Quantifying Lifetime Productivity Changes: A Longitudinal Study of 320,000 Late-Career Scientists

Marek Kwiek
(1) Center for Public Policy Studies (CPPS), Adam Mickiewicz University of Poznan, Poznan, Poland, and
(2) German Center for Higher Education Research and Science Studies (DZHW), Berlin, Germany
kwiekm@amu.edu.pl, ORCID: orcid.org/0000-0001-7953-1063, corresponding author

Lukasz Szymula
(1) Faculty of Mathematics and Computer Science, Adam Mickiewicz University of Poznan, Poznan, Poland, and

(2) Department of Computer Science, University of Colorado Boulder, USA
ORCID: orcid.org/0000-0001-8714-096X

Abstract

The present study focuses on persistence in research productivity over the course of an individual's entire scientific career. We track "late-career" scientists—scientists with at least 25 years of publishing experience (N=320,564)—in 16 STEMM (science, technology, engineering, mathematics, and medicine) and social science disciplines from 38 OECD countries for up to five decades. Our OECD sample includes 79.42% of late-career scientists globally. We examine the details of their mobility patterns as early-career, mid-career, and late-career scientists between decile-based productivity classes, from the bottom 10% to top 10% of the productivity distribution. Methodologically, we turn a large-scale bibliometric dataset (Scopus raw data) into a comprehensive, longitudinal data source for research on careers in science. The global science system is highly immobile: half of global top performers continue their careers as top performers and one-third of global bottom performers as bottom performers. Jumpers-Up and Droppers-Down are extremely rare in science. The chances of moving radically up or down in productivity classes are marginal (1% or less). Our regression analyses show that productivity classes are highly path dependent: there is a single most important predictor of being a top performer, which is being a top performer at an earlier career stage.

Introduction

The focus of the present study is persistence in top and bottom individual research productivity from a lifetime perspective—that is, over the course of an entire scientific career. We are tracking "late-career" scientists (N=320,564) from 38 OECD countries for up to five decades to examine their mobility patterns between decile-based productivity classes, from the bottom 10% to top 10%.

We turn a large-scale publication and citation bibliometric dataset (Scopus raw data) into a global, comprehensive, multidimensional, and longitudinal data source for research on careers in science; this practice follows several previous global studies on gender and self-citations (King et al., 2017), women in science (Huang et al., 2020; Larivière et al., 2013; Sugimoto & Larivière, 2023; West et al. 2013), global citation inequality (Nielsen & Andersen, 2021), continuing publishing core of global science (Ioannidis et al., 2018), collaboration with top scientists (Li et al., 2019), the nature of international collaboration (Wagner, 2018), or academic careers viewed from the perspective of science of science (Wang & Barabàsi, 2021). So far, however, individual publishing productivity has not been examined globally from a longitudinal perspective. Most importantly, we move from individual publications (and their properties) to individual scientists (and their characteristics) as a unit of analysis. We construct individual lifetime publication and citation histories for every late-career scientist in our sample, restricting our research to 16 STEMM (science, technology, engineering, mathematics, and medicine) and social science disciplines. In our context, "late-career" scientists are defined as scientists with at least 25 years of publishing experience.

Because our study is of a longitudinal nature, we use a global bibliometric dataset to study scientific careers and how they change over time (Menard, 2002; Rowland, 2014; Ruspini, 1999): The same individuals are tracked over the multiple decades of their publishing careers. In global academic career research, ever more datasets are currently tested (e.g., integrated datasets with administrative and biographical, commercial and noncommercial, national and global data; see, e.g., King et al., 2017; Larivière et al., 2013; Nielsen, 2021). In the present research, we test the usefulness of publication and citation metadata for examining the global science profession from a longitudinal perspective; these metadata are digital traces left by scientists throughout their professional lives (or as long as they keep publishing in academic journals). Digital traces (Liu et al., 2023; Salganik, 2018) allow for the emergence of a whole new multidisciplinary field of science of science (Clauset et al., 2017; Wang & Barabási, 2021; Zeng et al., 2017), hence allowing science career studies to radically move beyond traditional small-scale surveys and interviews (Hermanowicz, 2012; Leišytė & Dee, 2012). The digital traces left by scientists in global publication and citation datasets allow academic career researchers to change their focus from single national science systems to a global science system (Huang et al., 2020; King et al., 2017; Ni et al., 2021).

The present study explores mobility between the 10 individual productivity classes (constructed according to the 10 decile-based classes) throughout long academic careers, here encompassing early-, mid-, and late-career periods. Our initial hypotheses, which are based on research productivity literature (Allison et al., 1982; Fox, 1983; Turner & Mairesse, 2005), especially high research productivity literature focused on "top performers" and "prolific" scientists (Abramo et al., 2009; Aguinis & O'Boyle, 2014; Fox & Nikivincze, 2021; Kwiek, 2016; Li et al., 2019) are, first, that scientists are generally locked in within their productivity classes for years (Kelchtermans & Veugelers, 2013; Turner & Mairesse, 2005); second, we argue that the elite strata of highly productive scientists often continue their entire careers as being highly productive (Allison & Stewart, 1974; David, 1994); and, finally, we argue that radical changes in productivity classes, especially upward, although popular in narratives about academic careers, are highly improbable in practice because of the cumulative nature of the advantages and disadvantages in careers, as shown over the decades in the traditional sociology of science (Cole & Cole, 1973; DiPrete & Eirich, 2006; Merton, 1973).

The current study follows the research lines explored in the field of science of science, which provides data-driven insights into the inner workings of science (Wang & Barabási, 2021). A shift toward new digitalized data sources allows for the exploration of new questions about scientists (Liu et al., 2023). In this case, traditional cross-sectional studies can be complemented with longitudinal studies (Lutter &

Schröder, 2016; Ma et al., 2020) in which individuals are tracked over time. The career histories of thousands of individual scientists can change the way we think about science and scientists because of the unprecedented level of detail that can be obtained. As a result, the various aspects of academic careers have recently been examined both globally (gender disparities in careers Larivière et al., 2013; Huang et al., 2020; continuous publishing Ioannidis et al., 2014; collaboration with top scientists Li et al., 2019; gendered nature of authorship Ni et al., 2021; women in science Sugimoto & Larivière, 2023) and nationally, especially in the US (e.g., productivity across career stages Way et al., 2017; long-term effects on careers of initial setbacks Wang et al., 2019; careers in elite universities Zhang et al., 2022; credit distribution in academic publishing Ross et al., 2022) at a scale unthinkable in career studies before.

Productivity Classes in Single-Nation Studies

We have tested our initial hypotheses in a previous national-level strand of research under the general label of "once highly productive, forever highly productive" (Kwiek & Roszka, 2024a, 2024b). The patterns found about OECD scientists consistently supported our initial intuitions about immobility in the system based on our single-nation research: The majority of highly productive early-career scientists and mid-career scientists continued their careers as highly productive mid-career scientists and late-career scientists; radical upward and downward mobility was either at zero or at marginal levels. In the present research, we develop our methodological approach to study mobility between the productivity classes of scientists from 38 OECD countries that are often powerfully involved in international collaboration as part of the ongoing globalization of science (Kwiek, 2023; Marginson, 2022). We track 320,564 late-career scientists from a wide variety of research systems, giving us the potential to test hypotheses about the scientific profession more generally.

There are three small-scale longitudinal single-nation studies similar to ours. First, for 497 French physicists, Turner and Mairesse (2005) showed that 66% of the most productive researchers (defined as quartile 1 scientists) and 67% of the least productive researchers (defined as quartile 4 scientists) remained as such for the period 1986–1997, underlying a stability of the relative positions of the researchers in the distribution of publication counts over time. Second, in a study of a single Belgian university, Kelchtermans and Veugelers (2013) discussed top research productivity and its persistence over time by using a panel dataset comprising the publications of 1,040 biomedical and exact scientists for the period 1992–2001; they studied how researchers switch between productivity categories over time, showing strong support for an accumulative process. However, this places scientists with a low initial output at a disadvantage while giving highly productive scientists a greater advantage. Finally, Abramo et al. (2017) studied Italian scientists in three consecutive four-year periods of 2001–2012; they identified 2,883 top performers in the first period and followed them over time. About one-third of top performers retained their top ranking for three consecutive periods, and about half retained it for two periods (35% and 55%, respectively).

Our research explores a different scale, scope, and methodology: We track a large number of late-career scientists from 38 OECD countries from all science sectors (including higher education); we examine productivity changes over a prolonged period of time (25–50 years) across all STEMM and selected social science disciplines; and we use a longitudinal and classificatory approach combined with two-dimensional analyses and logistic regression models.

Persistence in High (and Low) Productivity

Productive scientists are likely to be "even more productive in the future, while scientists who produce little original work are likely to decline further in their productivity" (Allison & Stewart, 1974: 596). Substantial predetermined differences among scientists may have a powerful impact on careers (Cole & Cole, 1973; Fox, 1983). An "initial success" may lead to increased productivity; in contrast, a "bad start" may lead to the scientist leaving science (Turner & Mairesse, 2005). In other words, "there are substantial, predetermined differences among scientists in their ability and motivation to do creative scientific research" (Allison & Stewart, 1974: 596). Some scientists are always very productive, and a differential distribution of talent affects inequality in productivity more than the recognition system in science (Stephan & Levin, 1992).

As a result, productivity stratification in science leads to "persistent hierarchies of productivity": "once scientists enter the current productivity elite, it is rare for them to exit from it in the next period; and the same holds true at the lower extreme of the productivity distribution" (David, 1994). Top performers tend to try hard not to disappoint their colleagues and themselves; bottom performers, in contrast, tend to lose confidence in their research capabilities. Previous top performance significantly and positively affects current top performance (Kelchtermans & Veugelers, 2013).

We realize that productivity is a narrow measure of scientists' success: What is much more important in the long term is scholarly impact. Citations tell more about individual academic success than publications and publishing patterns and intensity. However, the focus in the present paper is on changing productivity classes rather than changing impact classes (which would also be possible based on individual-level, field-weighted citation impact (FWCI) of all articles published within specific career periods); in the current research, our reference to impact based on citations is twofold: through Scopus journal citation ranks (range: 0–100) in productivity computations and through the variable of FWCI 4y in regression models (or the impact of every publication within four-year windows).

Research Questions

At a global scale, we quantify the persistence in research productivity over an academic lifetime by following previous research (using analyses based on small-scale surveys and a limited number of interviews) across a span of decades (Allison, 1980; Allison et al., 1982; Cole & Cole, 1973; Merton, 1973). Tracking the career trajectories of thousands of scientists, we seek otherwise invisible, global mobility patterns between research productivity classes (whenever we use the term "global," we refer to 38 OECD countries). Our sample of OECD late-career scientists includes 79.42% of all late-career scientists globally (from all countries, N=403,653); and their research output (30,695,679 research articles) includes 83.03% of all research articles produced by this category of scientists globally (N=36,969,473).

We have posed the following research questions regarding changing publishing productivity over the course of individuals' academic careers: First, what is the scale of horizontal transitions (top to top, bottom to bottom) and radical vertical transitions (bottom to top, top to bottom) between global productivity classes? Second, what is the scale of jumping up (and dropping down) in science in terms of research productivity—radically changing productivity classes upward or downward globally? Third, what are the cross-disciplinary differences in mobility patterns between global productivity classes? Finally, what are the predictors of belonging to the classes of highly productive and bottom productive scientists (the top 10%, the bottom 10%), and how do they differ between academic disciplines?

