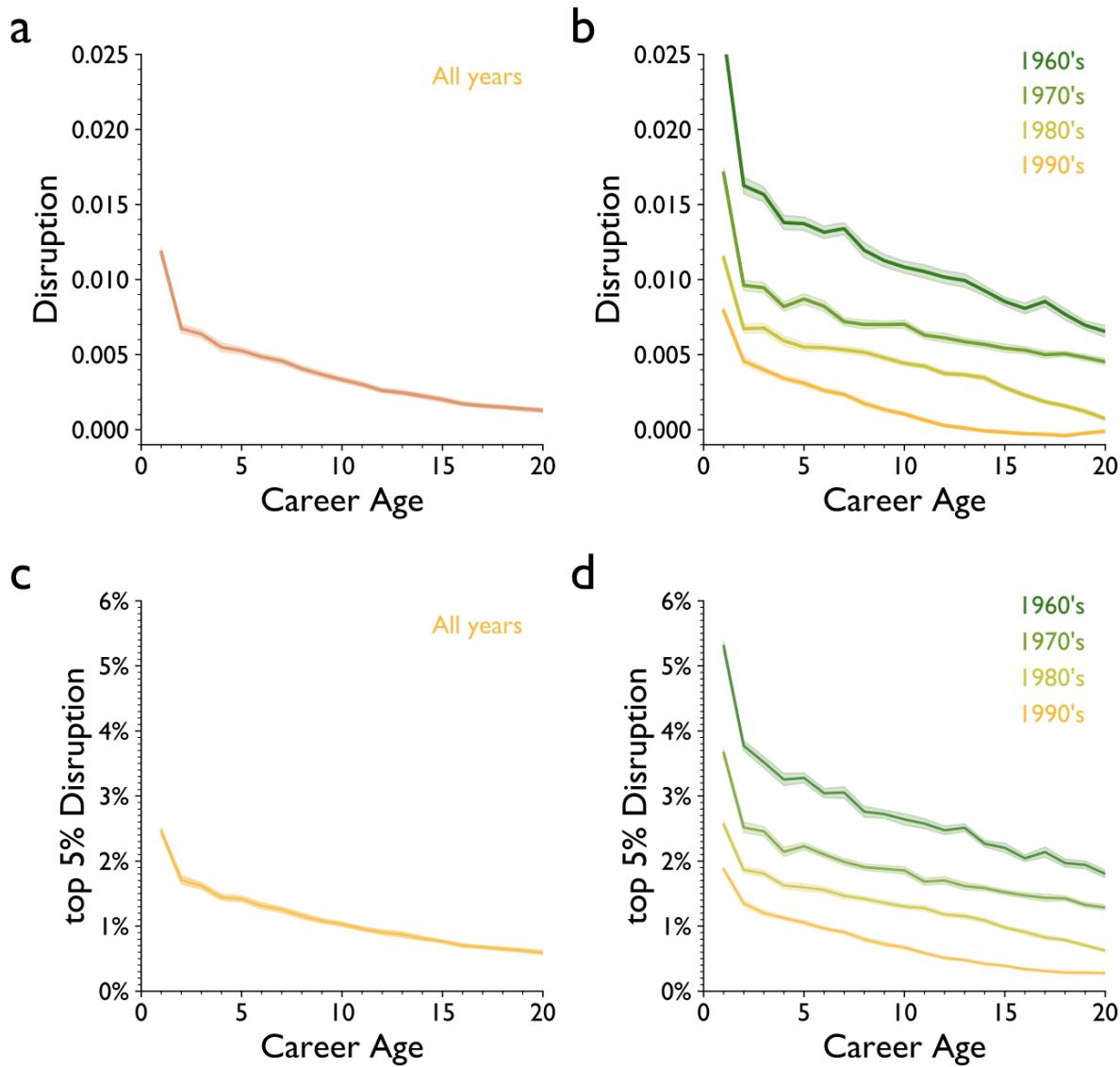
**Extended Data Fig. 4. Association between reference and career age across fields; and comparison of age structure across fields.** Heatmaps of reference age (a) and population density (b) tracking their distribution across fields. This provides a more systematic comparison between fields to validate conclusions from Figure 1, in which only the average reference age was analyzed and displayed. The transition of blue into red from left to right in (a) shows that aging scientists tend to cite old papers and the increase in red cells from low to high reveals that fields of aging scientists, as confirmed in (b), use more canonical literature. (c), We plot the distribution of scientists by five career-age groups in 2015: 0-10 years, 10-20 years, 20-30 years, 30-40 years, and 40 years or more. We calculate and display the fraction of old scientists with a career age of twenty years or older. Social Sciences have the highest fraction of old scientists, nearly twice as much as that in Medicine or Computer Science (23% vs 13%).



