The misleading narrative of the canonical faculty productivity trajectory

Samuel F. Way[a,1], Allison C. Morgan[a], Aaron Clauset[a,b,c,2], and Daniel B. Larremore[a,b,c,1,2]

[a]Department of Computer Science, University of Colorado, Boulder, CO 80309; [b]BioFrontiers Institute, University of Colorado, Boulder, CO 80303; and [c]Santa Fe Institute, Santa Fe, NM 87501

Edited by Kenneth W. Wachter, University of California, Berkeley, CA, and approved September 18, 2017 (received for review February 13, 2017)

A scientist may publish tens or hundreds of papers over a career, but these contributions are not evenly spaced in time. Sixty years of studies on career productivity patterns in a variety of fields suggest an intuitive and universal pattern: Productivity tends to rise rapidly to an early peak and then gradually declines. Here, we test the universality of this conventional narrative by analyzing the structures of individual faculty productivity time series, constructed from over 200,000 publications and matched with hiring data for 2,453 tenure-track faculty in all 205 PhD-granting computer science departments in the United States and Canada. Unlike prior studies, which considered only some faculty or some institutions, or lacked common career reference points, here we combine a large bibliographic dataset with comprehensive information on career transitions that covers an entire field of study. We show that the conventional narrative confidently describes only one-fifth of faculty, regardless of department prestige or researcher gender, and the remaining four-fifths of faculty exhibit a rich diversity of productivity patterns. To explain this diversity, we introduce a simple model of productivity trajectories and explore correlations between its parameters and researcher covariates, showing that departmental prestige predicts overall individual productivity and the timing of the transition from first- to last-author publications. These results demonstrate the unpredictability of productivity over time and open the door for new efforts to understand how environmental and individual factors shape scientific productivity.

data analysis | career trajectory | computer science | productivity | sociology

Scholarly publications serve as the primary mode of communication through which scientific knowledge is developed, discussed, and disseminated. The amount that an individual researcher contributes to this dialogue—their scholarly productivity—thus serves as an important measure of the rate at which they contribute units of knowledge to the field, and this measure is known to influence the placement of graduates into faculty jobs (1), the likelihood of being granted tenure (2, 3), and the ability to secure funding for future research (4).

The trajectory of productivity over the course of a researcher's lifetime has been studied for at least 60 years, with the common observation being that a researcher's productivity rises rapidly to a peak and then slowly declines (5–9), which has inspired the construction of mechanistic models with a similar profile (7, 9–12). These models have included factors like cognitive decline with age, career age, finite supplies of human capital, and knowledge advantages conferred by recent education, as well as skill deficits among the young (among others), and have been supported by the observation that individual productivity curves feature both long- and medium-term fluctuations (12) and are not well described by even fourth-degree polynomial models (9). Indeed, every study we found to date proposes or confirms a "rise and decline," "curvilinear," or "peak and tapering" productivity trajectory, regardless of whether researchers are binned by chronological age (5–8, 10–13), career age (9, 10), or (only for young researchers) years since first publication (14). The pattern may even extend to mentorship, supported by a finding that the protégés of early-career mathematicians tended to mentor more students, themselves, than protégés trained by those same faculty late in their careers (15). In fact, this conventional narrative of the life course is not restricted to academia, with similar trajectories observed in criminal behavior and artistic production in 1800s France (16) and even in productivity of food acquisition by hunter-gatherers (17).

While these past studies have firmly established that the conventional academic productivity narrative is equally descriptive across fields and time, their analyses are based on averages over hundreds or thousands of individuals (5–11, 13–17). This raises two crucial and previously unanswered questions: Is this average trajectory representative of individual faculty? And how much diversity is hidden by a focus on a central tendency over a population? To answer these questions, we combine and study two comprehensive datasets that span 40 years of productivity for nearly every tenure-track professor in a North American PhD-granting computer science department. By introducing a simple mathematical description of the shape of a scientist's productivity over time, we map individuals' publication histories to a low-dimensional parameter space, revealing substantial diversity in the publication trends of individual faculty and showing that only a minority follow the conventional narrative of productivity. In fact, even among the conventional trajectories, individuals exhibit large fluctuations in their productivity around the average trend. Together, these results reveal that population averages provide a dramatically inaccurate picture of intellectual contributions

Significance

Scholarly productivity impacts nearly every aspect of a researcher's career, from their initial placement as faculty to funding and tenure decisions. Historically, expectations for individuals rely on 60 years of research on aggregate trends, which suggest that productivity rises rapidly to an early-career peak and then gradually declines. Here we show, using comprehensive data on the publication and employment histories of an entire field of research, that the canonical narrative of "rapid rise, gradual decline" describes only about one-fifth of individual faculty, and the remaining four-fifths exhibit a rich diversity of productivity patterns. This suggests existing models and expectations for faculty productivity require revision, as they capture only one of many ways to have a successful career in science.

Author contributions: S.F.W., A.C., and D.B.L. designed research; S.F.W., A.C.M., and D.B.L. performed research; S.F.W., A.C.M., and D.B.L. analyzed data; and S.F.W., A.C.M., A.C., and D.B.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: daniel.larremore@colorado.edu or samuel.way@colorado.edu.

[2]A.C. and D.B.L. contributed equally to this work.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1702121114/-/DCSupplemental.

over time and that productivity patterns are both more diverse and less predictable than previously thought. These findings were preliminarily described in a recent review (18) which provides additional context for the results reported fully here.

Moreover, while we show that the distribution of productivity trajectories resists natural categorization, it is nevertheless possible to explore covariates that are associated with different regions of its parameter space. The literature on such associations has avoided detailed trajectories and instead focused on the complicated relationship between prestige, productivity, and hiring. Past studies have found that researchers trained at prestigious institutions are likely to remain productive (19), regardless of where they place as faculty (20). Other results link the prestige of the doctorate and the advisor to early-career productivity but not long-term productivity (21), which is at odds with other studies (22, 23) that found early-career productivity predicts long-term productivity. Disagreement about hiring exists as well, with multiple studies finding that doctoral prestige and not productivity drives the initial placement of faculty (24, 25), while recent work based on comprehensive data in multiple fields suggests that prestige alone is insufficient to fully explain faculty placement (1, 26). This, too, is complicated by hypotheses of mutual causality, where departments both select for and facilitate high productivity (27). Unfortunately, while such studies shed light on a complicated system, they tend to restrict their analyses to unusual scientists, such as Nobel laureates or faculty at elite departments, rather than typical researchers. In contrast, the data analyzed here are comprehensive, covering faculty across the prestige hierarchy, which enables us to move beyond total productivity to study publication trajectories in light of prestige, hiring, and past productivity alike.