Methods

The data were collected from the Scopus bibliometric database and were obtained through a multiyear collaborative agreement with the International Center for the Study of Research (ICSR) Lab, a cloud computing platform provided for research purposes by Elsevier. Our final sample included all late-career scientists who were research active in 2023 (with at least 25 years of publishing experience) located in 16 STEMM and social science disciplines and coming from 38 OECD countries (N=320,564 scientists with N=16,345,891 research articles; see the major steps in data preprocessing in Figure 1; the "authors" in our dataset were defined as having publications of any type and scientists as authors with articles in journals and conference proceedings only. For our calculations, we utilized the Scopus database dated March 29, 2024. Our sample (Supplementary Tables 1 and 2) included 12,585 social scientists (from BUS, ECON, and PSYCH) and 307,979 STEMM scientists, the latter comprising 95.76% of our sample. About a quarter of our sample included women scientists (26.34%), the number of which was slightly more in the social sciences than STEMM fields. The largest academic discipline represented in our sample was MED (40.89%), followed by BIO (14.29%) and PHYS (9.13%). The percentage of women in our sample was about one-third in three disciplines (the two largest: MED and BIO) and exceeded 40% in only one (PSYCH: 41.44%). The three largest countries represented were the USA, Japan, and Italy, comprising about a half of all scientists in the sample. In terms of academic age, there are about 20,000–25,000 scientists in the youngest cohorts and about 2,000–3,000 in the oldest cohorts, with the share of women decreasing with each subsequent cohort, from about one-third for the youngest cohort (25 years of academic experience: 32.77%) to 13–15% for the oldest cohorts. In practical terms, we are working with the census of a population with clearly defined inclusion criteria rather than with a sample of scientists – with methodological implications for testing for statistical significance (which are not needed in the present research). The probability that the observed relationships and differences occurred by pure chance are zero because we do not work with samples drawn from the population of late-career scientists but with their population.

To achieve aggregate-level results, the ICSR Lab employed the Databricks environment, which facilitates the management and execution of cloud computing with Amazon EC2 services. The scripts for generating the results were developed using the PySparkSQL library. Runs were carried out using a cluster in standard mode with Databricks Runtime version 11.2 ML, Apache Spark technology version 3.3.0, Scala 2.12, and an i3.2xlarge instance with 61 GB memory, eight cores, one to six workers for the worker type, and a c4.2xlarge instance with 15 GB memory and four cores for the driver type. The execution time took six hours, and this operation was initiated on June 25, 2024. We obtained the results in CSV format.

The academic lives of all late-career scientists from 38 OECD countries that were research active in 2023 were retrospectively divided into three stages: early-, mid-, and late-career stages. All late-career scientists, by definition, were initially both early-career scientists (in their publishing years 5–14) and mid-career scientists (in their publishing years 15–24). We analyzed their current five-year publishing behavior (2019–2023) and looked back into their past publishing behavior to examine how they may have changed their productivity classes.

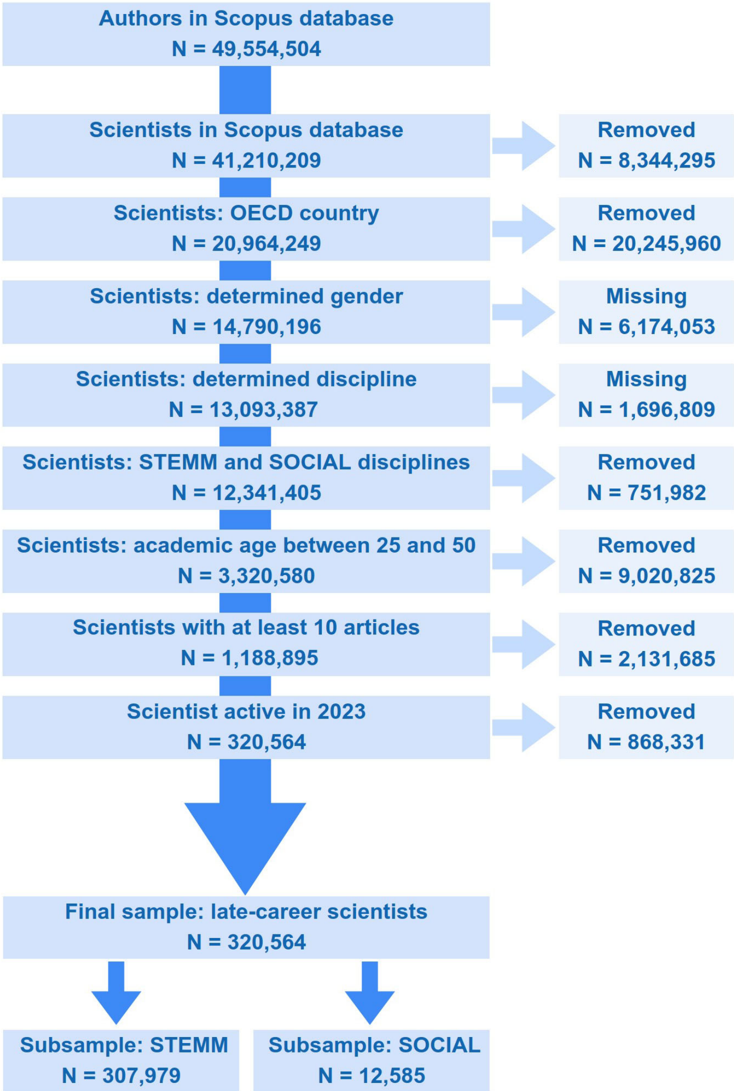

Figure 1. Flowchart and major steps in data preprocessing: from all scientists in the Scopus database to late-career scientists in our sample.

At each career stage, current late-career scientists showed their annual individual productivity. Consequently, their productivity was calculated for the recent five-year period and for two earlier periods: when they were early-career scientists and mid-career scientists. Our analyses are based on the idea of subsequent distributions of scientists into classes: Late-career scientists are first distributed by current productivity classes (separately within each of the 16 disciplines) and then, retrospectively, by past productivity classes in the two earlier career periods.

Early-career scientists may retain or change their decile-based productivity classes while being mid-career scientists as mid-career scientists may do while being late-career scientists. In the present study, we tracked scientists for 25–50 years and compared their productivity with the productivity of their peers (the same academic career stage and the same discipline).

For each scientist in our sample, an individual publication and citation portfolio was constructed. The portfolio included Scopus-derived publication metadata and their various constructs that accompanied individual authors from their first publication in the dataset to the year 2023. Within the portfolios, all metadata and their specially computed constructs were linked to the three career periods (e.g., annual productivity), individual publications (e.g., field-weighted four-year citation impact), or the entire lifetime careers of scientists (e.g., gender, discipline, international collaboration rate, and median team size) (see Variables in Table 9).

Our approach to individual research productivity is longitudinal (Menard, 2002; Rowland, 2014; Ruspini, 1999) and classificatory (or class based) (Costas & Bardons, 2007; Costas et al., 2010). First, we tracked the productivity of late-career scientists as individuals ever since they became early-career scientists, that is, five years after their first globally indexed publication. Second, we did not compare productivity changes over time (as individual scientific careers develop) in terms of changing publication numbers—we compared productivity in terms of the stable or changing membership in the productivity classes while scientists grow older and move up the professional ladder. Scientists can always be allocated to the top and bottom classes, regardless of actual publication numbers, cut-off points permitting, so that both terms could be used not to judge the level of productivity but rather classify it.

We used a journal prestige–normalized, full counting method of calculating productivity. This approach refers to the quantity and quality of globally indexed publications at the level of individuals (as opposed to quantity only in non-normalized approaches). Prestige normalization refers to journal percentile ranks used in the Scopus database (CiteScore ranking, range: 1–99), and it highlights the difference in average scholarly efforts between preparing and revising publications in generally less selective and more selective journals, here with different peer review procedures and acceptance rates. Prestige normalization is determined by the number of citations received by the journal (43,092 journals in 2024) in the previous four years. In a prestige-normalized approach, the weight of publications depends on their location in a vertically stratified system of academic journals (for more on the role of journal stratification in academic careers, see Hammarfelt, 2017; Heckman & Moktan, 2018; Kwiek 2021; Lindahl, 2018; Shibayama & Baba, 2015).

Our focus is on scientific careers rather than on publications. Therefore, the unit of analysis is individual scientists, with their unambiguously defined individual publication- and citation-related attributes (rather than publications, with their properties). A global publication-focused bibliometric dataset (raw Scopus dataset owned by Elsevier) was used to define individual attributes of all scientists in our sample. The productivity classes of individuals were traced over their lifetime—as early-, mid- and late-career scientists.

In the present research, we have used a global bibliometric dataset to define scientists' individual attributes. The determination of some attributes has already been described in detail in our previous research: gender determination (binary: male or female), discipline determination (using all cited references from all publications, lifetime), determining the country of affiliation (using a modal value of all affiliations in all publications, lifetime), determination of scientists' nonoccasional status in global science (using a minimum output of 10 research articles), and determining academic age (using the distance in years between the first publication, of any type, and 2023; Kwiek & Szymula, 2023, 2024). Three other individual attributes were used in individual publication and citation portfolios (their construction is described in Table 9): international collaboration rate (lifetime), field-weighted four-year citation impact (FWCI 4y), and median team size (lifetime). The distribution of the sample by academic age (i.e., publishing experience) is shown in Figure 2, with further details in Supplementary Table 2.