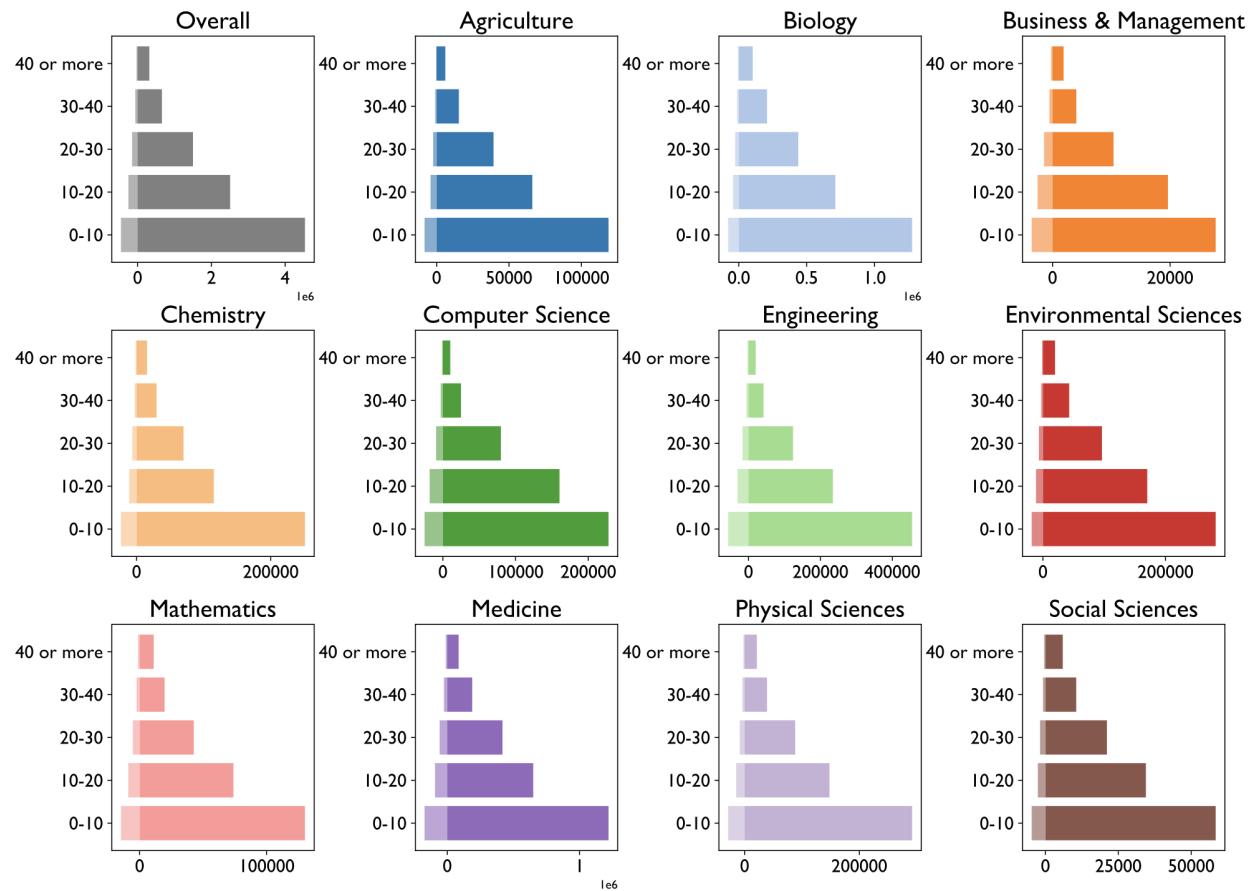
**Extended Data Fig. 5. Decreasing fraction of new references across fields.** Derek de Solla Price proposed studying the fraction of new references published within five years to quantify the rate at which new ideas are drawn upon to advance science<sup>10</sup>. Previous studies observed a decrease in Price index for the entire universe of science<sup>11</sup>. Here we verify and display the universality of decrease in new references across fields. We experimented with five and ten years as the criteria for new references and found both results display consistent decrease. We show the result for ten years, which necessarily manifests a somewhat more stable trend due to the longer time window for new references.



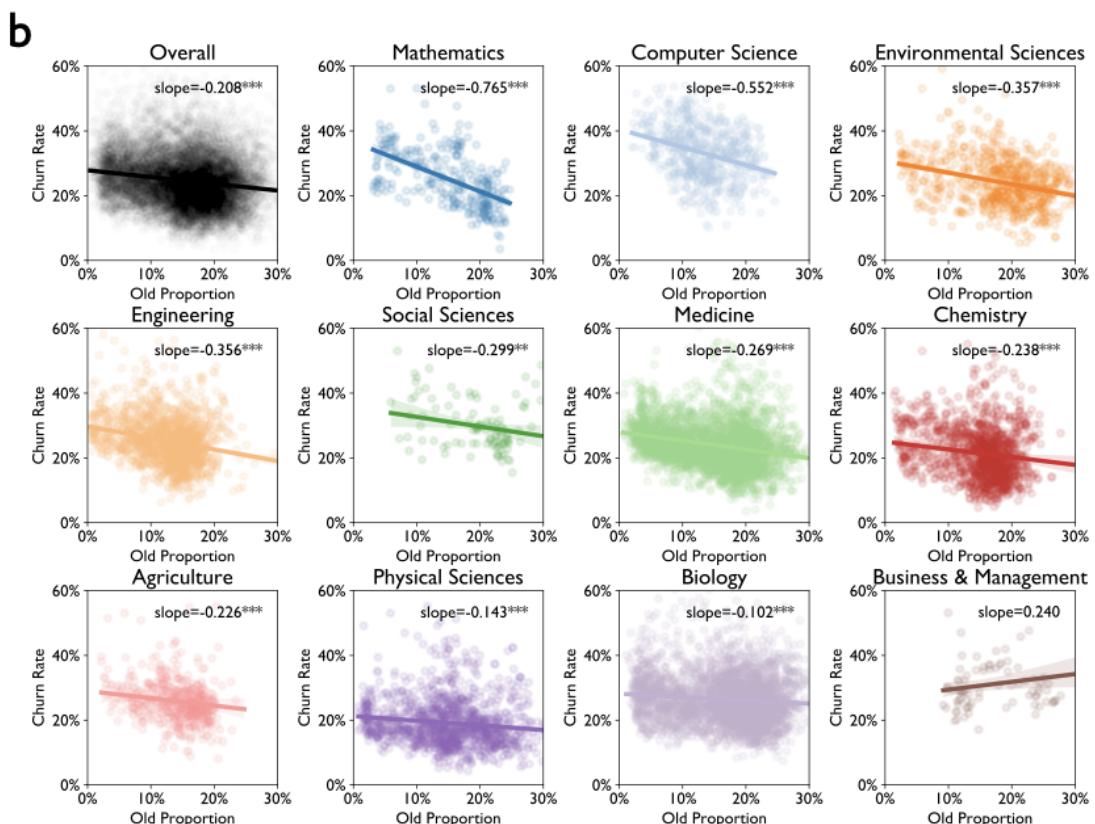
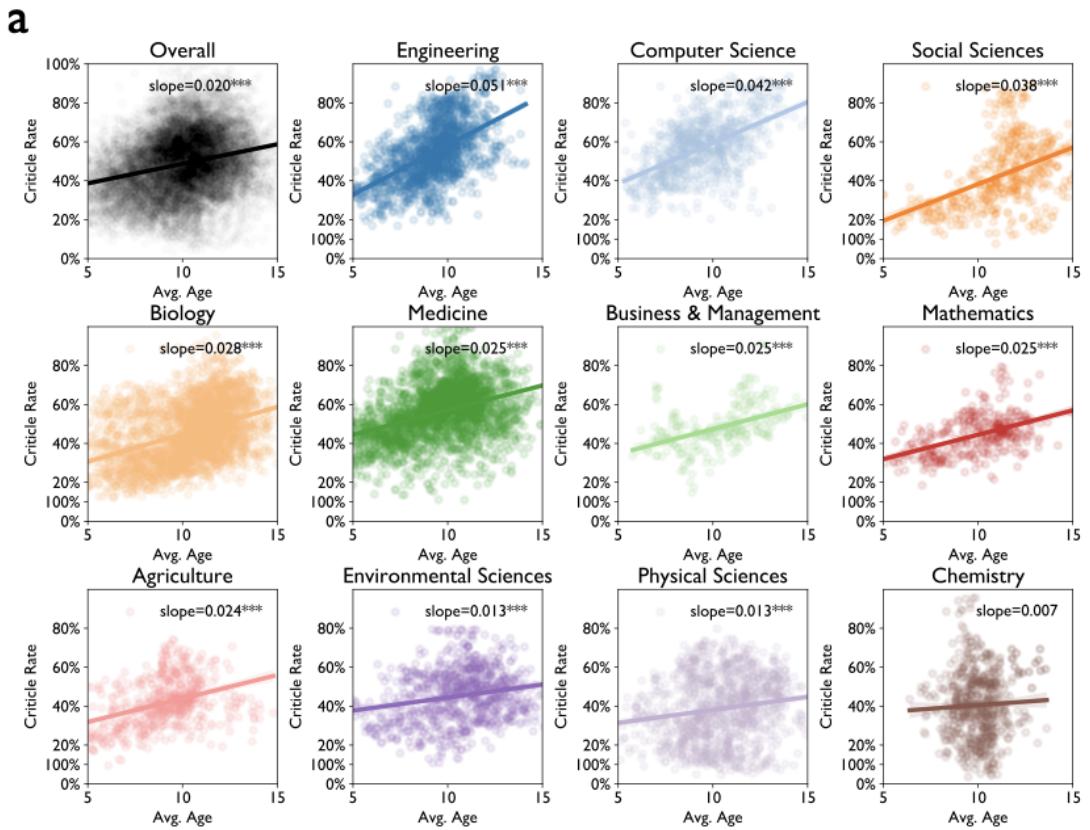
**Extended Data Fig. 6. Increase of self-similarity in research topics and increase in distance to field front over career.** We quantify self-similarity for scholars in successive years with the Wasserstein distance between paper keywords. Self-similarity increases with career age for both the scientific taxonomy provided by MAG (**a**) and the keywords we extract from abstracts using term frequency ( $\geq 2$ ) (**c**) (see Methods for details). We quantify the “scientific front” for each field in a given year by selecting the top 1% keywords across all published articles. The overlap between individual keywords and the most popular keywords in the field decreases with career age for both the scientific taxonomy provided by MAG (**b**) and the keywords we extract from abstracts using term frequency ( $\geq 2$ ) (**d**). In all panels, bootstrapped 95% confidence intervals are displayed as the envelope around each line but are barely invisible. These demonstrate the substantial and statistical significance of growing self-similarity and distance from the scientific frontier with age.