This study exploits and combines two large datasets related to faculty productivity. The first one is a comprehensive, hand-curated collection of education and academic appointment histories for tenure-track and tenured computer science faculty (26). This dataset spans all 205 departmental or school-level academic units on the Computing Research Association's Forsythe List of PhD-granting departments in computing-related disciplines in the United States and Canada (archive.cra.org/reports/forsythe.html).

For each department, the dataset provides a complete list of regular faculty for the 2011–2012 academic year, and for each of the 5,032 faculty in this collection, it provides partial or complete information on their education and academic appointments, obtained from public online sources, mainly résumés and homepages. Of these, we selected the 2,583 faculty who both received their PhD from and held their first assistant professorship at one of these institutions and for whom the year of that hire is known and occurred in 1970–2011. The first requirement ensured that we modeled the relatively closed North American faculty market; roughly 87% of computing faculty received their PhD from one of the Forsythe institutions, and past analysis has shown that Canada and the United States are not distinct job markets in computer science (26). A number of faculty were removed in this step because the location of their first assistant professorship was not known; these were mainly senior faculty.

The first dataset also provides a ranking of institutional prestige π, derived from patterns in the PhD-to-faculty hiring network between departments. In short, π is a consensus of ordinal rankings (lower is better) in which prestige is defined recursively: Prestigious departments are those whose graduates are hired as faculty in prestigious departments. Networks, code, and rankings are available in ref. 26.

The second dataset, constructed around the first one, is a complete publication history as listed in the Digital Bibliography and Library Project (DBLP; dblp.uni-trier.de), an online database that provides open bibliographic information for most journals and conference proceedings relevant to computing research, using manual name disambiguation as necessary. For each paper in a faculty's publication history, we recorded the paper's title, author list (preserving author order), and year of publication. By following this procedure, we collected data for 200,476 publications which covered 2,453 (95.0%) faculty in our sample. Of those, we manually collected records of all peer-reviewed conference and journal publication histories from the publicly available curricula vitae (CVs) of 109 faculty, a randomly selected 10% of the 1,091 faculty with career lengths between 10 years and 25 years, providing a benchmark dataset to evaluate the accuracy of DBLP data (*Collection of CV Data*).

Our combined dataset consists of the career trajectories of these 2,453 tenure-track faculty as of 2011–2012: each professor's publicly accessible metadata, their time-stamped PhD and employment history, and the annotated time series of their publications. We note that this dataset does not include information on faculty who have retired or left academia before 2011. Implications of these data limitations for the conclusions that can be drawn from our analyses are explored in *Discussion*. Finally, this study was not reviewed by an institutional review board because all data used were collected from publicly available sources. All results are presented anonymously or in aggregate to avoid revealing personally identifiable information about individual scientists.

Results

General Trends in Productivity. Two broad trends characterize scholarly productivity in academic computer science. First, publication rates have been increasing over the past 45 years, and second, higher publication rates are correlated with higher prestige. These two observations are intertwined and underpin a number of subsequent analyses, so we explore them briefly in more depth.

Past studies have found that researchers at more prestigious institutions tend to be more productive (20, 21, 24, 25, 27, 28). Our data corroborate this finding, but we also find that the typical productivity advantage associated with greater prestige holds regardless of whether an institution is public or private, for both early-career publications (first 10 years; Fig. 1) and lifetime publications (Fig. S4). Regressing the median number of time-adjusted publications (below) among faculty in a department against departmental prestige indicates that the relationships between prestige and productivity are statistically indistinguishable for public and private institutions, with expected increases

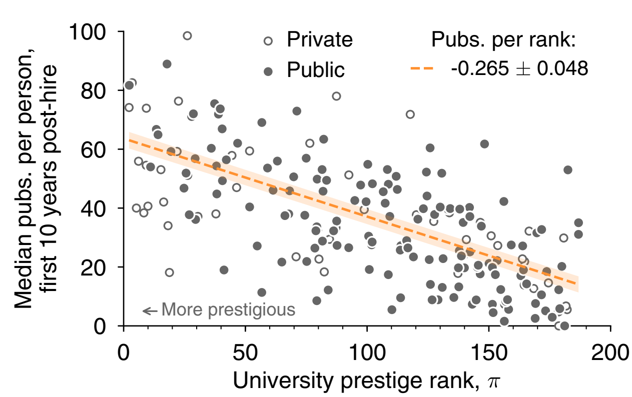

Fig. 1. Publications correlate with institution prestige. Circles indicate median number of publications per person per institution for researchers' first 10 years posthire, adjusted for growth in publication rates over time (Figs. S1–S3 and *General Trends in Productivity Data*) and ordered by institutional prestige, π (26). Effects of prestige are similar for private (open circles) and public (closed circles) institutions ($P = 0.146$, t test; main text), increasing at a rate of nearly 2.7 publications per 10-rank improvement in prestige. Shaded region denotes the 95% confidence interval for least-squares regression. pubs, publications.

in the first decade of a career of roughly 2.7 publications for every 10-rank improvement in prestige. In fact, when comparing public and private institutions, neither the prestige–productivity slope nor productivity overall is significantly different ($P = 0.150$, 0.148, respectively, two-tailed t test), contradicting the conventional wisdom that private universities enjoy a productivity advantage over public ones. The conventional wisdom is likely skewed by a focus on elite departments, as 8 of the top 10 computer science departments are private (26), but in fact, private institutions are distributed evenly across all ranks. Expanding this analysis to include lifetime publications increases the prestige–publication slope to 3.28 publications per 10-rank improvement in prestige but does not alter the nonsignificance of public/private status ($P = 0.714$, 0.346, two-tailed t test).