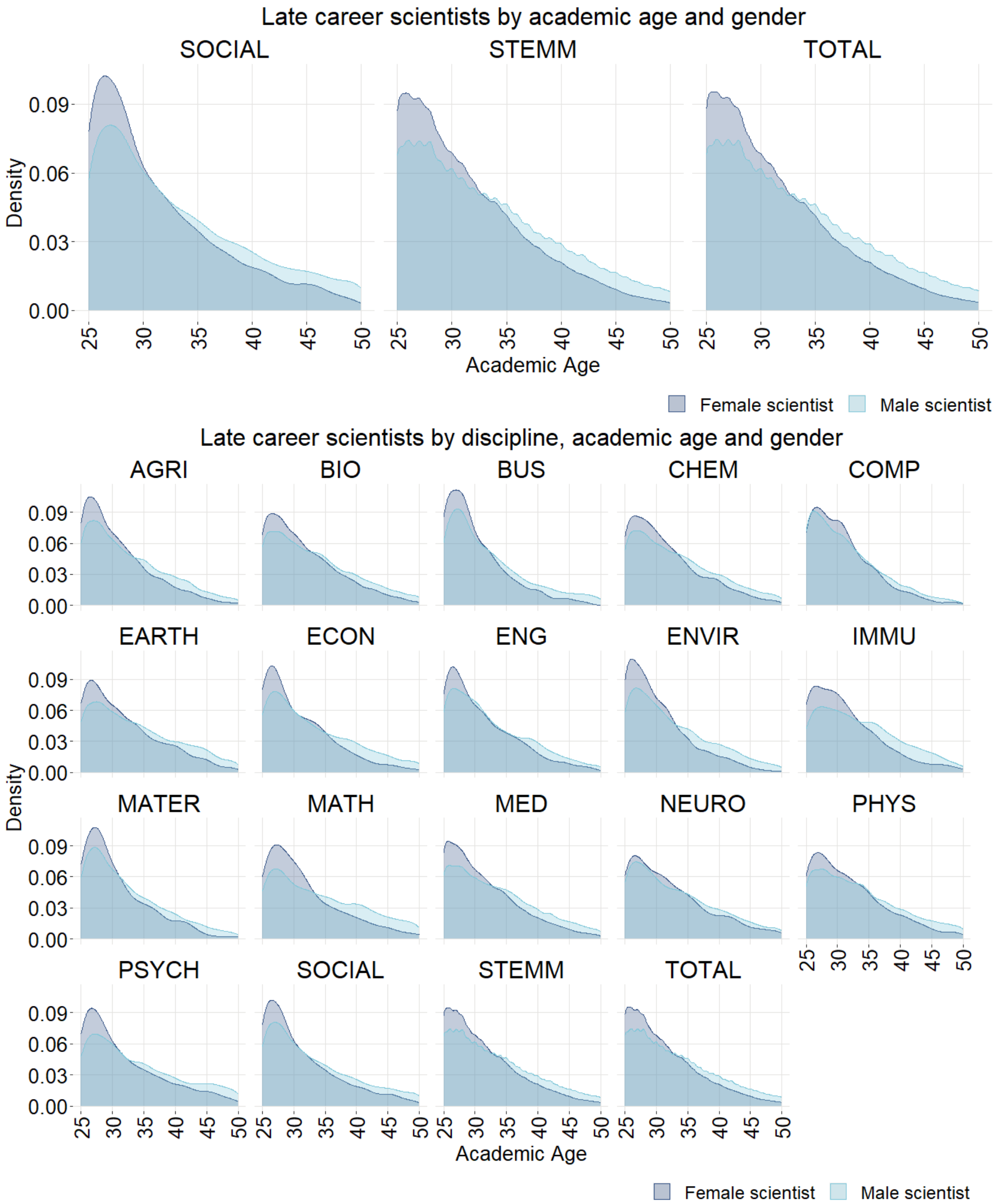

Figure 2. Distribution of academic age: Kernel density plots. Late-career scientists, all academic disciplines combined (top panel) by gender. Late-career scientists by academic discipline (bottom panel) and gender (N = 320,564)

Results

Changing Productivity Classes Over Academic Careers

Our focus is on analyzing the mobility between productivity classes, particularly the transitions between the top and bottom classes and the adjacent classes: productivity deciles 8, 9, and 10 at the top and deciles 1, 2, and 3 at the bottom of the productivity distribution. We examine three stages of academic careers—early career, mid-career, and late career—and the transitions from early career to mid-career and from mid-career to late career.

Early-career scientists in the top and bottom productivity classes can change their productivity classes as they progress to the mid-career stage, moving to any decile. Similarly, mid-career scientists in the top and bottom productivity classes can experience changes in their productivity levels as they transition to the late-career stage, moving up, down, or staying in the same productivity decile. We want to understand how productivity can evolve over the course of an academic career and discuss the extent to which individuals move between different levels of productivity.

We examine three primary types of mobility (across all academic disciplines and within specific disciplines):

1. *Top-to-top mobility*: Early-career scientists who are in the highest productivity decile remain in the highest decile as they progress to the mid-career stage, and similarly, from the mid-career to late-career stages. This reflects consistency in high productivity from one career stage to the next (mobility from decile 10 to decile 10).

2. *Bottom-to-bottom mobility*: Early-career scientists in the lowest productivity decile stay in the lowest decile as they advance to the mid-career stage and, likewise, from the mid-career to late-career stages. This indicates a persistent low productivity level across career stages (mobility from decile 1 to decile 1).

3. *Extreme downward and extreme upward mobility* (*top-to-bottom mobility* and *bottom-to-top mobility*): This includes both downward and upward mobility. Scientists who start in the top productivity decile in their early careers and who drop to the bottom decile by the mid- or late-career stage (top-to-bottom mobility); and scientists who begin in the lowest decile and rise to the highest by mid- or late-career stage (bottom-to-top mobility). This represents significant shifts in productivity, either from decile 10 to decile 1 or from decile 1 to decile 10.

In addition to examining the basic mobility between the highest (decile 10) and lowest (decile 1) productivity deciles, we will also explore a broader perspective of mobility that considers the transition between the upper deciles (8–10) and lower deciles (1–3). Some scientists are positioned just above the decile 1 threshold and others just below the decile 10 threshold (as illustrated in Table 1 for late-career scientists; Supplementary Tables 3 and 4 provide cut-off points for late-career scientists at the early- and mid-career stages). A more comprehensive approach that includes adjacent deciles (1–3 and 8–10) seems useful.

Table 1. Cut-off points (publication numbers: articles and chapters in conference proceedings) for membership in the productivity deciles, late-career scientists at a late-career stage, by discipline (N = 320,564)

Discipline	**Min**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Max**
AGRI	0.00	0.41	0.63	0.84	1.08	1.39	1.77	2.27	3.02	4.42	101.78
BIO	0.00	0.51	0.72	0.90	1.12	1.40	1.74	2.19	2.91	4.26	61.31
BUS	0.00	0.42	0.60	0.75	0.90	1.01	1.26	1.57	2.00	2.78	18.22
CHEM	0.00	0.48	0.73	0.98	1.29	1.66	2.17	2.83	3.82	5.84	69.92
COMP	0.00	0.36	0.58	0.75	0.92	1.16	1.46	1.83	2.39	3.59	74.89
EARTH	0.00	0.50	0.73	0.95	1.23	1.57	1.94	2.49	3.32	4.89	81.94
ECON	0.00	0.25	0.39	0.53	0.65	0.79	0.95	1.16	1.52	2.12	25.40
ENG	0.00	0.33	0.55	0.76	0.96	1.25	1.61	2.10	2.89	4.51	56.65
ENVIR	0.00	0.48	0.74	0.97	1.25	1.57	2.01	2.58	3.43	5.18	51.51
IMMU	0.00	0.48	0.74	0.95	1.22	1.55	1.92	2.41	3.31	4.98	87.91
MATER	0.00	0.45	0.72	0.99	1.32	1.74	2.30	3.05	4.18	6.52	57.61
MATH	0.00	0.19	0.34	0.47	0.60	0.74	0.92	1.16	1.52	2.27	171.49
MED	0.00	0.41	0.65	0.90	1.19	1.59	2.11	2.86	4.08	6.53	178.31
NEURO	0.00	0.47	0.68	0.88	1.10	1.35	1.66	2.10	2.79	4.09	57.73
PHYS	0.00	0.45	0.70	0.96	1.29	1.69	2.25	3.02	4.34	7.88	272.31
PSYCH	0.00	0.39	0.61	0.80	1.03	1.34	1.70	2.19	2.88	4.26	87.52
SOCIAL	**0.00**	**0.25**	**0.39**	**0.53**	**0.65**	**0.79**	**0.95**	**1.16**	**1.52**	**2.12**	**18.22**
STEMM	**0.00**	**0.19**	**0.34**	**0.47**	**0.60**	**0.74**	**0.92**	**1.16**	**1.52**	**2.27**	**51.51**
TOTAL	**0.00**	**0.19**	**0.34**	**0.47**	**0.60**	**0.74**	**0.92**	**1.16**	**1.52**	**2.12**	**18.22**

Our general question is how top-performing (productivity decile 10, N=32,063) mid-career scientists were distributed by productivity percentile ranks (range: 0–100) when they were early-career scientists in the past and, analogously, how top-performing (productivity decile 10, N=32,063) late-career scientists were distributed by productivity percentile ranks when they were mid-career scientists.

In addition, we are interested in how bottom-performing (productivity decile 1, N=32,063) mid-career scientists were distributed by productivity percentile ranks (range: 0–100) when they were early-career scientists in the past. In addition, analogously, how the current bottom-performing (productivity decile 1, N=32,075) late-career scientists were distributed by productivity percentile ranks when they were mid-career scientists. In all cases studied, we retrospectively examine current late-career scientists (who make up our sample): when they were early-career scientists and when they were mid-career scientists.

In the mobility from the early-career to mid-career stage, as could be expected, the median value of the original percentile rank (as early-career scientist) is very close to the target percentile rank (as mid-career scientist): The median is the 90th percentile for global top performers and the 15th percentile for global bottom performers (Table 2), with limited discipline-related variability for top performers (from 89th in COMP, ENG, and IMMU to 92nd in MATH) and slightly higher discipline-related variability for bottom performers (from 12th in PHYS to 17th in COMP and ECON). Differences in both the median and mean values between social science academic disciplines combined (SOCIAL) and STEMM disciplines combined are limited.

Table 2. How top performers at a mid-career stage (productivity decile 10) (Left panel) and bottom performers at a mid-career stage (productivity decile 1) (Right panel) were distributed by productivity percentiles (range: 0–100) when they were in their early-career stage. Top/bottom performers in their mid-career stage, initial (as early-career stage) percentile distribution statistics by academic discipline (N_{top} = 32,063, N_{bottom} = 32,063)

	Top performers at a mid-career stage – distribution at an early-career stage				Bottom performers at a mid-career stage - distribution at an early-career stage			
Discipline	**N**	**Mean**	**Std dev**	**Median**	**N**	**Mean**	**Std dev**	**Median**
AGRI	2,373	85.03	15.84	**91**	2,373	18.66	17.13	**14**
BIO	4,582	82.95	18.69	**90**	4,582	21.05	19.58	**15**
BUS	326	82.00	19.74	**89.5**	326	24.43	21.44	**18**
CHEM	1,490	84.52	17.55	**91**	1,490	17.57	16.68	**13**
COMP	765	80.80	20.73	**89**	765	23.12	20.07	**17**
EARTH	1,437	85.01	16.27	**91**	1,437	18.32	17.48	**13**
ECON	385	82.32	19.85	**90**	385	22.30	19.20	**17**
ENG	1,282	82.40	18.84	**89**	1,282	20.85	19.02	**15**
ENVIR	652	84.82	16.78	**91**	652	20.99	19.48	**15.5**
IMMU	315	82.28	19.21	**89**	315	21.24	20.46	**14**
MATER	584	83.26	17.91	**90**	584	17.89	16.18	**14**
MATH	701	85.58	16.79	**92**	701	19.90	18.01	**15**
MED	13,108	83.97	17.04	**90**	13,108	20.15	18.49	**15**
NEURO	587	83.72	18.50	**90**	587	19.90	18.88	**14**
PHYS	2,928	83.25	20.90	**92**	2,928	18.49	19.15	**12**
PSYCH	548	83.60	18.35	**91**	548	18.79	17.74	**13**
SOCIAL	**1,259**	**82.79**	**19.17**	**90**	**1,259**	**21.32**	**19.14**	**17**
STEMM	**30,804**	**83.78**	**17.77**	**90**	**30,804**	**19.88**	**18.53**	**15**
TOTAL	**32,063**	**83.74**	**17.83**	**90**	**32,063**	**19.94**	**18.55**	**15**

Similarly, regarding the mobility from the mid-career to late-career stage, the median value of the original percentile rank (as mid-career scientist) for the current late-career top performers is very close to the target percentile rank (as late-career scientist): The median is the 90th percentile for top performers and the 19th percentile for bottom performers (Table 3), with limited discipline-related variability for top performers (from 89th in BUS to 92nd in CHEM) and slightly higher discipline-related variability for bottom performers (from 15th in PSYCH and EARTH to 21st in BIO and ECON). Again, the differences in both the median and mean values between social science academic disciplines (SOCIAL) and STEMM academic disciplines combined are limited.