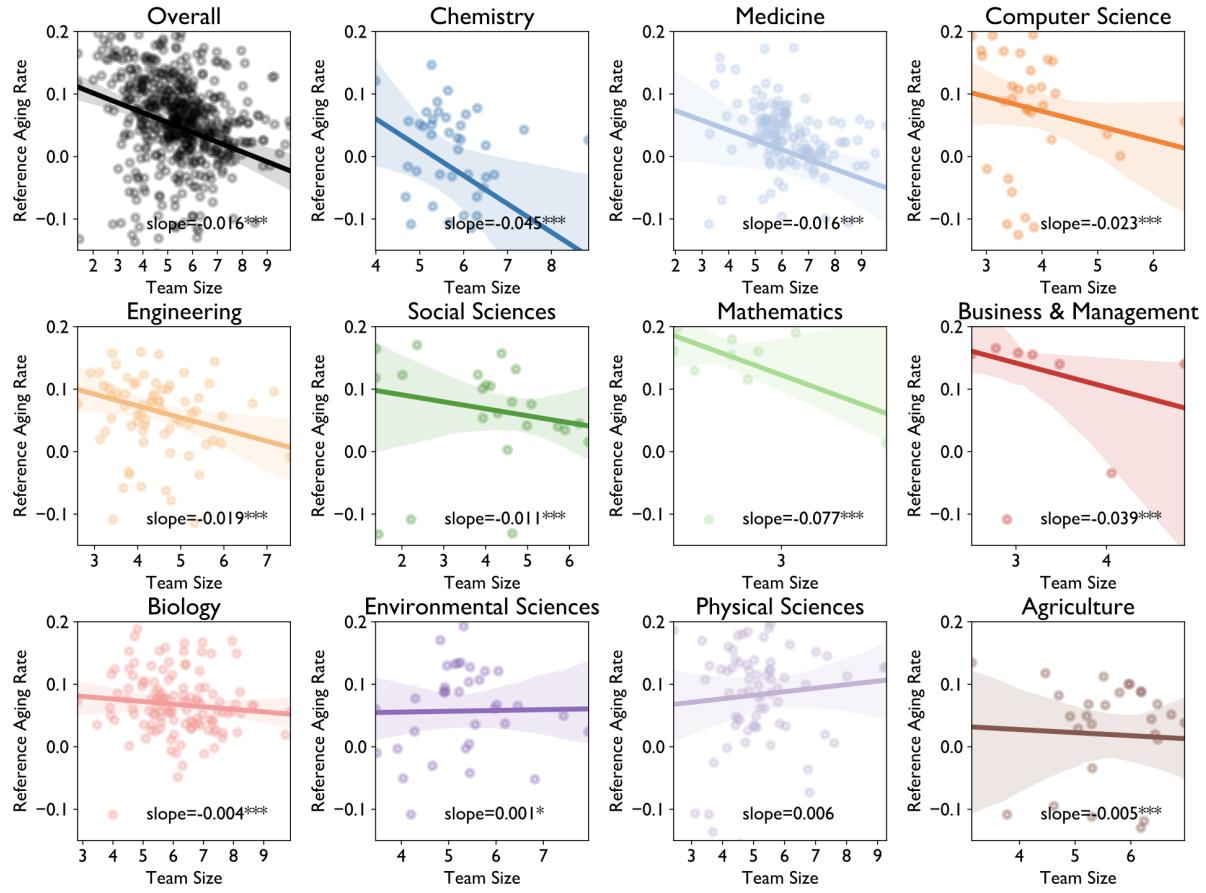
**Extended Data Fig. 7. Decrease in disruption score over individual academic careers.** The annual average of disruption across published papers decreases with career age for scientists in all years (**A**) and across historical times including the 1960s, 1970s, 1980s, and 1990s (**B**). Probability of publishing a top 5% disruptive paper decreases with career age for scientists in all years (**C**) and across historical times (**D**). Scientists who start publishing after 2000 are not shown as their career length is less than twenty years and thus do not meet our selection criteria designed to analyze long-term career performance. In all panels, bootstrapped 95% confidence intervals are displayed as the envelope around each line but are barely invisible. These demonstrate the substantial and statistical significance of decreasing disruption with age.