Past studies have also found that publication rates have increased over time (29, 30). However, before investigating whether changes in publication rates apply to computer science, we used the manually collected CV data to probe the extent of DBLP's coverage. Indeed, the fraction of publications indexed by DBLP is nonuniform over time, increasing linearly from around 55% in the 1980s to over 85% by 2011 ($R^2 = 0.685$, $P < 0.001$, two-tailed t test; Fig. S1 and *General Trends in Productivity Data*). Because DBLP's coverage of the published literature varies over time, in the analyses that follow we use data from hand-collected faculty CVs whenever possible and otherwise apply a statistical correction to DBLP's data to account for its lower coverage.

Knowing already that there are substantial differences in productivity by prestige, we separated universities by prestige into five groups of approximately equal size and investigated whether the growth of publication rate varies by prestige. We find that the average number of publications per person produced in each calendar year has been increasing at all five strata of prestige at rates between 0.72 and 1.23 publications per decade, for 45 years (Fig. S3). Because we have used data from hand-collected faculty CVs to adjust DBLP-derived paper counts for DBLP's steadily improving coverage over time, these estimated growth rates represent a real increase in publication rates over this 40-year period. Moreover, the observed steady increase in productivity is not uniform across prestige, and the difference between production growth rates between higher- and lower-prestige departments has widened slightly but significantly over time ($P < 0.05$, two-tailed t test). In other words, prestigious and nonprestigious institutions have contributed to the overall growth at different rates. Not only are there small but significant differences in productivity by prestige (Fig. 1) but also those differences are slowly growing (Fig. S3).

To investigate the productivity patterns of individual researchers and test the conventional narrative of rapidly rising productivity followed by a gradual decline, for the remainder of this paper we focus on time series of individual productivity. However, due to both the observed growth in productivity and the variability in DBLP coverage, it would be misleading to directly compare a 1975 publication with a 2011 publication. Thus, hereafter we use "adjusted" publication counts, which corrects the raw DBLP counts to account for both the changing DBLP coverage and the increasing mean publication rate over time (*General Trends in Productivity Data*). All publication counts are hence reported as 2011-equivalent counts.

Individual Productivity Trajectories. Examining the productivity trajectories of individual researchers, we find that they too exhibit substantial and significant differences in their publication rates. Early studies of scholarly productivity noted profound imbalance in the number of articles published by individual researchers (31, 32). Cole (33) and Reskin (34) in the 1970s noted that about 50% of all scholarly articles were produced by about 15% of the scientific workforce. Our data reflect similar levels of imbalance, with approximately half of all contributions in the dataset authored by only 20% of all faculty. Stratifying by decade, however, the Gini coefficients for productivity imbalance have been declining, from 0.62 in the 1970s to 0.40 in the 2000s (Fig. S5). This trend persists when researchers are restricted to only the publications within the first 5 years of their careers.

There are several possible explanations for the trend of decreasing inequality in individual productivity. For instance, the lower end of the productivity distribution could have become relatively more productive over time, perhaps as more institutions shifted focus from teaching to research. Or, it may reflect a strengthening selective filter on highly productive faculty, perhaps as community expectations for continual productivity rose. It may also reflect nonuniform errors in the DBLP data, although the correction for DBLP coverage should account for these (*General Trends in Productivity Data*).

We now focus on testing the conventional productivity narrative that has been described in various disciplines and at many points in time (5–11, 13, 14, 16, 17): Productivity climbs to a peak and then gradually declines over the course of the researcher's career. Across computer science faculty, we find that the average number of publications per year over a faculty career is highly stereotyped (Fig. 2), with a rising productivity that peaks after around 5 years, declines slowly for another 5 years, and then remains roughly constant for any remaining years. Although departmental prestige correlates with productivity in several ways (Fig. 1 and Fig. S3), it does not alter this stereotypical pattern, which appears essentially unchanged across departments with different levels of prestige, except for a roughly constant shift up as prestige increases (Fig. 2).

The suggestion that productivity grows in the early years of a career has intuitive appeal. Professors settle into their research environments, begin training graduate students, and build their cases for promotion and tenure. Similarly, many reasons have been suggested for why productivity might decrease after promotion, including increased service and nonresearch commitments, declining cognitive abilities, and increased levels of distraction from outside work due to health issues and childcare obligations (35). Although an average over faculty appears to confirm the stereotyped trajectory of rapid growth, peak, and slow decline, it does not reveal whether this average is representative of the many individual trajectories it averages over, nor does it show how much diversity there might be around the average and whether that diversity correlates with other factors of interest.

Fig. 2. Average publications follow conventional narrative across prestige. Five average productivity curves are shown, partitioning universities according to prestige rank π such that each quintile represents $\sim$20% of all faculty in the full dataset. Averages over researchers at all levels of institutional prestige follow similar productivity trajectories, in agreement with the conventional narrative, but at differing scales of output.

To characterize the productivity pattern within an individual career, we fit a simple stereotypical model of productivity over time to the number of papers published per year,

$$f(t) = \begin{cases} b + m_1 t & 0 \leq t \leq t^* \\ b + m_1 t^* + m_2(t - t^*) & t > t^* \end{cases}, \qquad [1]$$

a piecewise linear function in which t^* is the change point between the two lines, m_1 and m_2 are the rates of change in productivity before and after the change point, respectively, and b is the initial productivity (Fig. 3). We apply this model to the $N = 1{,}091$ faculty who have been used for 10–25 years. By fitting these four parameters to each individual's publication trajectory, we map that trajectory into a low-dimensional description of its overall pattern [fitting done by least squares; see *Least-Squares Fit of f(t)* for optimal numerical methods and *Modeling Framework* for detailed discussion of statistical models].

However, before interpreting the distributions of parameters, we subjected each trajectory to two additional tests to ensure that its best-fit parameters were meaningful. First, to avoid overfitting linear trajectories with a piecewise linear model, we performed model selection, asking whether the Akaike information criterion (AIC) with finite-size correction favored a straight line or the more complex $f(t)$ (*Model Selection*). This process conservatively selected only 33.3% ($N = 363$) of researchers who are more confidently modeled by the piecewise function.