A useful way to visualize the distribution of current top-performing and bottom-performing late-career scientists across productivity deciles during their mid-career and early-career stages is presented through kernel density plots (Figures 3 and 4). These plots utilize kernel density estimation to generate a smooth, continuous curve that represents the underlying data distribution. Unlike histograms, kernel density plots are not influenced by the number of bins or significant differences between them, making them more effective in illustrating the shape of a distribution; they also allow for a more flexible comparison between multiple datasets. When considering all academic disciplines combined (TOTAL), the majority of top performers were previously in productivity deciles 8 through 10, while most bottom performers were previously in deciles 1 through 3. Notably, the highest concentration of top-performing (as well as bottom-performing) mid-career scientists can be found in PHYS and CHEM.

Table 3. How top performers in their late-career stage (productivity decile 10) (Left panel) and bottom performers in their late-career stage (productivity decile 1) (Right panel) were distributed by productivity percentiles (range: 0–100) when they were in their mid-career stage. Top/bottom performers in a late-career stage, initial (as mid-career stage) percentile distribution statistics by academic discipline (N_{top} = 32,063, N_{bottom} = 32,075)

	Top performers at a late-career stage - distribution at a mid-career stage				**Bottom performers at a late-career stage - distribution at a mid-career stage**			
Discipline	**N**	**Mean**	**Std dev**	**Median**	**N**	**Mean**	**Std dev**	**Median**
AGRI	2,373	84.34	17.98	**91**	2,373	21.90	19.08	**17**
BIO	4,582	82.97	19.52	**90**	4,582	26.37	22.31	**21**
BUS	326	80.33	21.99	**89**	326	25.52	20.82	**20**
CHEM	1,490	85.20	17.92	**92**	1,490	21.92	19.39	**16**
COMP	765	81.76	21.08	**90**	767	24.55	20.68	**19**
EARTH	1,437	82.50	19.38	**90**	1,437	21.70	19.70	**15**
ECON	385	80.38	22.31	**90**	385	25.44	20.76	**21**
ENG	1,282	83.06	19.74	**91**	1,282	23.38	19.26	**19**
ENVIR	652	83.58	19.25	**91**	652	24.69	20.37	**19**
IMMU	315	82.78	19.81	**90**	319	24.12	22.07	**17**
MATER	584	85.03	17.84	**91**	590	22.17	18.93	**17**
MATH	701	84.21	19.00	**91**	701	22.07	19.02	**18**
MED	13,108	83.01	19.61	**90**	13,108	24.50	21.09	**19**
NEURO	587	84.53	18.35	**91**	587	24.55	21.75	**18**
PHYS	2,928	83.28	21.94	**92**	2,928	22.37	20.84	**16**
PSYCH	548	83.84	18.10	**91**	548	21.29	19.69	**15**
SOCIAL	**1,259**	**81.87**	**20.39**	**90**	**1,259**	**23.65**	**20.31**	**20**
STEMM	**30,804**	**83.29**	**19.56**	**90**	**30,816**	**23.98**	**20.78**	**19**
TOTAL	**32,063**	**83.23**	**19.60**	**90**	**32,075**	**23.96**	**20.76**	**19**

With our dataset, we can analyze the mobility between productivity deciles (at the level of individuals) in great detail. Table 4 shows initial productivity deciles (as early-career scientists) in the past of top-performing mid-career scientists across the various academic disciplines.

Over half of these top mid-career scientists were in productivity decile 10 in their early-career stage (52.39%), with 20.94% starting off in decile 9 and 10.33% decile 8. Altogether, more than 80% were in productivity deciles 8–10 during their early-career stage (83.66%). Only a small fraction of these scientists moved up from the lowest three deciles, with just 162 making a significant leap from decile 1 to decile 10 (0.51%, referred to as Jumpers-Up) and 232 from decile 2 to decile 10 (0.72%). We have full lifetime biographical and publishing profiles of every scientist, including these few hundreds of outliers. Overall, only 2.2% (717 scientists out of 32,063) from deciles 1–3 reached decile 10. In social science disciplines, the likelihood of such extreme upward mobility was slightly higher compared with STEMM disciplines (3.13% vs. 2.22%) but still relatively rare.

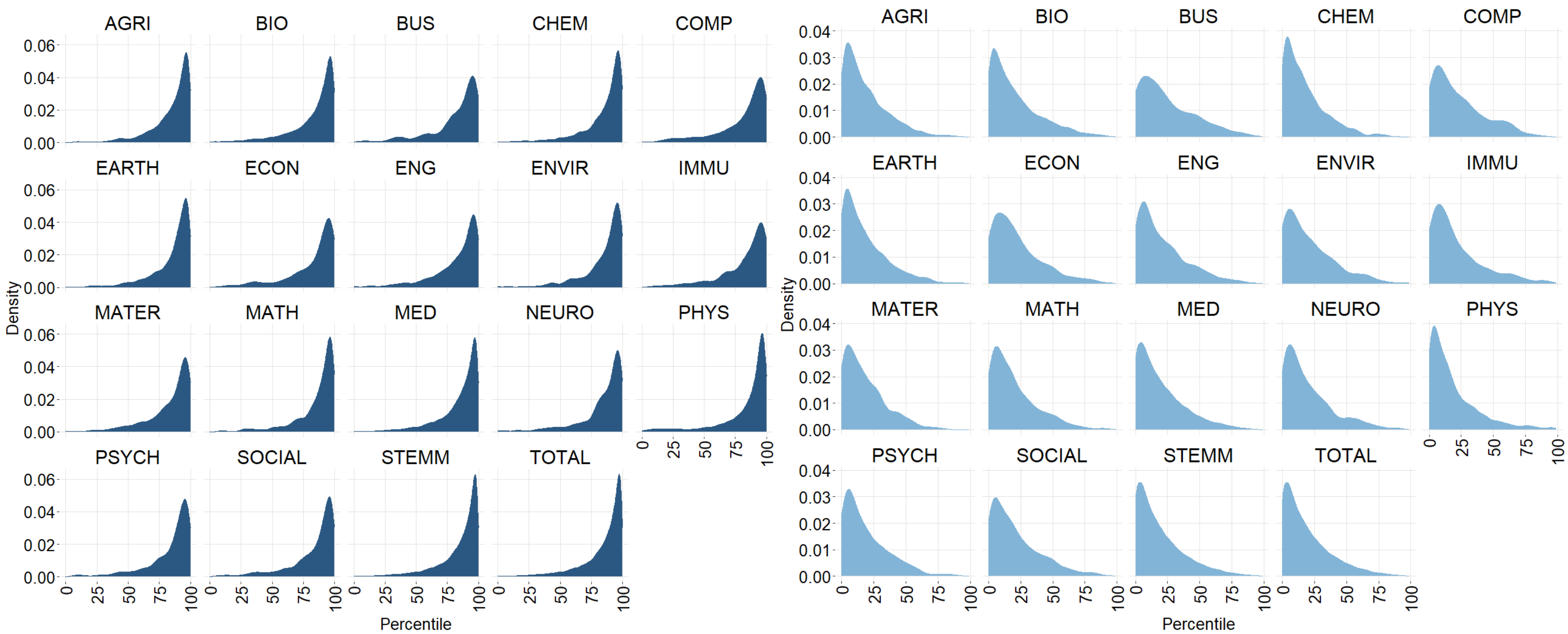

Figure 3. (Left panel) How were the top-performing (N=32,063, productivity decile 10) mid-career scientists distributed by productivity percentiles (range: 0–100) when they were in their early-career stage? (Right panel) How were the bottom-performing (N=32,063, productivity decile 1) mid-career scientists distributed by productivity percentiles (range: 0–100) when they were in their early-career stage? Kernel density plots, initial percentile distribution, by discipline.

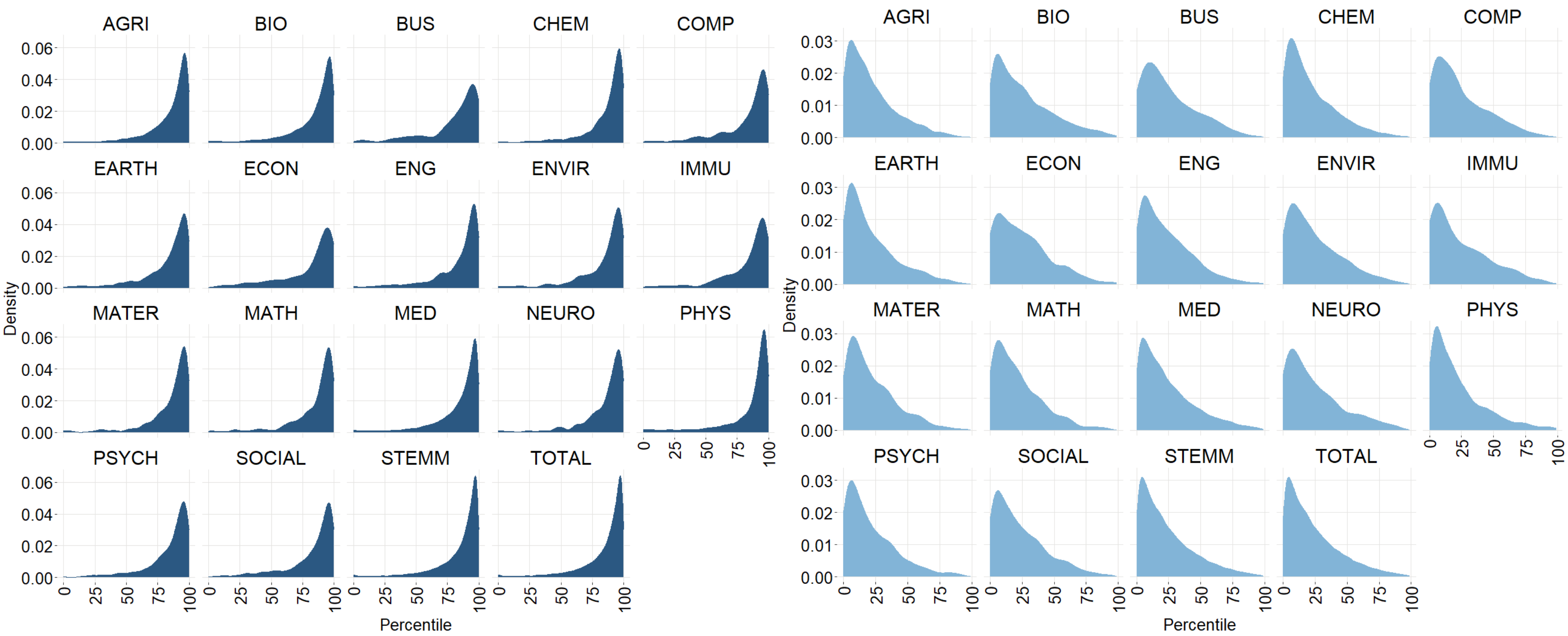

Figure 4. (Left panel) How were the top-performing (N=32,063, productivity decile 10) late-career scientists distributed by the productivity percentiles (range: 0–100) when they were in their mid-career stage? (Right panel) How were the bottom-performing (N=32,075, productivity decile 1) late-career scientists distributed by productivity percentiles (range: 0–100) when they were in their mid-career stage? Kernel density plots, initial percentile distribution, by discipline.