**Extended Data Fig. 8. Distribution of ‘Critical’ citation (left bars) and ‘Uses’ citation (right bars) for different age groups among the top-level twelve fields.** For each field, we collect the set of scholars who published papers in 2015. We group them into five age groups and graph the percentage of the citation criticism and citation use.



**Extended Data Fig. 9. Relationship between average age of field with critical rate; and churn rate for fields decreases with increasing old proportion.** **a**, Critical Rate is the critical rate of citation in each year for a subfield (271 subfields in total). Average age is the average scholar age from first publication. The first panel is the relation between Average Age across all fields with Critical Rate for the entire dataset. Lines represent OLS regressions of critical rate on average age, with 95% bootstrapped confidence intervals on the slope, demonstrating a consistent and statistically significant ( $^{***}p<.001$ ) upward trend for all fields. **b**, Churn Rate is the changing rate of keywords among the top 10% in focal year  $t$  and the immediately prior year  $t-1$ . Old proportion is the percentage of scholars whose career age is twenty years or more. Here, we weight scholars by their frequency of publication in a year. We divide the data into twelve groups according to the field of papers and draw the relation between churn rate and old proportion in different fields. The first panel is the relation between churn rate and old proportion across the entire dataset. Lines also represent OLS regressions of churn rate on average age, with 95% bootstrapped confidence intervals on the slope, demonstrating a consistent and statistically significant ( $^{***}p<.001$ ) downward trend for all fields except for business and management, which manifests no significant trend.



**Extended Data Fig. 10. Subfields with larger teams more likely cite more recent literature.** Across the analyzed 271 subfields, further organized into eleven fields, subfields of larger teams are more likely to cite more recent literature. We calculate reference aging rate as the annual increase in reference age with career age. The upper left panel details the relationship for all fields in the dataset. Lines represent OLS regressions of reference aging rate on team size, with 95% bootstrapped confidence intervals on the slope, demonstrating a consistent and statistically significant ( $^{***}p<.001$ ) downward trend for all fields except the physical sciences, which demonstrate no significant trend, and the environmental sciences where the relationship posts a mildly positive effect.

## **Acknowledgments**

The authors thank members of the University of Chicago's Knowledge Lab and the Swarma Club (Beijing) for helpful discussions. We are grateful for support from AFOSR #FA9550-19-1-0354 (J.E.), NSF #1829366 and #1800956 (J.E.), DARPA #HR00111820006 (J.E.), the Kaifeng Foundation for Young Scholars Award (H.C.), the Pitt Cyber Institute (L.W.) and Richard King Mellon Foundation (L.W.).

## References

1. Blau, D. M. & Weinberg, B. A. Why the US science and engineering workforce is aging rapidly. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 3879–3884 (2017).
2. Zivin, J. G., Azoulay, P. & Fons-Rosen, C. Does Science Advance One Funeral at a Time? *Am. Econ. Rev.* **109**, 2889–2920 (2019).
3. Jones, B. F. & Weinberg, B. A. Age dynamics in scientific creativity. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 18910–18914 (2011).
4. Jones, B. F. The Burden of Knowledge and the ‘Death of the Renaissance Man’: Is Innovation Getting Harder? *Rev. Econ. Stud.* **76**, 283–317 (2009).
5. Bloom, N., Jones, C. I., Van Reenen, J. & Webb, M. Are Ideas Getting Harder to Find? *Am. Econ. Rev.* **110**, 1104–1144 (2020).
6. Einstein, A. On the Generalized Theory of Gravitation. *Scientific American* vol. 182 13–17 (1950).
7. Einstein, A., Podolsky, B. & Rosen, N. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Phys. Rev.* **47**, 777–780 (1935).
8. Planck, M. Max Planck: Scientific Autobiography and Other Papers. (1949).
9. Kuhn, T. S. The structure of scientific revolutions: University of Chicago press. *Original edition* (1962).
10. Price, D. J. Citation measures of hard science, soft science, technology, and nonscience. *Communication among scientists and engineers* **3**, 22 (1970).
11. Larivière, V., Archambault, É. & Gingras, Y. Long-term patterns in the aging of the scientific literature, 1900--2004. in *Proceedings of the 11th conference of the international society for scientometrics and informetrics (ISSI), Madrid, Spain* 449–456 (2007).
12. Milojević, S. How are academic age, productivity and collaboration related to citing behavior of researchers? *PLoS One* **7**, e49176 (2012).
13. Collison, P. & Nielsen, M. Science is getting less bang for its buck. *Atlantic* (2018).
14. Milojević, S. Quantifying the cognitive extent of science. *J. Informetr.* **9**, 962–973 (2015).