Second, to address the possibility that a researcher's best-fit parameters may be sensitive to small changes in the years of their publications, we conducted a sensitivity analysis in which we repeatedly refit model parameters to productivity trajectories, adding a small amount of noise to shift some publications into adjacent years (*Sensitivity to Timing of Publications*). This procedure places each professor's noise-free trajectory within a distribution of nearby noisy trajectories, enabling two different (but ultimately concordant) analyses. The primary sensitivity analysis focuses on individual faculty, computing whether the parameters of each professor's noise-free trajectory are similar to their noisy distribution. This approach revealed that a majority (77.2%) of trajectories are well represented by their noise-free parameters, each consistently falling into the same region of parameter space for over 75% of resampled trajectories. We refer to these trajectories as "stable" in subsequent analyses, meaning that their noise-free parameters are representative and interpretable. The alternative sensitivity analysis focuses on the population of faculty, combining all noise-free trajectories with their noise-added distributions into a single expanded ensemble of conceivable productivity trajectories (*Sensitivity to Timing of Publications*). Although this ensemble is unable to support analyses of individual faculty, we use it to corroborate the findings that follow. Combining the individual stability and AICs, we find that 32.3% ($N = 352$) of researchers possess productivity trajectories that are both stable and nonlinear. All analyses and discussions of model parameters hereafter refer to stable, nonlinear trajectories unless otherwise noted.

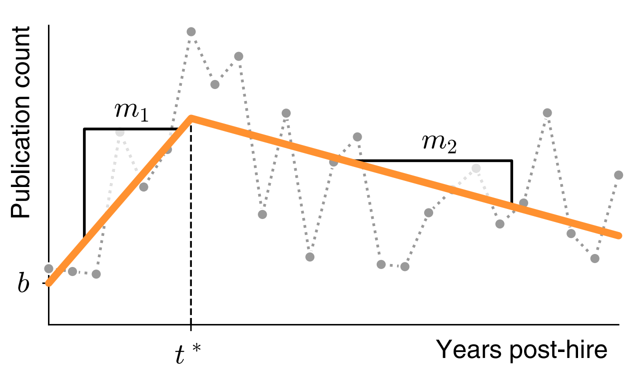

Fig. 3. Example trajectory and piecewise model. Circles represent empirical annual publications. Orange line shows best fit of piecewise linear model **1** with slopes m_1 and m_2, change-point t^*, and intercept b annotated.

The narrative of "early growth in productivity, followed by a slow decline" implies four conditions on the inferred parameters: While the conditions of growth ($m_1 > 0$) and decline ($m_2 < 0$) are straightforward, we interpret "early growth" to mean that inferred peak productivity comes within the first decade after hiring ($t^* \leq 10$) and "slow" to mean that the slope of decline is smaller in magnitude that the slope of growth ($|m_2| < m_1$). After fitting individual trajectory models to the 1,091 faculty in our sample, we find that only 20.1% follow the stereotypical trajectory. Even dropping the aforementioned restriction on t^* increases the fraction meeting the stereotype to only 20.3%. To ensure that these results were not sensitive to our definition of stability in the presence of noise, we generated an ensemble with 200 noise-added trajectories for each professor (Fig. S8 and *Sensitivity to Timing of Publications*), subjected each to the AIC for nonlinearity, and found that only 19.7% of ensemble trajectories are reliably categorized as adhering to the conventional narrative. In other words, the average trajectory, which has been held up as established fact for more than 50 years, describes the behavior of only a minority of researchers, while a large majority of researchers follow qualitatively different trajectories.

Publication trajectories can be divided into four general classes based on the signs of the two slope parameters, m_1 and m_2, corresponding to the quadrants shown in Fig. 4. Individual trajectory shapes exhibit substantial diversity, spanning all four quadrants. Even among faculty whose publication rates grew and then declined (Fig. 4, *Bottom Right* quadrant, 28.6%), the conventional narrative includes only the 20.3% of individuals whose rate of growth exceeds their rate of decline ($m_1 > |m_2|$; shaded region, Fig. 4). Additionally, researchers were distributed similarly across the four quadrants, comparing parameters extracted from DBLP data vs. hand-collected CV data ($P = 0.14$, χ^2), confirming that the dispersion shown in Fig. 4 represents the true diversity of careers.

The cloud of faculty trajectory parameters shown in Fig. 4 does not naturally separate into coherent clusters. In their absence, what are the covariates that predict which region of the plot an individual is likely to occupy? First, early-career growth rate of yearly publications m_1 is significantly correlated ($P < 0.001$, t test) with the prestige of researchers' institutions. This is particularly true for researchers at "elite" institutions, which we define as being in the top 20% of universities according to prestige rank and adjusting for number of faculty (same partitions as in Fig. 2). Specifically, researchers' productivity grows by a median of 2.02 additional papers per year at elite institutions compared with 1.19 for others ($P < 0.001$, one-tailed Mann–Whitney test). Perhaps as a result—what goes up must come down—the slope after the point of change, m_2, correlates significantly with prestige and is more negative for researchers at higher-ranked institutions, compared with those at lower-ranked institutions ($P < 0.05$, t test). Additionally, researchers who received their doctorates from elite institutions exhibit faster early-career growth than those who trained at lower-ranked institutions ($P < 0.05$, one-tailed Mann–Whitney test).