The variation in extreme mobility from decile 1 to decile 10 (Jumpers-Up) across disciplines is significant, with rates ranging from 0.26% in ECON to 1.33% in PHYS. Only one economist (0.26%) and one immunologist (0.32%) made the leap from decile 1 to decile 10 (out of 385 and 315, respectively). We know a high number of details of their specific academic careers based on our bibliometric data (but not their scientific biographies based on administrative data from national registries of scientists as in single-nation studies—not available for a multicountry study). The mobility from decile 10 to decile 10 shows variability as well, with less than 50% of scientists in COMP, ENG, and IMMU remaining in decile 10 compared with 57–58% in MATH and PHYS.

Table 4. Mobility of top performers between two career stages: early career (initial stage) and mid-career (target stage): From which initial productivity deciles (at an early-career stage) do top-performing scientists at a mid-career stage come from? Late-career scientists who were top performers at a mid-career stage (N=32,063) by academic discipline and initial productivity decile (frequencies and percentages)

		Total	Bottom 10%	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Top 10%
Late-career scientists who were top performers at a mid-career stage												
TOTAL	**N**	**32,063**	**162**	**232**	**323**	**540**	**799**	**1,175**	**2,007**	**3,312**	**6,714**	**16,799**
	%	**100**	**0.51**	**0.72**	**1.01**	**1.68**	**2.49**	**3.66**	**6.26**	**10.33**	**20.94**	**52.39**
SOCIAL	**N**	**1,259**	**7**	**14**	**12**	**35**	**35**	**48**	**65**	**149**	**234**	**660**
	%	**100**	**0.56**	**1.11**	**0.95**	**2.78**	**2.78**	**3.81**	**5.16**	**11.83**	**18.59**	**52.42**
STEMM	**N**	**30,804**	**155**	**218**	**311**	**505**	**764**	**1,127**	**1,942**	**3,163**	**6,480**	**16,139**
	%	**100**	**0.50**	**0.71**	**1.01**	**1.64**	**2.48**	**3.66**	**6.30**	**10.27**	**21.04**	**52.39**
AGRI	N	2,373	7	11	9	25	57	72	147	273	510	1,262
	%	100	0.29	0.46	0.38	1.05	2.40	3.03	6.19	11.50	21.49	53.18
BIO	N	4,582	24	41	50	100	121	187	296	476	933	2,354
	%	100	0.52	0.89	1.09	2.18	2.64	4.08	6.46	10.39	20.36	51.37
BUS	N	326	2	4	2	12	7	16	15	39	66	163
	%	100	0.61	1.23	0.61	3.68	2.15	4.91	4.60	11.96	20.25	50.00
CHEM	N	1,490	7	11	14	21	34	50	85	149	306	813
	%	100	0.47	0.74	0.94	1.41	2.28	3.36	5.70	10.00	20.54	54.56
COMP	N	765	4	9	20	20	25	29	57	82	150	369
	%	100	0.52	1.18	2.61	2.61	3.27	3.79	7.45	10.72	19.61	48.24
EARTH	N	1,437	5	4	14	14	32	51	81	149	297	790
	%	100	0.35	0.28	0.97	0.97	2.23	3.55	5.64	10.37	20.67	54.98
ECON	N	385	1	5	5	15	11	13	23	44	64	204
	%	100	0.26	1.30	1.30	3.90	2.86	3.38	5.97	11.43	16.62	52.99
ENG	N	1,282	7	12	14	26	32	59	88	153	259	632
	%	100	0.55	0.94	1.09	2.03	2.50	4.60	6.86	11.93	20.20	49.30
ENVIR	N	652	4	2	4	7	17	25	36	64	137	356
	%	100	0.61	0.31	0.61	1.07	2.61	3.83	5.52	9.82	21.01	54.60
IMMU	N	315	1	3	5	7	11	10	26	28	69	155
	%	100	0.32	0.95	1.59	2.22	3.49	3.17	8.25	8.89	21.90	49.21
MATER	N	584	3	2	5	11	19	27	38	64	115	300
	%	100	0.51	0.34	0.86	1.88	3.25	4.62	6.51	10.96	19.69	51.37
MATH	N	701	3	3	9	11	9	22	37	60	141	406
	%	100	0.43	0.43	1.28	1.57	1.28	3.14	5.28	8.56	20.11	57.92
MED	N	13,108	46	63	113	205	330	491	869	1,360	2,902	6,729
	%	100	0.35	0.48	0.86	1.56	2.52	3.75	6.63	10.38	22.14	51.34
NEURO	N	587	5	6	3	12	16	18	34	51	137	305
	%	100	0.85	1.02	0.51	2.04	2.73	3.07	5.79	8.69	23.34	51.96
PHYS	N	2,928	39	51	51	46	61	86	148	254	524	1,668
	%	100	1.33	1.74	1.74	1.57	2.08	2.94	5.05	8.67	17.90	56.97
PSYCH	N	548	4	5	5	8	17	19	27	66	104	293
	%	100	0.73	0.91	0.91	1.46	3.10	3.47	4.93	12.04	18.98	53.47

Similarly, when examining the mobility from mid-career to late-career for top-performing scientists (Table 5), the persistence in top productivity is even more pronounced, with 53.83% of scientists who started in decile 10 remaining in decile 10. Only one in six (16.61%) of scientists in decile 10 did not originate from deciles 8–10, and about 3% came from deciles 1–3 (3.39%). Among all current top-performing late-career scientists, there are only four Jumpers-Up: two economists (0.52%) and two psychologists (0.36%) who experienced extreme mobility from decile 1 to decile 10 (out of 385 and 548, respectively). For these individuals, we have comprehensive lifetime data on their demographics, publishing and collaboration patterns, and scholarly impact.

Table 5. Mobility of top performers between two career stages: mid-career (initial stage) and late career (target stage): From which initial productivity deciles (at a mid-career stage) do top-performing scientists at a late-career stage come from? Late-career scientists who were top performers at a late-career stage (N=32,063) by academic discipline and initial productivity decile (frequencies and percentages)

		Total	Bottom 10%	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Top 10%
Late-career scientists who are top performers at a late-career stage												
TOTAL	**N**	**32,063**	**436**	**305**	**345**	**541**	**756**	**1,125**	**1,816**	**3,139**	**6,339**	**17,261**
	%	**100**	**1.36**	**0.95**	**1.08**	**1.69**	**2.36**	**3.51**	**5.66**	**9.79**	**19.77**	**53.83**
SOCIAL	**N**	**1,259**	**10**	**14**	**21**	**32**	**44**	**55**	**58**	**131**	**249**	**645**
	%	**100**	**0.79**	**1.11**	**1.67**	**2.54**	**3.49**	**4.37**	**4.61**	**10.41**	**19.78**	**51.23**
STEMM	**N**	**30,804**	**426**	**291**	**324**	**509**	**712**	**1,070**	**1,758**	**3,008**	**6,090**	**16,616**
	%	**100**	**1.38**	**0.94**	**1.05**	**1.65**	**2.31**	**3.47**	**5.71**	**9.76**	**19.77**	**53.94**
AGRI	N	2,373	23	16	17	34	49	81	131	245	474	1,303
	%	100	0.97	0.67	0.72	1.43	2.06	3.41	5.52	10.32	19.97	54.91
BIO	N	4,582	65	38	45	85	104	170	266	467	952	2,390
	%	100	1.42	0.83	0.98	1.86	2.27	3.71	5.81	10.19	20.78	52.16
BUS	N	326	6	3	4	11	14	16	9	38	70	155
	%	100	1.84	0.92	1.23	3.37	4.29	4.91	2.76	11.66	21.47	47.55
CHEM	N	1,490	14	7	17	22	32	42	60	133	291	872
	%	100	0.94	0.47	1.14	1.48	2.15	2.82	4.03	8.93	19.53	58.52
COMP	N	765	11	8	9	19	32	25	55	59	145	402
	%	100	1.44	1.05	1.18	2.48	4.18	3.27	7.19	7.71	18.95	52.55
EARTH	N	1,437	12	21	12	23	42	61	83	155	308	720
	%	100	0.84	1.46	0.84	1.60	2.92	4.24	5.78	10.79	21.43	50.10
ECON	N	385	2	8	8	13	14	21	21	31	69	198
	%	100	0.52	2.08	2.08	3.38	3.64	5.45	5.45	8.05	17.92	51.43
ENG	N	1,282	13	12	19	26	33	43	81	118	240	697
	%	100	1.01	0.94	1.48	2.03	2.57	3.35	6.32	9.20	18.72	54.37
ENVIR	N	652	8	9	3	11	12	20	47	58	128	356
	%	100	1.23	1.38	0.46	1.69	1.84	3.07	7.21	8.90	19.63	54.60
IMMU	N	315	3	4	5	6	3	16	22	30	61	165
	%	100	0.95	1.27	1.59	1.90	0.95	5.08	6.98	9.52	19.37	52.38
MATER	N	584	8	1	6	10	5	15	30	63	117	329
	%	100	1.37	0.17	1.03	1.71	0.86	2.57	5.14	10.79	20.03	56.34
MATH	N	701	8	3	12	11	14	14	44	71	123	401
	%	100	1.14	0.43	1.71	1.57	2.00	2.00	6.28	10.13	17.55	57.20
MED	N	13,108	187	131	125	200	312	480	771	1,344	2,671	6,887
	%	100	1.43	1.00	0.95	1.53	2.38	3.66	5.88	10.25	20.38	52.54
NEURO	N	587	9	1	7	6	14	10	33	62	119	326
	%	100	1.53	0.17	1.19	1.02	2.39	1.70	5.62	10.56	20.27	55.54
PHYS	N	2,928	65	40	47	56	60	93	135	203	461	1,768
	%	100	2.22	1.37	1.61	1.91	2.05	3.18	4.61	6.93	15.74	60.38
PSYCH	N	548	2	3	9	8	16	18	28	62	110	292
	%	100	0.36	0.55	1.64	1.46	2.92	3.28	5.11	11.31	20.07	53.28

Analogous analyses were performed for the lowest productivity deciles. Supplementary Table 5 shows the decile origins of currently bottom-performing scientists (productivity decile 1), tracing their initial productivity deciles during their early-career stages and across various academic disciplines. The observed patterns are similar to, but less pronounced than, those seen for the top performers. Approximately three-fourths (75.31%) of bottom performers originated from the three lowest productivity deciles (deciles 1, 2, and 3), with over one-third coming from the lowest decile (37.41%). Conversely, only 2.28% (728 scientists) came from the top three deciles, including just 0.26% (82 scientists) from the highest decile 10 (referred to as Droppers-Down). Also for these scientists, we have full data (demographics, publishing and collaboration patterns, scholarly impact).

Supplementary Table 6 shows the decile origins of current bottom-performing late-career scientists during their mid-career stage, revealing similar mobility patterns. The majority of current decile 1 scientists came from the bottom three deciles (68.40%), while only 4.34% originated from the top three deciles. A mere 222 scientists (0.69%) experienced a drop from decile 10 to decile 1 (Droppers-Down). Upward mobility is of particular interest to science policy; downward mobility, in contrast, may often be attributed to personal life circumstances, such as health issues or family problems, which cannot be fully analyzed through bibliometric datasets.

Changing Productivity Classes: All Academic Disciplines Combined

The Sankey diagram (Figure 5) serves as a visual guide to better help understand the concept of scientists' mobility across productivity classes throughout their careers. This diagram illustrates the movement of scientists between productivity deciles at different career stages: early career (left: top and bottom), mid-career (middle: top and bottom), and late career (right: top and bottom). Our focus is on horizontal top-to-top and bottom-to-bottom mobility as well as the transitions involving extreme downward mobility from the top-to-bottom and extreme upward mobility from bottom-to-top.