15. Jurgens, D., Kumar, S., Hoover, R., McFarland, D. & Jurafsky, D. Measuring the evolution of a scientific field through citation frames. *Trans. Assoc. Comput. Linguist.* **6**, 391–406 (2018).
16. Cohan, A., Ammar, W., van Zuylen, M. & Cady, F. Structural Scaffolds for Citation Intent Classification in Scientific Publications. *arXiv [cs.CL]* (2019).
17. Sinha, A. *et al.* An Overview of Microsoft Academic Service (MAS) and Applications. in *Proceedings of the 24th International Conference on World Wide Web* 243–246 (Association for Computing Machinery, 2015).
18. Wang, K. *et al.* A Review of Microsoft Academic Services for Science of Science Studies. *Front Big Data* **2**, 45 (2019).
19. Funk, R. J. & Owen-Smith, J. A Dynamic Network Measure of Technological Change. *Manage. Sci.* **63**, 791–817 (2016).
20. Wu, L., Wang, D. & Evans, J. A. Large teams develop and small teams disrupt science and technology. *Nature* **566**, 378–382 (2019).
21. Giddens, A. *The Constitution of Society: Outline of the Theory of Structuration*. (University of California Press, 1984).
22. Chen, C. *Mapping Scientific Frontiers: The Quest for Knowledge Visualization*. (Springer, London, 2013).
23. Monechi, B., Ruiz-Serrano, Á., Tria, F. & Loreto, V. Waves of novelties in the expansion into the adjacent possible. *PLoS One* **12**, e0179303 (2017).
24. Tshitoyan, V. *et al.* Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature* **571**, 95–98 (2019).
25. Hill, R., Yin, Y., Stein, C., Wang, D. & Jones, B. F. Adaptability and the Pivot Penalty in Science. (2021) doi:10.2139/ssrn.3886142.
26. Merton, R. K. The Matthew effect in science. The reward and communication systems of science are considered. *Science* **159**, 56–63 (1968).
27. Wang, D., Song, C. & Barabási, A.-L. Quantifying long-term scientific impact. *Science* **342**,

- 127–132 (2013).
28. Adams, D. *The Salmon of Doubt: Hitchhiking the Universe One Last Time*. (Harmony Books, 2002).
  29. Shen, Z., Ma, H. & Wang, K. A Web-scale system for scientific knowledge exploration. *arXiv [cs.CL]* (2018).
  30. Higgs, P. W. Spontaneous Symmetry Breakdown without Massless Bosons. *Phys. Rev.* **145**, 1156–1163 (1966).

**Supplementary Information for  
Aging Scientists and Slowed Advance**

Haochuan Cui, Lingfei Wu, James A. Evans

Correspondence to: [jevans@uchicago.edu](mailto:jevans@uchicago.edu) (J.A.E.)

## Methods

### Data sets

We analyze two types of datasets: (i) Publication record for 2,310,301 scholars selected from the Microsoft Academic Graph (MAG). MAG has disambiguated the names of scholars and their institutions<sup>17</sup>. These scholars published 46,003,252 papers in 48,953 journals across 19 different scientific fields from 1800 to 2020. As data are sparse before the 1950s and after 2015, we focused on papers published 1960-2015. Results presented in the main figures show the behavior of a subset of 1,025,084 scholars, who had a career age of twenty or more years and wrote at least ten papers. This data set is verified using the Web of Science<sup>20</sup>. (ii) Citation graph containing 70,558,203 papers published during 1960-2015, which contains 611,483,153 references from or citations to these papers.

### Identifying early-successes and late-bloomers

Recognition arrives at different stages of the career, differentiating “early successes” from “late bloomers”<sup>26</sup> in academia. Here we define a hit paper as those among the top 1% most cited within the same field and year. We then divide scholars who own these field-definitive papers into three groups according to the publication timing of their first hit paper, including “earlier successes” (within the first five years), “late-bloomers” (after twenty years), and the rest (between early and late). Our pattern of results is not sensitive to these thresholds.

### Mapping the scientific fields of papers and topics studied by scholars

We used the scientific taxonomy created by the Microsoft Academic Graph team<sup>17</sup> to identify the scientific fields of papers and the topics studied by scholars. The MAG taxonomy has six levels. Level zero comprises 19 coarse-grained fields, level one lists 292 subfields, and level 2-5 contains 543,454 unique keywords (called MAG keywords hereafter). Each MAG paper is labeled by one or more keywords, which permits grouping papers into fields to map change in topics (keywords) for these fields. Using this taxonomy, we also identified the home field for 1,151,907 scholars based on where their productivity is concentrated (if over a half of their papers are published within a field) to explore how scholars change topics at different rates across fields.

### Quantifying keyword coverage and repetition for scholars and the churn rate of fields

To calculate the annual popular keyword coverage of a scholar, we identify the top 1% popular MAG keywords from his/her home field in the given year, and calculate the fraction of field keywords covered by the scholar’s publications. We also compare paper keywords between successive years for each scholar’s publication history to calculate the rate of keyword repetition. We calculate field “churn rate” by applying the same keyword repetition analysis to fields.

### Detecting citation functions and calculating claim contrasting probability

We use a deep learning method<sup>16</sup> developed based on the pioneering work by Jurgens et al.<sup>15</sup> to infer the function or role of citations based on surrounding words. This NLP model successfully classifies citations into six groups, including “Background,” “Extends,” “Uses,” “Motivation,” “Compare/Contrast,” and “Future work.” We identified the functions of 235,598,500 citations of 12,177,040 papers published during 1895 to 2020 based on their discourse context provided in

the MAG data. 8.16% of total citations are labeled as “Compare/Contrast” by our model. 43% of the analyzed papers had at least one outbound “Compare/Contrast” citation and were identified as “claim contrasting” papers. We calculate the annual fraction of claim-contrasting papers for each scholar and analyze how it changes over the career. In Figure 2a, we calculated the 2-sample *K-S* statistic between 100 samples each from the distribution of citations from older and younger scholars. (Across 20 such sampling estimations, the statistic was always significant). In Figure 2b, we calculated it between the randomized null model and the age distribution of critical citations for each field.

### **Analyzing anti-aging contingent on collaborators, institutions, and topics**

We select 2,310,301 scholars who contributed to ten or more research articles over two decades or longer as the focal group and analyze their configuration of collaborators, institutions, and topics throughout their career. To identify collaborator preference, we calculate the median career age of collaborators for all scholars from the focal group at a given career year, and separate papers as either “collaborating with seniors” or “collaborating with juniors” based on whether the mean career age of collaborators within that paper is higher or lower than the group average. In identifying institution preference, we keep track of all institutions for a target scholar from the focal group, and analyze whether a paper was written before (“old institutions”) or after (“new institutions”) a new institution was reported. To identify preferences in topic, we track the expanding vocabulary of MAG keywords for each scholar from the focal group and label a paper as either “old topics” or “new topics” according to whether it introduces keywords new to the researcher.