Second, the early-career initial productivity b is significantly higher for faculty who graduated from elite departments ($P < 0.005$, one-tailed Mann–Whitney test). We also find that researchers who place into elite departments or who have postdoctoral experience tend to start out more productive; however, these differences are not statistically significant ($P > 0.05$,

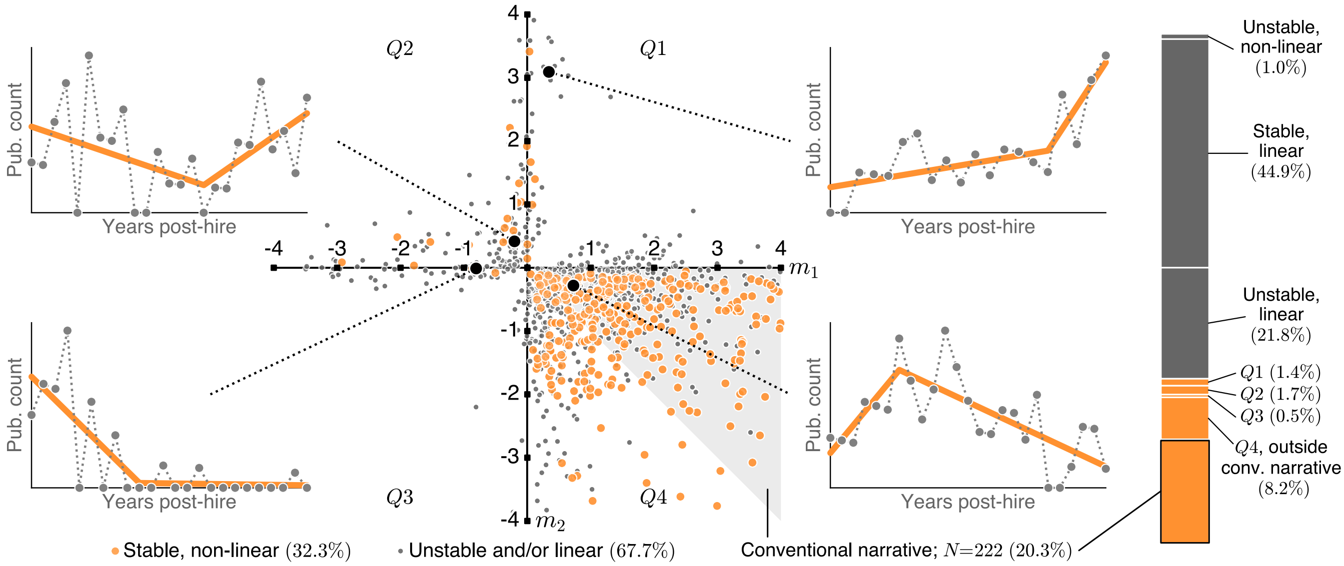

Fig. 4. Distribution of individuals' productivity trajectory parameters. Diverse trends in individual productivity fall into four quadrants based on their slopes m_1 and m_2 in the piecewise linear model Eq. **1**. Plots show example publication trajectories to illustrate general characteristics of each quadrant. The shaded triangular region (*Bottom Center*) corresponds to the conventional narrative of early increase followed by gradual decline. Color distinguishes trajectories in two classes: those that are stable and nonlinear (orange) and those that are either unstable or linear (gray). The plot at *Right* describes how researchers are distributed within these two classes. conv. narrative, conventional narrative; Pub, publication.

Mann–Whitney test). These findings regarding m_1 and b combine to suggest that current academic environment correlates with—and perhaps influences—productivity, while prior academic environment does not. Finally, faculty at top-ranked departments are statistically no more or less likely to be found within this triangular region, a result robust to alternative cutoffs for "top-ranked" institutions.

The relationship between trajectories and gender is more complicated. First, trajectories of male and female researchers were similarly distributed across the four quadrants ($P = 0.94$, χ^2 test), and gender was uncorrelated with the likelihood of meeting the four criteria of the canonical narrative ($P = 0.39$, χ^2 test). Further, within this canonical subset, the women's initial productivity grew at a rate indistinguishable from the men's ($P = 0.15$, Mann–Whitney test) and peaked in similar years ($P = 0.305$, Kolmogorov–Smirnov test). Women's initial productivity, however, was 0.46 publications lower than the men's ($P = 0.032$, Mann–Whitney test) in general, and this difference exists despite the fact that men and women in this subset trained and were hired at similarly ranked institutions ($P > 0.05$, Kolmogorov–Smirnov test) and completed postdoctoral training at similar rates ($P = 0.89$, χ^2 test).

The change point within a career may indicate regime shifts in productivity, regardless of which type of trajectory an individual may follow. While the change-point parameter t^* does not correlate with the other parameters of $f(t)$, its distribution reveals that for most faculty, the inferred change point in productivity rates occurs at approximately year 5. Fig. 5 translates each selected faculty member's career length and inferred change point into an ordered pair, creating a heat map of career change points. Shown in the accompanying marginal distribution, the modal value for t^* is year 5 with the median at 6 years, closely preceding tenure decisions at most institutions. Nevertheless, there is still rich diversity in career transitions, and the average remains misleading as the descriptor of a majority of individuals. In particular, faculty at the top 20% of institutions have significantly earlier t^* than the remaining 80%, with medians of 4.1 years and 6.4 years, respectively ($P < 0.001$, Mann–Whitney test). There is no such difference between the faculty whose doctorates are from the top 20% of institutions and those whose doctorates are from the remaining 80% (medians of 5.9 years vs. 6.0 years; $P = 0.37$).

The trends and diversity observed in t^* distributions remain true even when models are avoided entirely. A direct empirical examination of all DBLP and CV publication time series reveals that a computer science professor's productivity is also most likely to peak in the fifth year, yet peak productivity can nevertheless occur in any year of a professor's career (Fig. 6). While the marginal distribution shows that 41.9% of faculty have their peak productivity within the first 6 years, with the modal peak year in year 5, there is substantial variance. Note, for example, that individuals along the bottom of Fig. 6 published the most in their first year as faculty, while individuals along the diagonal published the most in their most recent recorded year as faculty.

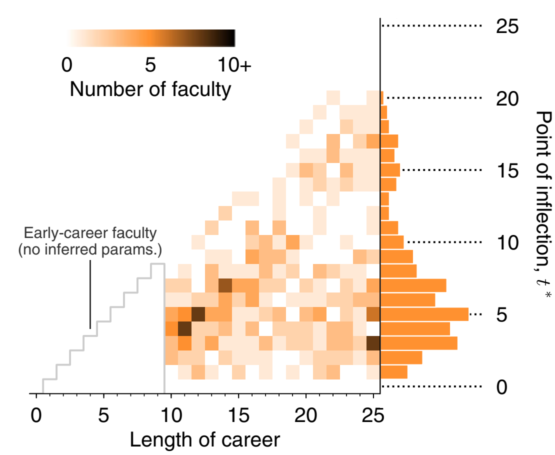

Fig. 5. Heat map of researchers' inferred change points. Each researcher's inferred change-point parameter t^* is plotted as a heat map, sorted by the length of their career in our dataset and restricted to individuals whose productivity trajectories are both stable under the addition of noise (main text) and better modeled by Eq. **1** than a straight line, determined by the AIC (*Model Selection*). params, parameters.