Figure 5 displays the mobility of scientists across all academic disciplines combined (N=320,564), Figure 6 shows the mobility for all social science disciplines combined (N=12,585), and Figure 7 shows all STEMM disciplines combined (N=307,979). The left columns of the diagrams represent the distribution of early-career scientists within the top- and bottom productivity deciles (each decile totaling 100%), the middle columns represent mid-career scientists, and the right columns represent late-career scientists within the same two productivity classes. To enhance clarity, deciles 2 through 9 are excluded from the diagram.

The horizontal top-to-top and bottom-to-bottom mobility between the early- and mid-career stages is represented by thick flows: More than a half of global top performers continue as global top performers (52.39%), and more than one-third of global bottom performers continue as global bottom performers (37.41%). Extreme vertical top-to-bottom and bottom-to-top mobilities are rare and represented as thin downward and upward flows: Only 0.26% of top productive early-career scientists (82 scientists) land in the bottom productivity mid-career scientists—and only 0.51% bottom productive early-career scientists (162 scientists) land in the class of top productive mid-career scientists.

The mobility patterns between mid-career and late-career stage are very similar. The mobility patterns do not differ between social science disciplines combined (Figure 6) and STEMM disciplines combined (Figure 7): Surprisingly, despite different publishing and collaboration patterns, the top-to-top mobility for the early- to mid-career transition is almost exactly the same; and for the mid- to late-career transition, it is slightly higher for STEMM disciplines. It is as rare in the SOCIAL disciplines to experience extreme upward mobility—from decile 1 to decile 10—as it is in the STEMM disciplines. In our sample of SOCIAL scientists (N=12,585), there are only seven scientists involved in the first transition and only 10 scientists involved in the second transitions out of 1,259 scientists (Table 6). Here and elsewhere, statistical significance of differences are not shown because we work with a population of scientists (all scientists meeting the inclusion criteria) rather than with their sample (a selection of all scientists), which is an essential difference between small-scale and large-scale studies.

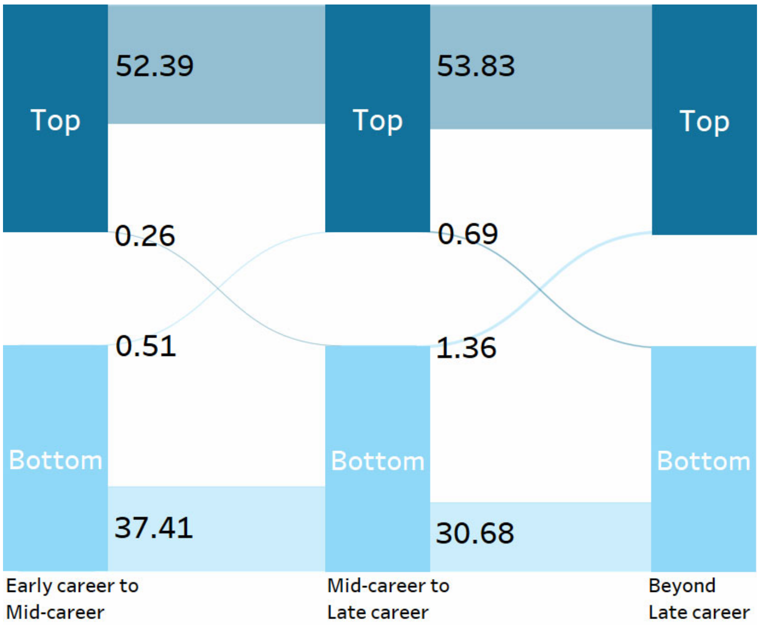

Figure 5. Scientists' mobility between productivity classes in the three stages of a scientific career. All academic disciplines combined (TOTAL), current late-career scientists. All observations ranked and clustered into productivity deciles, top- (upper 10%, productivity decile 10) and bottom- (bottom 10%, productivity decile 1) productivity classes only (N = 320,564) (percentages, top class, and bottom class 100% each)

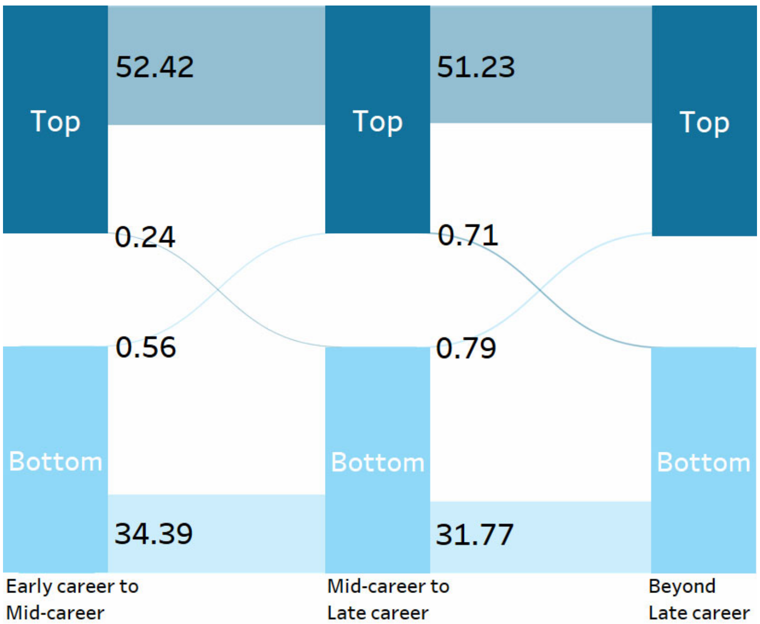

Figure 6. Scientists' mobility between productivity classes in the three stages of a scientific career. All social science disciplines combined (SOCIAL), current late-career scientists. All observations ranked and clustered into productivity deciles, top- (upper 10%, productivity decile 10) and bottom- (bottom

10%, productivity decile 1) productivity classes only (N = 12,585) (percentages, top class and bottom class, 100% each)

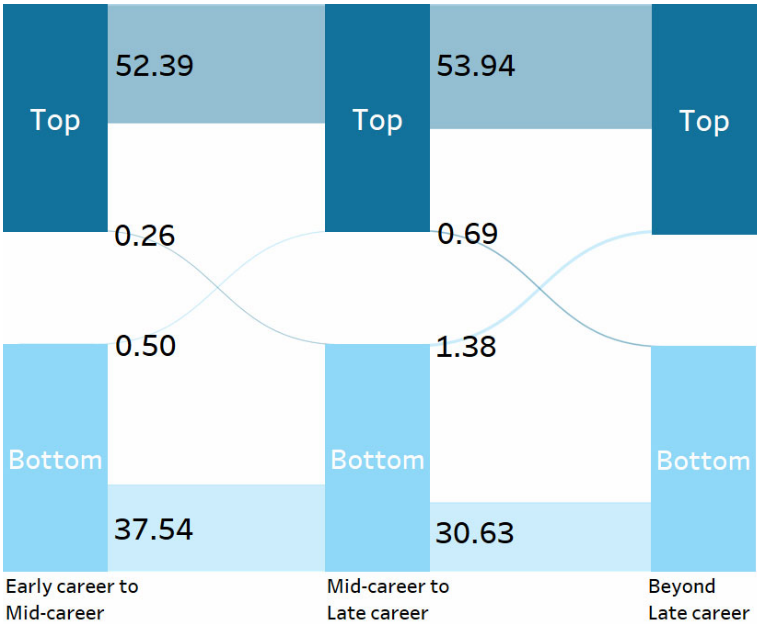

Figure 7. Scientists' mobility between productivity classes in the three stages of a scientific career. All STEMM academic disciplines combined, current late-career scientists. All observations ranked and clustered into productivity deciles, top- (upper 10%, productivity decile 10) and bottom- (bottom 10%, productivity decile 1) productivity classes only (N = 307,979) (percentages, top class and bottom class, 100% each)

From an aggregated perspective of all academic disciplines combined (Table 6), the mobility patterns are clear: Over half (52.39% and 53.83%) of the scientists who achieve top productivity (decile 10) remain in this top category. Similarly, about one-third (37.41% and 30.68%) of those in the bottom productivity category (decile 1) continue to stay in the same class. This indicates an intriguing "locking-in" mechanism within academic careers that warrants further scholarly investigation.

Importantly, our approach does not rely on publication numbers because productivity across OECD countries has generally increased over the past decades, especially when computed on a full counting rather than fractional counting basis. Instead, we rank all current late-career scientists by productivity, assigning them to specific productivity classes within their respective academic disciplines. We then retrospectively rank these late-career scientists based on their productivity during their early- and mid-career stages by using four-year periods to measure their productivity at these times ("initial academic decile" in productivity mobility in Table 6).

The likelihood of experiencing extreme upward mobility (moving from decile 1 to decile 10) or extreme downward mobility (moving from decile 10 to decile 1) between productivity classes is very low. According to our data (Table 6), in the context of our prestige-normalized counting approach, the chances of radical change in publishing behavior compared with peers within an academic discipline are minimal.

Specifically, only 162 scientists (0.51%) who became top performers in their mid career stage were initially in the bottom productivity class during their early career stage. Similarly, only 436 scientists (1.36%) who were top performers in their late career stage started as bottom performers in their mid career stage. The chances for extreme downward mobility are also rare, with less than 1% of scientists moving from the top- to the bottom productivity class in successive career stages (0.26% and 0.69%, respectively). These data indicate that radical changes in publishing behavior that lead to such significant shifts in productivity classes are quite uncommon.

Table 6. Mobility between top- (decile 10) and bottom- (decile 1) productivity classes while moving up from the early-career stage to mid-career stage and from the mid-career to late-career stage by all academic disciplines combined (N = 320,564, top panel), social science disciplines combined (N = 12,585, middle panel), and STEMM disciplines combined (N = 307,979, bottom panel) (frequencies and percentages)

Career stage (transition from)	Initial productivity decile	Career stage (transition to)	Target productivity decile	Number of scientists in transition	Number of scientists in productivity class	%
TOTAL (ALL DISCIPLINES COMBINED)						
Early career	Bottom	Mid-career	Bottom	11,996	32,063	37.41
Early career	Bottom	Mid-career	Top	162	32,063	0.51
Early career	Top	Mid-career	Bottom	82	32,063	0.26
Early career	Top	Mid-career	Top	16,799	32,063	52.39
Mid-career	Bottom	Late career	Bottom	9,836	32,063	30.68
Mid-career	Bottom	Late career	Top	436	32,063	1.36
Mid-career	Top	Late career	Bottom	222	32,063	0.69
Mid-career	Top	Late career	Top	17,261	32,063	53.83
SOCIAL SCIENCE DISCIPLINES COMBINED						
Early career	Bottom	Mid-career	Bottom	433	1,259	34.39
Early career	Bottom	Mid-career	Top	7	1,259	0.56
Early career	Top	Mid-career	Bottom	3	1,259	0.24
Early career	Top	Mid-career	Top	660	1,259	52.42
Mid-career	Bottom	Late career	Bottom	400	1,259	31.77
Mid-career	Bottom	Late career	Top	10	1,259	0.79
Mid-career	Top	Late career	Bottom	9	1,259	0.71
Mid-career	Top	Late career	Top	645	1,259	51.23
STEMM DISCIPLINES COMBINED						
Early career	Bottom	Mid-career	Bottom	11,563	30,804	37.54
Early career	Bottom	Mid-career	Top	155	30,804	0.50
Early career	Top	Mid-career	Bottom	79	30,804	0.26
Early career	Top	Mid-career	Top	16,139	30,804	52.39
Mid-career	Bottom	Late career	Bottom	9,436	30,804	30.63
Mid-career	Bottom	Late career	Top	426	30,804	1.38
Mid-career	Top	Late career	Bottom	213	30,804	0.69
Mid-career	Top	Late career	Top	16,616	30,804	53.94

Changing Productivity Classes: Cross-Disciplinary Differentiation

The aggregated pictures of all academic disciplines combined, all STEMM disciplines combined, and all SOCIAL disciplines combined hide a much more nuanced picture of individual academic disciplines with their distinct mobility patterns between productivity classes.