### **Estimating relationships between focal variables**

We initially estimated OLS models to explore the relationship between reference and career age. We selected 1,047,637 scholars who published over twenty years and contributed to ten or more research articles then estimated the effect for each field. The average coefficient across all fields is approximately 0.09, i.e., a one month increase of reference age for each year across the publishing career.

We then estimated Hierarchical Linear Models (HLMs) to evaluate the same relationship, but accounting for mean age in the field (Table S2). We analyze the annual statistics of reference age and career age for all 4,677 scientists who belong to the 1 million focal scientists under analysis with trackable records in the citation context datasets. We assign a scholar’s “home” if half or more papers authored by that scholar over the entire career were published within one of the eleven scientific fields under study. For each scholar in a year, we calculated the average career age of their “home” field in the same year. The HLM comprises three levels including annual statistics of a scientist (level-1), scientists (level-2), and fields (level-3). Annual reference age (*RefAge*) is the dependent variable for level-1, career age (*Age*) and average team size are the independent variables for level-1, and mean career age in field (*MeanAge*) and mean teamsizes in field (*MeanTS*) are the independent variables for level-3 as shown below. Number of subscripts designate a variable at that level of the model (e.g.,  $r_{100}$  at level-3 of the models has 3 subscripts).

Level-1

$$RefAge = \pi_0 + \pi_1 Age + \pi_2 TS + e$$

Level-2

$$\begin{aligned}\pi_0 &= \beta_{00} + r_0 \\ \pi_1 &= \beta_{10} + r_1 \\ \pi_2 &= \beta_{20} + r_2\end{aligned}$$

Level-3

$$\begin{aligned}\beta_{00} &= r_{000} + r_{001}(MeanAge) + r_{002}(MeanTS) + u_{00} \\ \beta_{10} &= r_{100} + r_{101}(MeanAge) + r_{102}(MeanTS) + u_{10} \\ \beta_{20} &= r_{200} + u_{20}\end{aligned}$$

Our findings from this estimated model (see Table S2) confirm that holding constant mean age within field and mean team size, increased scientist age continues to manifest a statistically significant and substantial positive relationship with reference age, consistent with the OLS model above. Mean age posts an additional positive relationship, suggesting that aging scientists in older fields cite the oldest work.

We then estimated HLMs to explore the relationship between career age and critical rate, accounting for the mean age of scholars in the field (Table S3). As above, we analyzed these data for all 4,677 scientists who belong to the focal 1 million scientists under analysis with trackable records in the citation context datasets, assigning home discipline in the same way. This HLM involves three levels including annual statistics of a scientist (level-1), that scientist over the career (level-2), and that scientist's field (level-3). Annual reference age (*RefAge*) is the dependent variable for level-1, career age (*Age*) and avg. team size are the independent variables for level-1, and mean career age in field (*MeanAge*) and mean teamsize in field (*MeanTS*) are the independent variable for level-3 as shown below.

Level-1

$$CriRate = \pi_0 + \pi_1 Age + \pi_2 TS + e$$

Level-2

$$\begin{aligned}\pi_0 &= \beta_{00} + r_0 \\ \pi_1 &= \beta_{10} + r_1 \\ \pi_2 &= \beta_{20} + r_2\end{aligned}$$

Level-3

$$\begin{aligned}\beta_{00} &= r_{000} + r_{001}(MeanAge) + r_{002}(MeanTS) + u_{00} \\ \beta_{10} &= r_{100} + r_{101}(MeanAge) + r_{102}(MeanTS) + u_{10} \\ \beta_{20} &= r_{200} + u_{20}\end{aligned}$$

Our findings from this estimated model (see Table S3) confirm that holding constant mean age within field and team size, increased scientist age continues to manifest a substantial and statistically significant negative relationship with critical rate. As discussed in the main manuscript, mean age for the field posts a negative relationship with critical rate because in older, slow-moving fields there are less new discoveries to criticize. Moreover, as mean team size grows, critical rate decreases with larger teams focus on constructive advances.

We then performed Granger Causality tests to explore whether average age in field more likely anticipated churn rate, as we suspect, or churn rate better predicts average scholar age in field.

For all 11 fields in the Microsoft Academic Graph, we calculate churn rate in the field and old proportion (percentage of scholars whose career year > 20 in each field) year by year, then predict each other with one and two year lags in both directions. Next, we applied Granger causality tests on subfield churn rate and subfield old proportion, as detailed Table S4. When we alternatively weighted scholars according to their number of annual papers, the pattern of findings remained constant.

Finally, we estimated OLS models tracing the relationship between churn rate on old proportion across 271 subfields and eleven fields (see Table S5). Churn rate is negatively associated with old scientist proportion, both for unweighted (model-1) and weighted data (model-2 in which scientists are weighted according to their number of published papers.)

**Table S1. OLS models regressing reference age on career age.** Ordinary Least Square (OLS) regression coefficients for the data underlying Figure 1 for the 1,047,637 scholars who published over twenty years and contributed to ten or more research articles. The average coefficient across all fields is approximately 0.09, i.e., a one month increase of reference age for each year across the publishing career.

| Field                  | Intercept | Coefficient | R <sup>2</sup> |
|------------------------|-----------|-------------|----------------|
| Mathematics            | 12.699*** | 0.143***    | 0.986          |
| Agriculture            | 10.771*** | 0.075***    | 0.952          |
| Social Sciences        | 9.938***  | 0.102***    | 0.971          |
| Business & Management  | 9.974***  | 0.107***    | 0.983          |
| Environmental Sciences | 9.797***  | 0.078***    | 0.973          |
| Chemistry              | 8.778***  | 0.110***    | 0.978          |
| Physical Sciences      | 8.762***  | 0.092***    | 0.979          |
| Engineering            | 8.310***  | 0.089***    | 0.976          |
| Biology                | 8.443***  | 0.072***    | 0.957          |
| Medicine               | 8.554***  | 0.045***    | 0.895          |
| Computer Science       | 7.119***  | 0.096***    | 0.957          |

\* p-value < 0.05; \*\* p-value < 0.01; \*\*\* p-value < 1.0 × 10<sup>-40</sup>

**Table S2. HLM model regressing reference age on career age.** HLM comprising three levels for the annual statistics of a scientist (level-1), that scientist over the lifecourse (level-2), and the scientist's field (level-3). Annual reference age (*RefAge*) is the dependent variable, with career age (Age) and team size (TS) independent variables in level-1. Mean career age in the field (MEAN\_AGE) and mean team size in the field (MEAN\_TS) are the independent variables in level-3 as shown below, revealing that as mean age goes up, references get older, but as mean team size increases, references become more recent. Controlling for these higher-level factors, scientist career age retains by the most significant and substantial relationship with reference age.