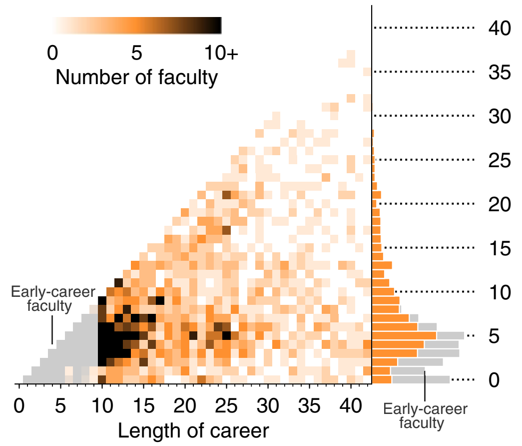

Fig. 6. Heat map of researchers' most productive years. Each researcher's most productive year (empirically; not model fit) is plotted as a heat map, sorted by the length of their career in our dataset. White box indicates researchers with fewer than 10 years of experience, whose most productive year is necessarily early. The marginal distribution (*Right*) shows the empirically most productive year for all faculty in the dataset, separated by early career (first 10 years; gray) or later career (orange). The most common peak-productivity year is year 5, and only about half of senior faculty exhibit peak productivity in year 5 or earlier.

Transitions in Authorship Roles. Finally, other transitions exist that are not quantifiable in publication counts alone, yet these are surprisingly well synchronized with the transitions noted above. As faculty train graduate students, their roles ordinarily shift from lead researcher to senior advisor or principal investigator, and this transition is commonly reflected in a shift from first author to last author. While common, this first/last convention is not universal. For example, papers in theoretical computer science typically order authors alphabetically, so the relative position of these researchers in the author list will not exhibit any consistent pattern over a career. To investigate career-stage transitions in author position, we first identified the set of journals or conferences that list authors alphabetically by computing whether each venue's authors are alphabetized significantly more often than is expected by chance ($\alpha = 0.05$) and exceeding twice the expected rate (*Detection of Alphabetized Publication Venues*). These conditions selected 11.2% of publication venues, accounting for 15.4% of all papers in the dataset, which we manually verified includes all top theoretical computer science conferences and excludes all top machine-learning and data-mining conferences. We then discarded these alphabetically biased venues from the following analysis. The remaining data show clear evidence of a progressive shift toward last-authorship position over time, with the relative first/last proportion reaching stability around year 8 (Fig. 7). Interestingly, the onset of this change is earlier among faculty at high-prestige institutions, and their average proportion of last-author papers is significantly higher than those of other faculty, consistent with a hypothesis that faculty at elite institutions tend to begin working with students earlier and have larger or more productive research groups.

As with the aggregate trend in productivity over a faculty career (Fig. 2), the transition from first- to last-author publications (Fig. 7) is based on averaging across many faculty and thus may not reflect the pattern of any particular individual. To characterize individual performances, we compared the fraction of first-authored papers in the first 3 years posthire to the same fraction in the second 3 years, for faculty with careers longer than 6 years ($N = 2{,}036$). A substantial drop in this fraction across these two periods would be consistent with the average trend reflecting individual patterns. For this analysis, we treated single-author papers as first-author publications. Overall 70.1% of researchers undergo this transition, publishing a larger fraction of first-author publications in the first 3 years of their faculty career than in the second 3 years. These fractions are consistent for faculty at top-50 institutions (70.2%) and those at other institutions (70.1%), but individuals at top-ranked institutions appear to make the transition more quickly and completely by the end of the 6-year period (Fig. 8). Despite these trends, there remains substantial diversity among first/last author transitions, reinforcing the notion that averages may be poor descriptors of many individuals.

Discussion

The conventional narrative of faculty productivity over a career is pervasive, with repeated findings reinforcing a canonical trajectory where productivity rises rapidly to a peak early in one's career and then declines slowly (5–11, 13, 14, 16, 17). This narrative shapes expectations of faculty across career stages, and publication counts have been shown to impact both tenure decisions (2, 3) and the ability to secure funding for future research (4). In this study, we showed that the conventional narrative, while intuitive and certainly applicable to averages of many professors, is a remarkably inaccurate description of most professors' trajectories. By applying a simple piecewise-linear model to a comprehensive dataset of academic appointment histories and publication records, we found that only about one-fifth of tenured or tenure-track computer science faculty resemble the average, regardless of their department's prestige.

While diverse, some aspects of a trajectory are nevertheless partially predictable. For example, although the diversity of trajectories remains unaffected, productivity does tend to scale with prestige: Researchers who graduated from or were hired by top-ranked institutions are significantly more productive at the onset of their careers, and, furthermore, productivity of high-prestige faculty tends to grow at faster rates and achieve higher peaks than that of researchers used by other institutions. Together, these results support previously suggested hypotheses that top-ranked universities both select for and facilitate productivity (27). In fact, our results suggest that the early-career transition to leadership roles, a phenomenon also found in other disciplines (36), takes place more quickly at top-ranked institutions, further implicating facilitation effects in addition to selection.