Focusing on the horizontal top-to-top (productivity decile 10 to decile 10) mobility first, for almost all academic disciplines, more than 50% of the top productivity scientists continue as top productivity scientists (Tables 7 and 8). The highest share is observed for MATH and PHYS in both transitions (as well as CHEM in the second transition), reaching as much as 60.38% for PHYS in the second transition. Scientists representing bottom-to-top mobility (Jumpers-Up), which is of great interest in productivity studies, are extremely rare across all academic disciplines: Their share ranges from 0.29%

for AGRI to 1.33% for PHYS in the first career transition and from 0.36% for PSYCH and 1.84% for BUS in the second career transition.

Table 7. Four mobility types by academic discipline between the early-career and mid-career stage, by discipline, percentages (N = 320,564)

Academic discipline	Top-to-top mobility		Bottom-to-bottom mobility		Top-to-bottom mobility		Bottom-to-top mobility	
	Scientists at early-career stage: top to top (%)	As % of top scientists at mid-career stage	Scientists at early-career stage: bottom to bottom (%)	As % of bottom scientists at mid-career stage	Scientists at early-career stage: top to bottom (%)	As % of bottom scientists at mid-career stage	Scientists at early-career stage: bottom to top (%)	As % of top scientists at mid-career stage
AGRI	53.18	53.18	38.60	38.60	0.17	0.17	0.29	0.29
BIO	51.37	51.37	36.56	36.56	0.33	0.33	0.52	0.52
BUS	50.00	50.00	30.37	30.37	0.31	0.31	0.61	0.61
CHEM	54.56	54.56	41.01	41.01	0.07	0.07	0.47	0.47
COMP	48.24	48.24	32.42	32.42	0.13	0.13	0.52	0.52
EARTH	54.98	54.98	40.36	40.36	0.21	0.21	0.35	0.35
ECON	52.99	52.99	29.87	29.87	0.26	0.26	0.26	0.26
ENG	49.30	49.30	36.97	36.97	0.23	0.23	0.55	0.55
ENVIR	54.60	54.60	35.58	35.58	0.46	0.46	0.61	0.61
IMMU	49.21	49.21	35.87	35.87	0.63	0.63	0.32	0.32
MATER	51.37	51.37	38.70	38.70	-	-	0.51	0.51
MATH	57.92	57.92	36.23	36.23	0.29	0.29	0.43	0.43
MED	51.34	51.34	36.37	36.37	0.16	0.16	0.35	0.35
NEURO	51.96	51.96	39.18	39.18	0.34	0.34	0.85	0.85
PHYS	56.97	56.97	42.25	42.25	0.75	0.75	1.33	1.33
PSYCH	53.47	53.47	39.96	39.96	0.18	0.18	0.73	0.73
SOCIAL	**52.42**	**52.42**	**34.39**	**34.39**	**0.24**	**0.24**	**0.56**	**0.56**
STEMM	**52.39**	**52.39**	**37.54**	**37.54**	**0.26**	**0.26**	**0.50**	**0.50**
TOTAL	**52.39**	**52.39**	**37.41**	**37.41**	**0.26**	**0.26**	**0.51**	**0.51**

Note: "-" = no scientists involved in this transition

Model Approach: Logistic Regression

In this subsection, we introduce a multidimensional approach and analyze the odds ratio estimates of membership in the classes of global top- and bottom productive scientists for current late-career scientists and, retrospectively, for current late-career scientists when they were mid-career scientists (the upper 10% and the bottom 10%, or decile 10 and decile 1, with separate models for each discipline, N=320,564).

We use a single demographic variable (gender, binary: male or female) and four variables we have computed using micro-level data about individual scientists. One variable is related to the article-level metrics of individual scholarly impact (field-normalized citations received within the first four years after each article has been published); two other variables are related to individual publishing and collaboration patterns (lifetime median team size and lifetime international collaboration rate); and one variable is related to publishing productivity in earlier career stages (prior membership in the global top- and bottom productivity classes at the early-career stage and at the mid-career stage).

Table 8. Four mobility types by academic discipline between the mid-career and late-career stage, percentages (N = 320,564)

Academic discipline	**Top-to-top mobility**		**Bottom-to-bottom mobility**		**Top-to-bottom mobility**		**Bottom-to-top mobility**	
	Scientists at mid-career stage: top to top (%)	**As % of top scientists at late-career stage**	**Scientists at mid-career stage: bottom to bottom (%)**	**As % of bottom scientists at late-career stage**	**Scientists at mid-career stage: top to bottom (%)**	**As % of bottom scientists at late-career stage**	**Scientists at mid-career stage: bottom to top (%)**	**As % of top scientists at late-career stage**
AGRI	54.91	54.91	32.20	32.20	0.25	0.25	0.97	0.97
BIO	52.16	52.16	28.26	28.26	0.87	0.87	1.42	1.42
BUS	47.55	47.55	26.38	26.38	0.61	0.61	1.84	1.84
CHEM	58.52	58.52	33.42	33.42	0.40	0.40	0.94	0.94
COMP	52.55	52.55	28.37	28.29	0.26	0.26	1.44	1.44
EARTH	50.10	50.10	35.00	35.00	0.28	0.28	0.84	0.84
ECON	51.43	51.43	29.09	29.09	0.78	0.78	0.52	0.52
ENG	54.37	54.37	29.88	29.88	0.47	0.47	1.01	1.01
ENVIR	54.60	54.60	28.68	28.68	0.46	0.46	1.23	1.23
IMMU	52.38	52.38	35.87	35.42	-	-	0.95	0.95
MATER	56.34	56.34	32.53	32.20	0.34	0.34	1.37	1.37
MATH	57.20	57.20	32.38	32.38	0.29	0.29	1.14	1.14
MED	52.54	52.54	29.48	29.48	0.82	0.82	1.43	1.43
NEURO	55.54	55.54	31.35	31.35	0.34	0.34	1.53	1.53
PHYS	60.38	60.38	34.53	34.53	1.13	1.13	2.22	2.22
PSYCH	53.28	53.28	36.86	36.86	0.73	0.73	0.36	0.36
SOCIAL	**51.23**	**51.23**	**31.77**	**31.77**	**0.71**	**0.71**	**0.79**	**0.79**
STEMM	**53.94**	**53.94**	**30.63**	**30.62**	**0.69**	**0.69**	**1.38**	**1.38**
TOTAL	**53.83**	**53.83**	**30.68**	**30.67**	**0.69**	**0.69**	**1.36**	**1.38**

Note: “-“ = no scientists involved in this transition

All publications (lifetime) and cited references (lifetime) were used to compute a unique discipline to which every scientists was ascribed (see Dataflow in Figure 1). The field-weighted four-year citation impact is computed for each individual publication separately and subsequently averaged to all publications over one’s lifetime; publishing and collaboration pattern variables are also computed from a lifetime perspective of individual scientists: All journal articles and articles in conference proceedings published throughout one’s lifetime are examined. In contrast, membership in productivity classes has been computed for two specific periods of careers (early- and mid-career periods). In addition, we also use in our regression models an institutional variable (TOP 200 institution globally, for logistic regression analysis for top productive late-career scientists only). The variables and their short descriptions are presented in Table 9. The inverse correlation matrices and main diagonals are shown in Electronic Supplementary Material (Supplementary Tables 7 through 10).

Extensive previous research on individual productivity has suggested that the most important predictors of high research productivity at the individual level are international collaboration (Dusdal & Powell, 2021), working in teams (Wagner, 2018), gender (Fox & Nikivinicze, 2021; Larivière, 2013), career stage and academic experience (Jung, 2014; Shin & Cummings, 2010; Kwiek, 2016), and publication productivity earlier in academic careers (Horta & Santos, 2016; Kwiek & Roszka 2024a, 2024b), consistently with the Matthew effect in science (the rich get disproportionately richer while the poor get poorer; DiPrete & Eirich, 2006). As a result—and specifically in the context of our two-dimensional results visualized through Sankey diagrams that show powerful top-to-top and bottom-to-bottom mobility in terms of productivity—we have added the prior membership in global top- and bottom productivity classes as variables used in the regression analysis.

Table 9. Variables used in regression analysis

No.	Variable	Description
1.	Gender	Gender (binary: female/male) provided by Elsevier's ICSR Lab. Variable classified based on the first name, last name, and dominant country from the first year of publishing using the Namsor tool. Gender accepted with the probability score >= 0.85 only.
2.	Field-weighted four-year citation impact (FWCI 4y)	Average of the FWCI 4y metric values assigned to each publication in author's lifetime publication portfolio. The FWCI 4y metric value of a publication means the ratio of the number of citations of that publication (obtained in the publication year and three consecutive years) to the average number of citations for a similar publication (publication from the same discipline group in Scopus 4-digit ASJC discipline classification) in the same time frame.
3.	International collaboration rate (lifetime)	Share of author's international collaborative publications among all collaborative publications (solo publications excluded). For a publication to be considered collaborative, the number of all authors in the paper had to be greater than or equal to two. For a publication to be considered international, the number of affiliation countries in the paper had to be greater than or equal to two.
4.	Median team size (lifetime)	Median of the number of authors for each publication (author + number of collaborators) in author's lifetime publication portfolio. For publications with the number of authors greater than 10, the number of authors is 10.
5.	Discipline	Dominant discipline based on the modal value from all disciplines assigned to the journals of all cited references in all papers in scientists' lifetime publication portfolios.
6.	TOP200 institutional affiliation	Binary value indicating belonging (true/false) to one of the 200 top institutions. The list of top institutions was ranked based on the institutions' total scholarly output in the 10-year period of 2014–2023. Each author has been assigned to one institution as the dominant one based on the modal value from institutions indicated in author's lifetime publication portfolio. Used only for second transitions: mid-career to late career (affiliation in early career is too distant in time).
7. 8.	Early-career / mid-career top class	Membership in the global top 10% of scientists among early-career /mid-career scientists in terms of research productivity, separately within 16 disciplines.
9. 10.	Early-career / mid-career bottom class	Membership in the global bottom 10% of scientists among early-career / mid-career scientists in terms of research productivity within 16 disciplines.

First, we analyze the top productivity classes: Odds ratio estimates of membership in the class of global top productive mid-career scientists (Table 10) and global top productive late-career scientists (Table 11). In the vast majority of disciplines, high productivity in an earlier stage of career is the most powerful predictor of high productivity in a later stage, with Exp(B) in the range of 11.68 (IMMU) to 22.97 (MATH) for the first career stage and 13.89 (EARTH) to 21.82 (MATH) for the second career stage (in all cases: all other things being equal; the predictor is not statistically significant for two disciplines in the first stage and six disciplines in the second stage).