**Final estimation of fixed effects (with robust standard errors)**

| Fixed Effect               | Coefficient   | Standard error | T-ratio | Approx. d.f. | p-value |
|----------------------------|---------------|----------------|---------|--------------|---------|
| For INTRCPT1, $\pi_0$      |               |                |         |              |         |
| For INTRCPT2, $\beta_{00}$ |               |                |         |              |         |
| INTRCPT3, $\gamma_{000}$   | 9.515         | 0.425          | 22.404  | 9            | <0.001  |
| MEAN_AGE, $\gamma_{001}$   | 0.401         | 0.390          | 0.997   | 9            | 0.211   |
| MEAN_TS, $\gamma_{002}$    | 0.389         | 0.390          | 0.997   | 9            | 0.176   |
| For AGE slope, $\pi_1$     |               |                |         |              |         |
| For INTRCPT2, $\beta_{01}$ |               |                |         |              |         |
| INTRCPT3, $\gamma_{100}$   | <b>0.126</b>  | 0.017          | 7.630   | 9            | <0.001  |
| MEAN_AGE, $\gamma_{101}$   | <b>0.016</b>  | 0.010          | 1.488   | 9            | 0.171   |
| MEAN_TS, $\gamma_{102}$    | <b>-0.002</b> | 0.009          | -0.246  | 9            | 0.811   |
| For TS slope, $\pi_2$      |               |                |         |              |         |
| For INTRCPT2, $\beta_{02}$ |               |                |         |              |         |
| INYRCPT3, $\gamma_{200}$   | -0.0389       | 0.011          | -3.505  | 11           | 0.006   |

**Final estimation of level-1 and level-2 variance components**

| Random Effect    | Standard Deviation | Variance Component | d.f. | $\chi^2$  | p-value |
|------------------|--------------------|--------------------|------|-----------|---------|
| INTRCPT1, $r_0$  | 2.022              | 4.088              | 4642 | 44192.832 | <0.001  |
| Age slope, $r_1$ | 0.191              | 0.036              | 4642 | 9625.648  | <0.001  |
| TS slope, $r_2$  | 0.030              | 0.000              | 4642 | 5169.017  | <0.001  |
| level-1, e       | 2.427              | 5.891              |      |           |         |

Note: The chi-square statistics reported above are based on only 4654 of 4677 units that had sufficient data for computation. Fixed effects and variance components are based on all the data

**Final estimation of level-3 variance components**

| Random Effect               | Standard Deviation | Variance Component | d.f. | $\chi^2$ | p-value |
|-----------------------------|--------------------|--------------------|------|----------|---------|
| INTRCPT1/INTRCPT2, $u_{00}$ | 1.427              | 2.066              | 9    | 1741.678 | <0.001  |
| AGE/INTRCPT2, $u_{10}$      | 0.049              | 0.049              | 9    | 111.190  | <0.001  |
| TS/INTRCPT2, $u_{20}$       | 0.033              | 0.033              | 11   | 26.128   | <0.001  |

**Table S3. OLS Models regressing Churn Rate on Old Proportion.** Churn Rate relationship with old scientist proportion for unweighted (model-1) and weighed data (model-2; scientists are weighted according to their number of published papers.)

| Independent Variables    | Model 1: unweighted |             | Model 2: weighted by paper productivity |             |
|--------------------------|---------------------|-------------|---|-------------|
|                          | Intercept           | Coefficient | Intercept                               | Coefficient |
| <b>Mathematics</b>       | 0.184***            | -0.350***   | 0.186***                                | -0.282***   |
| <b>Geology</b>           | 0.197***            | -0.359***   | 0.194***                                | -0.266***   |
| <b>Economics</b>         | 0.136***            | -0.188***   | 0.135***                                | -0.155***   |
| <b>Psychology</b>        | 0.167***            | -0.263***   | 0.164***                                | -0.190***   |
| <b>Materials Science</b> | 0.165***            | -0.390***   | 0.164***                                | -0.282***   |
| <b>Physics</b>           | 0.120***            | -0.157***   | 0.117***                                | -0.098***   |
| <b>Chemistry</b>         | 0.128***            | -0.247***   | 0.133***                                | -0.187***   |
| <b>Biology</b>           | 0.165***            | -0.344***   | 0.167***                                | -0.254***   |
| <b>Medicine</b>          | 0.164***            | -0.409***   | 0.156***                                | -0.243***   |
| <b>Engineering</b>       | 0.179***            | -0.407***   | 0.176***                                | -0.311***   |
| <b>Computer Science</b>  | 0.202***            | -0.509***   | 0.202***                                | -0.363***   |

\* p-value < 0.05; \*\* p-value < 0.01; \*\*\* p-value < 0.001

**Table S4. Granger Causality Test on Churn Rate and Average Age in Field.** Predictive relationships between churn rate in the field and old proportion (percentage of scholars whose career year > 20 in each field) year by year with one and two year lags in both directions (e.g., Churn Rate predicting Old Proportion; Old Proportion predicting Churn Rate). Granger causality tests on subfield churn rate and subfield old proportion suggest that old proportion is much more likely to predict churn rate than the converse.