The relationship between productivity trajectories and gender is complicated and requires careful study. Gender has been shown to correlate with differences in productivity across fields

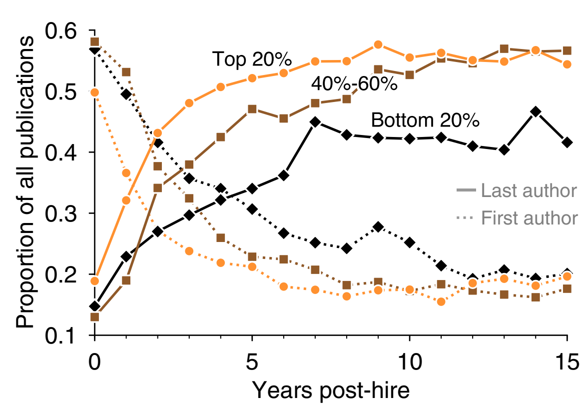

Fig. 7. Early-career transitions in authorship roles. Shown is the average proportion of first-author (dotted lines) and last-author (solid lines) papers as a fraction of the total, as a function of career age, separating researchers at institutions in the top, middle, and bottom quintiles according to prestige rank. Single-author publications are counted as first-author publications. On average, researchers at more prestigious institutions transition more quickly into senior-authorship roles.

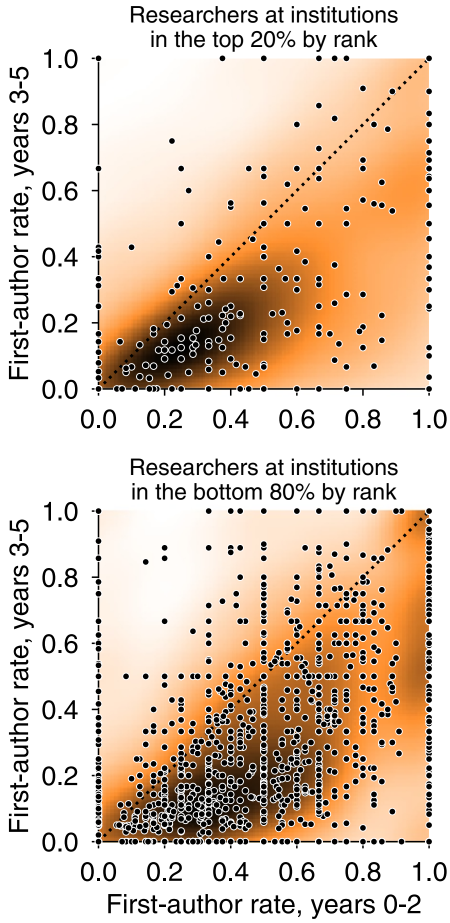

Fig. 8. First-author publication rates. First-author publications as a fraction of the total in the first 3 years posthire, and the 3 years thereafter, are shown separately for researchers who placed at an institution in the top 20% by rank (*Top*) and researchers placing outside of the top 20% (*Bottom*). Individual researcher data are plotted as points on top of a corresponding heat map in which darker color denotes higher density by Gaussian kernel-density estimation. Researchers at all levels of prestige tend to move out of first-authorship roles during this period, although researchers at more prestigious institutions transition more completely by years 3–5 than others.

(37–39), but these relationships are complicated by prestige (26) and have also changed over time (1). Other work has uncovered differences in collaboration patterns between subfields (40), as well as productivity differences that depend on both student and advisor genders (41). Here, we found that men and women follow the canonical productivity narrative at equal rates. However, among those who do, we found significant differences in initial and peak productivities between men and women. Given the complications revealed in past studies, the extent to which these differences reflect inequalities, past or present, and contribute to women's underrepresentation in computer science is an important topic of research and warrants future exploration.

Within the space of career trajectories, there is a noticeable tendency toward peak productivities and shifts in publication rates around 5 years after beginning as faculty. This is surely not a coincidence, given the fundamental role of tenure as a change point within the typical academic career, after which the total number of hours worked does not substantially change, but the time devoted to service tends to dramatically increase, with concomitant decreases in research and grant writing (42). However, our data cannot yet say how, from a mechanistic perspective, the existence of tenure requirements drives faculty to change or shape their productivity before or after promotion. If anything, the results in this paper make clear that there are numerous ways in which computer scientists meet promotion requirements, not all of which necessarily involve publishing a large number of papers. Indeed, in parallel with career shapes more broadly, there remains broad diversity in the distributions of productivity peaks and change points. This diversity in overall production, combined with the observation that an individual's highest-impact work is equally likely to be any of his or her publications (43), implies there are fundamental limits to predicting scientific careers (18).

Computer science is, itself, a multifaceted field, and previous studies of the DBLP dataset revealed that productivity rates differ by subfield (1). This observation, coupled with the menagerie of fluctuating trajectories revealed here, may suggest that year-to-year differences in individual trajectories are related to which subfields a researcher studies. Past work has revealed a first-mover advantage associated with entry into a rapidly growing field (44), so changes to individual research interests may contribute to noisy trajectories, particularly if they coincide with concentrated growth of popular new subfields.

Larger and higher-resolution datasets may improve our ability to identify expanding new subfields and other factors that could explain or predict trajectories. Although DBLP has the advantage of covering computer science journals and peer-reviewed conferences alike, we found that its coverage of those venues was incomplete in predictable ways. By manually collecting CV data for 10% of the scattered trajectories shown in Fig. 4, we adjusted DBLP data for missing publications and established the rate at which publishing rates have grown since 1970. Trajectories derived from DBLP data and benchmark CV data were statistically indistinguishable from each other. Investigations of productivity trajectories outside computer science will lack the field-specific DBLP database and may require additional calibration, name disambiguation, and data deduplication.

The misleading narrative of the canonical productivity trajectory is not likely to be unique to computer science. The rich diversity revealed here demands a reevaluation of the conventional narrative of careers across academia. Other studies that investigate the impact of this pervasive narrative on decisions of promotion, retention, and funding would be particularly valuable. Expectations, whether perceived or enforced through tenure decisions, might give rise to some of our results. If these expectations vary from field to field, it is possible that while diversity remains a feature that spans academia, some types of trajectories may be more common in certain fields. Regardless of whether this is borne out by studies of other fields, models of faculty productivity will need to be revisited and revised.

ACKNOWLEDGMENTS. The authors thank Mirta Galesic and Johan Ugander for helpful conversations. All authors were supported by National Science Foundation Award SMA 1633747; D.B.L. was also supported by the Santa Fe Institute Omidyar Fellowship.