| Independent Variables    | Old Proportion Predicts Churn Rate |             | Churn Rate Predicts Old Proportion |             |
|--------------------------|------------------------------------|-------------|------------------------------------|-------------|
|                          | SSR F-test                         |             | SSR F-test                         |             |
|                          | lags=1 year                        | lags=2 year | lags=1 year                        | lags=2 year |
| <b>Mathematics</b>       | 9.67***                            | 4.76**      | 8.17***                            | 5.55***     |
| <b>Geology</b>           | 16.76***                           | 7.74***     | 0.36                               | 4.69**      |
| <b>Economics</b>         | 11.45***                           | 9.16***     | 0.61                               | 1.59        |
| <b>Psychology</b>        | 18.66***                           | 6.65***     | 0.32                               | 0.96        |
| <b>Materials science</b> | 14.29***                           | 4.34**      | 5.49**                             | 3.65**      |
| <b>Physics</b>           | 9.08***                            | 2.19        | 4.05**                             | 2.27        |
| <b>Chemistry</b>         | 11.09***                           | 2.45*       | 1.29                               | 0.52        |
| <b>Biology</b>           | 13.82***                           | 4.01**      | 17.07***                           | 8.31***     |
| <b>Medicine</b>          | 17.27***                           | 5.07***     | 0.1                                | 0.57        |
| <b>Engineering</b>       | 16.38***                           | 4.98**      | 2.84*                              | 1.8         |
| <b>Computer science</b>  | 30.53***                           | 13.33***    | 4.19**                             | 0.75        |

\* p-value < 0.05; \*\* p-value < 0.01; \*\*\* p-value < 0.001

**Table S5. HLM model regressing critical rate on career age.** HLM modeling three levels including annual statistics of a scientist (level-1), that scientist over the lifecourse (level-2), and the scientist's field (level-3). Annual critical rate (*CriRate*) is the dependent variable with career age (Age) and team size (TS) independent variables in level-1. Mean career age in the field (MEAN\_AGE) and mean team size in the field (MEAN\_TS) are the independent variables in level-3 as shown below, revealing that as mean age and team size go up, the critical rate goes down, with less new work to criticize and teams collaborating around constructive new work. Controlling for these higher-level factors, scientist career age retains a significant and substantial relationship with critical rate.

**Final estimation of fixed effects (with robust standard errors)**

| Fixed Effect               | Coefficient   | Standard error | T-ratio | Approx. d.f. | p-value |
|----------------------------|---------------|----------------|---------|--------------|---------|
| For INTRCPT1, $\pi_0$      |               |                |         |              |         |
| For INTRCPT2, $\beta_{00}$ |               |                |         |              |         |
| INTRCPT3, $\gamma_{000}$   | 50.247        | 1.470          | 34.177  | 9            | <0.001  |
| MEAN_AGE, $\gamma_{001}$   | -3.259        | 1.251          | -2.606  | 9            | 0.029   |
| MEAN_TS, $\gamma_{002}$    | -1.810        | 0.763          | -2.373  | 9            | 0.042   |
| For AGE slope, $\pi_1$     |               |                |         |              |         |
| For INTRCPT2, $\beta_{01}$ |               |                |         |              |         |
| INTRCPT3, $\gamma_{100}$   | <b>0.652</b>  | 0.061          | 10.704  | 9            | <0.001  |
| MEAN_AGE, $\gamma_{101}$   | <b>-0.338</b> | 0.060          | -5.610  | 9            | <0.001  |
| MEAN_TS, $\gamma_{102}$    | <b>-0.142</b> | 0.044          | -3.215  | 9            | 0.011   |
| For TS slope, $\pi_2$      |               |                |         |              |         |
| For INTRCPT2, $\beta_{02}$ |               |                |         |              |         |
| INYRCPT3, $\gamma_{200}$   | 0.1617        | 0.088          | 1.836   | 11           | 0.093   |

**Final estimation of level-1 and level-2 variance components**

| Random Effect    | Standard Deviation | Variance Component | d.f. | $\chi^2$  | p-value |
|------------------|--------------------|--------------------|------|-----------|---------|
| INTRCPT1, $r_0$  | 12.790             | 163.592            | 4642 | 12024.477 | <0.001  |
| Age slope, $r_1$ | 1.342              | 1.802              | 4642 | 6073.891  | <0.001  |
| TS slope, $r_2$  | 1.323              | 1.752              | 4642 | 5118.783  | <0.001  |
| level-1, e       | 235.576            | 1265.681           |      |           |         |

Note: The chi-square statistics reported above are based on only 4654 of 4677 units that had sufficient data for computation. Fixed effects and variance components are based on all the data

**Final estimation of level-3 variance components**

| Random Effect               | Standard Deviation | Variance Component | d.f. | $\chi^2$ | p-value |
|-----------------------------|--------------------|--------------------|------|----------|---------|
| INTRCPT1/INTRCPT2, $u_{00}$ | 4.779              | 22.842             | 9    | 427.209  | <0.001  |
| AGE/INTRCPT2, $u_{10}$      | 0.131              | 0.0171             | 9    | 23.638   | 0.005   |
| TS/INTRCPT2, $u_{20}$       | 0.137              | 0.019              | 11   | 19.727   | 0.049   |

**Table S6. OLS models regressing reference age on career age, controlling for moderating influences.** We selected 1,047,637 scholars who published over twenty years and contributed to ten or more research articles (same as Table S1). We then construct a sample of 5,426,422 data points covering each scholar in each publishing year. For a scholar-publishing year, we calculate the average reference age and career age of that scholar, and constructed three dummy variables to describe their behavior, including whether they joined a new institution, published on a new topic, or collaborated with scholars younger than themselves. We regressed reference age on career age and these dummy variables using Ordinary Least Square (OLS) regression and reported the estimated coefficients and significances below.

| Independent Variables               | Dependent Variable: Reference Age |                |
|-------------------------------------|-----------------------------------|----------------|
|                                     | Coefficient                       | Standard Error |
| Intercept                           | 8.6596***                         | 0.004          |
| Career Age                          | 0.0511***                         | <0.001         |
| Joining in new institution          | -0.0823***                        | 0.01           |
| Expanding to new topics             | -0.1699***                        | 0.012          |
| Collaborating with younger scholars | -0.3249***                        | 0.004          |

\*  $p < 0.05$ ; \*\*  $p$ -value  $< 0.01$ ; \*\*\*  $p$ -value  $< 0.001$ ,  $N = 5,426,422$