1. Way SF, Larremore DB, Clauset A (2016) Gender, productivity, and prestige in computer science faculty hiring networks. *Proc 25th Intl Conf World Wide Web*, 1169–1179.
2. Cole S, Cole JR (1967) Scientific output and recognition: A study in the operation of the reward system in science. *Am Socio Rev* 32:377–390.
3. Massy WF, Wilger AK (1995) Improving productivity: What faculty think about it—and its effect on quality. *Change* 27:10–20.
4. Stephan PE (1996) The economics of science. *J Econ Lit* 34:1199–1235.
5. Lehman HC (1954) Men's creative production rate at different ages and in different countries. *Sci Mon* 78:321–326.
6. Dennis W (1956) Age and productivity among scientists. *Science* 123:724–725.
7. Diamond AM (1986) The life-cycle research productivity of mathematicians and scientists. *J Gerontol* 41:520–525.

8. Horner KL, Rushton JP, Vernon PA (1986) Relation between aging and research productivity of academic psychologists. *Psychol Aging* 1:319–324.
9. Bayer AE, Dutton JE (1977) Career age and research-professional activities of academic scientists: Tests of alternative nonlinear models and some implications for higher education faculty policies. *J Higher Educ* 48:259–282.
10. Simonton DK (1997) Creative productivity: A predictive and explanatory model of career trajectories and landmarks. *Psychol Rev* 104:66–89.
11. Levin SG, Stephan PE (1991) Research productivity over the life cycle: Evidence for academic scientists. *Am Econ Rev* 81:114–132.
12. Rinaldi S, Cordone R, Casagrandi R (2000) Instabilities in creative professions: A minimal model. *Nonlinear Dyn Psychol Life Sci* 4:255–273.
13. Cole S (1979) Age and scientific performance. *Am J Sociol* 84:958–977.
14. Symonds MR, Gemmell NJ, Braisher TL, Gorringe KL, Elgar MA (2006) Gender differences in publication output: Towards an unbiased metric of research performance. *PLoS One* 1:e127.
15. Malmgren R, Ottino J, Nunes AL (2010) The role of mentorship in protégé performance. *Nature* 465:622–626.
16. Quetelet LAJ (1835) *Sur l'Homme et le Développement de ses Facultés: Ou, Essai de Physique Sociale*; trans Knox R (2013) [*A Treatise on Man and the Development of his Faculties*], ed Simbert T (Cambridge Univ Press, New York).
17. Kaplan H, Hill K, Lancaster J, Hurtado AM (2000) A theory of human life history evolution: Diet, intelligence, and longevity. *Evol Anthropol* 9:156–185.
18. Clauset A, Larremore DB, Sinatra R (2017) Data-driven predictions in the science of science. *Science* 355:477–480.
19. Caplow T, McGee RJ (1958) *The Academic Marketplace* (Basic Books, Inc., New York).
20. Crane D (1965) Scientists at major and minor universities: A study of productivity and recognition. *Am Sociol Rev* 30:699–714.
21. Reskin BF (1979) Academic sponsorship and scientists' careers. *Sociol Educ* 52: 129–146.
22. Long JS, Allison PD, McGinnis R (1979) Entrance into the academic career. *Am Sociol Rev* 44:816–830.
23. Chubin DE, Porter AL, Boeckmann ME (1981) Career patterns of scientists: A case for complementary data. *Am Sociol Rev* 46:488–496.
24. Zuckerman HA (1970) Stratification in American science. *Sociol Inq* 40:235–257.
25. Long JS (1978) Productivity and academic position in the scientific career. *Am Sociol Rev* 43:889–908.
26. Clauset A, Arbesman S, Larremore DB (2015) Systematic inequality and hierarchy in faculty hiring networks. *Sci Adv* 1:e1400005.
27. Allison PD, Long JS (1990) Departmental effects on scientific productivity. *Am Sociol Rev* 55:469–478.
28. Pelz DC, Andrews FM (1976) *Scientists in Organizations: Productive Climates for Research and Development* (Institute for Social Research, Ann Arbor, MI).
29. Dey EL, Milem JF, Berger JB (1997) Changing patterns of publication productivity: Accumulative advantage or institutional isomorphism? *Sociol Educ* 70:308–323.
30. Larsen PO, Von Ins M (2010) The rate of growth in scientific publication and the decline in coverage provided by science citation index. *Scientometrics* 84:575–603.
31. Lotka AJ (1926) The frequency distribution of scientific productivity. *J Wash Acad Sci* 16:321–322.
32. Shockley W (1957) On the statistics of individual variations of productivity in research laboratories. *Proc IRE* 45:279–290.
33. Cole JR (1979) *Fair Science: Women in the Scientific Community* (Free Press, New York).
34. Reskin BF (1978) Scientific productivity, sex, and location in the institution of science. *Am J Sociol* 83:1235–1243.
35. Fox MF (1983) Publication productivity among scientists: A critical review. *Soc Stud Sci* 13:285–305.
36. Duch J, et al. (2012) The possible role of resource requirements and academic career-choice risk on gender differences in publication rate and impact. *PLoS One* 7: e51332.
37. Long JS (1992) Measures of sex differences in scientific productivity. *Soc Forces* 71: 159–178.
38. Xie Y, Shauman KA (1998) Sex differences in research productivity: New evidence about an old puzzle. *Am Sociol Rev* 63:847–870.
39. Fox MF (2005) Gender, family characteristics, and publication productivity among scientists. *Soc Stud Sci* 35:131–150.
40. Zeng XHT, et al. (2016) Differences in collaboration patterns across discipline, career stage, and gender. *PLoS Biol* 14:e1002573.
41. Pezzoni M, Mairesse J, Stephan P, Lane J (2016) Gender and the publication output of graduate students: A case study. *PLoS One* 11:e0145146.
42. Link AN, Swann CA, Bozeman B (2008) A time allocation study of university faculty. *Econ Educ Rev* 27:363–374.
43. Sinatra R, Wang D, Deville P, Song C, Barabási AL (2016) Quantifying the evolution of individual scientific impact. *Sci*ence 354:aaf5239.
44. Newman MEJ (2009) The first-mover advantage in scientific publication. *Europhys Lett* 86:68001.