The direction of the impact of the gender variable (being a male scientist) is consistent across disciplines; however, the impact is highly differentiated. Being male increases the probability of success in 13 out of 16 disciplines for mid-career scientists by as little as 3–7% in COMP, MATH, and ECON and by more than 100% (in IMMU and NEURO). For late-career scientists, gender is statistically significant in 10 disciplines, and its impact is smaller; being male increases the odds of success by 3–14% (in AGRI, EARTH, MATER) to 46–70% (in NEURO, BIO, and IMMU), with the exception of MATH, where it decreases the odds by 2%, again, all other things being equal.

Interestingly, the probability of success for men differs between highly mathematized disciplines, which traditionally have the lowest percentage of women (MATH, COMP, ENG, and PHYS), and the disciplines with high percentages of women (see Supplementary Table 1 for male and female participation data in our sample). In the former cluster, the impact of gender is marginal (COMP and MATH in the first stage; MATH in the second stage) or not statistically significant (ENG and PHYS in the first stage, and COMP, ENG, and PHYS in the second stage). In contrast, in the disciplines with high participation of women, the probability of success for men is 86% higher than for women in BIO, 112% higher in IMMU, and 127% higher in NEURO (for the first stage), and it is 70% higher in IMMU, 46% higher in NEURO, and 47% higher in BIO, the largest exception to the rule being MED, which has much lower increases for men compared with women.

For all disciplines, at both career stages, article-level citation metrics is statistically significant (except for PSYCH and COMP in the second stage): an increase of a field-weighted citation impact for all publications from an early-career period by one unit increases the probability of success from a few percentage points (as in BUS and PSYCH, as well as COMP and PHYS) to as much as 53% in AGRI, all other things being equal. In the second stage, the impact in the mid-career stage is the most consequential in MATER, increasing the odds by 50%.

In addition, international collaboration rate at an early-career period (and at a mid-career period) is statistically significant for all disciplines, both from the STEMM and SOCIAL clusters of academic disciplines. Generally, the probability of success increases by about 1% for a one-unit increase so that a 30% higher rate increases the odds by 30%. Finally, team size is statistically significant: The higher the median team size at an early-career period, the higher chances of success. The team size is especially consequential for membership in top productive mid-career class in traditionally sole-authored or small-team disciplines, such as the three social science disciplines (BUS, ECON, and PSYCH) and MATH in the STEMM cluster. In these disciplines, a one-unit increase (i.e., one more coauthor in publications in an early-career period) increases the probability of success by 25–40%. In other words, publishing on average with three additional co-authors—or working in larger teams—increases the probability of success by as much as 75–120%.

Table 10. Logistic regression statistics: odds ratio estimates of membership in the class of global top productive mid-career scientists (the top 10%, separately for each academic discipline) (N = 320,564)

Model	**AGRI**	**BIO**	**BUS**	**CHEM**	**COMP**	**EARTH**	**ECON**	**ENG**	**ENVIR**	**IMMU**	**MATER**	**MATH**	**MED**	**NEURO**	**PHYS**	**PSYCH**
R^2	0.25	0.22	0.21	0.26	0.18	0.26	0.23	0.20	0.26	0.21	0.24	0.28	0.22	0.25	0.34	0.25
Male	1.32 (3)	1.86 (3)	1.61 (3)	1.30 (3)	1.07 (3)	1.29 (3)	1.03 (1)	0.78	1.21	2.12 (3)	1.41 (1)	1.05 (1)	1.30 (3)	2.27 (3)	0.95	1.36 (2)
Avg. FWCI 4y Early	1.53 (3)	1.14 (3)	1.04 (3)	1.37 (3)	1.05 (3)	1.24 (3)	1.12 (3)	1.16 (3)	1.27 (3)	1.22 (3)	1.41 (3)	1.23 (3)	1.02 (3)	1.37 (3)	1.05 (3)	1.08 (3)
Inter. Collab. Rate Early	1.01 (3)	1.01 (3)	1.01 (3)	1.01 (3)	1.00 (3)	1.01 (3)	1.01 (3)	1.01 (3)	1.01 (3)	1.01 (3)	1.01 (3)	1.00 (3)	1.01 (3)	1.00 (3)	1.02 (3)	1.01 (3)
Median Team Size Early	1.06 (3)	1.01 (3)	1.31 (3)	0.99 (3)	1.12 (3)	1.05 (3)	1.35 (3)	0.96 (2)	1.01 (3)	1.02 (3)	1.06 (3)	1.40 (3)	1.08 (3)	1.16 (3)	1.17 (3)	1.25 (3)
Top Early	15.39 (3)	15.87 3)	15.10 (3)	17.81 (3)	14.64 (3)	17.17 (3)	17.49 (3)	14.33	18.16	11.68 (3)	13.53 (2)	22.97 (3)	16.39 (3)	14.25 (3)	11.92 (2)	18.52 (2)
Constant	0.02 (3)	0.02 (3)	0.02 (1)	0.02 (1)	0.03 (1)	0.02 (2)	0.02 (1)	0.06	0.02	0.02 (2)	0.02	0.02	0.03 (3)	0.01 (2)	0.01	0.02

Note: (1) = p-value ≤ 0.05; (2) = p-value ≤ 0.01; (3) = p-value ≤ 0.001

Table 11. Logistic regression statistics: odds ratio estimates of membership in the class of global top productive late-career scientists (the top 10%, separately for each academic discipline) (N = 320,564)

Model	**AGRI**	**BIO**	**BUS**	**CHEM**	**COMP**	**EARTH**	**ECON**	**ENG**	**ENVIR**	**IMMU**	**MATER**	**MATH**	**MED**	**NEURO**	**PHYS**	**PSYCH**
R^2	0.25	0.22	0.19	0.29	0.22	0.21	0.21	0.26	0.25	0.24	0.28	0.28	0.23	0.26	0.35	0.24
Male	1.07 (3)	1.47 (3)	1.18	1.27 (3)	0.97	1.14 (2)	1.26 (3)	0.89	1.09	1.70 (2)	1.13 (1)	0.98 (1)	1.24 (3)	1.46 (3)	0.85	0.95
Avg. FWCI 4y Mid	1.31 (3)	1.03 (3)	1.12 (3)	1.28 (3)	1.05	1.10 (3)	1.03 (3)	1.09 (2)	1.18 (3)	1.14 (3)	1.50 (3)	1.28 (3)	1.02 (3)	1.20 (3)	1.07 (3)	1.33
Inter. Collab. Rate Mid	1.01 (3)	1.01 (3)	1.01 (3)	1.01 (3)	1.01 (3)	1.01 (3)	1.00 (3)	1.02 (3)	1.01 (3)	1.01 (3)	1.01 (3)	1.00 (3)	1.01 (3)	1.01 (3)	1.02 (3)	1.01 (3)
Median Team Size Mid	0.99 (3)	1.03 (3)	1.17	0.93 (3)	1.05	1.00 (3)	1.34 (3)	0.89 (1)	0.92 (3)	1.04 (3)	0.95 (3)	1.35 (3)	1.10 (3)	1.02 (3)	1.10 (3)	1.04
TOP200	1.46 (3)	1.30 (3)	1.26	1.42 (3)	1.09	1.47 (3)	1.01 (3)	1.81	1.51	1.34 (1)	1.44 (1)	1.05 (2)	1.50 (3)	1.41 (3)	1.23	1.43
Top Mid	18.06 (3)	16.76 (3)	12.47	22.98 (3)	17.92	13.89 (3)	16.16 (3)	18.01	18.83	14.73 (2)	18.26 (3)	21.82 (3)	16.47 (3)	19.27 (3)	13.94	16.74
Constant	0.03 (1)	0.03 (3)	0.02	0.03(1)	0.04	0.03(1)	0.02(1)	0.05	0.04	0.02	0.02	0.02	0.02 (3)	0.02 (2)	0.01	0.03

Note: (1) = p-value ≤ 0.05; (2) = p-value ≤ 0.01; (3) = p-value ≤ 0.001

However, at a late career, stage, the role of larger collaboration teams is much more diversified: In some academic disciplines, larger median team size at a mid-career period actually decrease the chances of success; in several disciplines, the team size is statistically insignificant (including BUS, COMP, and PSYCH); and in others, team size continues to be a highly influential predictor of success (ECON and MATH where a one-unit increase, i.e., one additional collaborator, increases the odds by as much as one-third).

In models for the late-career stage, we have introduced a TOP200 variable (see Table 9). The TOP200 predictor substantially (by 30–50%) increases the odds of success in eight disciplines, with the highest impact in MED, the largest discipline (by 50%—which may be linked to better access to better training and costly infrastructure such as, e.g., university hospitals). The predictor is statistically insignificant in such disciplines as BUS and PSYCH as well as COMP and PHYS.

We have also analyzed the global bottom productivity classes (Supplementary Tables 11–12). The results are mirror images of those for global top productivity classes but only to some extent. Specifically, although the role of membership in the bottom productivity classes in previous career stages follows the expected pattern—that is, prior membership increases the odds of future membership—the independent variables of prior bottom productivity classes are statistically significant only for seven (first career period) and two (second career period) disciplines. Membership in the bottom productivity classes earlier (publishing years 5–14) increases the probability of membership in this class later (publishing years 15–24) 4–7 times in the case of the first career period and three times (BIO and MED only, the two largest disciplines in our sample, comprising together 55.18% scientists) in the case of the second career period.

The role of gender is not unequivocal: Being a female scientist substantially increases the odds of success only in AGRI and ENG in the first period and slightly (by 17%) in MED in the second stage; interestingly, for all disciplines but two (MED and BIO), gender is statistically insignificant in the second stage. What is notable is the contrast between BIO (with one of the highest shares of late-career women scientists) and ENG (with the lowest share of late-career women scientists) in the case of mid-career scientists. Being male in BIO decreases the odds of success by one-third on average, but being male in ENG increases the odds of success by half on average (28.0% and 48.0%, respectively).

Discussion and Conclusions

Our global results fully support our previous small-scale single-nation research in which half of the Polish top productive assistant professors continued as top productive associate professors, and half of the top productive associate professors continued as top productive full professors (52.6% and 50.8%) (we used a much less granular approach: a tripartite division into top, middle, and bottom productivity classes based on the 20/60/20 formula, N=2326 full professors, Kwiek & Roszka, 2024a). Similarly, publishing productivity during assistant professorship heavily influences productivity during associate professorship (N=4165 associate professors, Kwiek & Roszka, 2024b). Individual-level microdata from a national registry of scientists suggested that Polish associate professors tend to be stuck in their productivity classes for years: High performers tend to remain high performers, and low performers tend to remain low performers for up to 40 years of their careers. Our OECD study includes 79.42% of all late-career scientists globally (and 83.03% of all research articles produced by late-career scientists globally).

Following the results of our longitudinal study based on micro-level data on hundreds of thousands of late-career scientists, we suggest that, relatively early on in scientific careers, the productivity distribution within the global science profession at its two extremes (top 10% and bottom 10%) is already largely settled and that the early global distribution persists over time, that is, for years and decades. Exceptions are very rare: global bottom performers almost never become global top performers (our Jumpers-Up), and global top performers almost never become global bottom performers (our Droppers-Down).

However, global data on individual publishing careers show that attrition in science—leaving science for good or ceasing to publish—powerfully affects the global science workforce. Some scientists stay on in academic science and keep publishing, while others stop publishing (Geuna & Shibayama, 2015; Preston, 2004). About one-third disappears from academic publishing within five years—which is mostly the time spent in doctoral schools—and about a half within a decade (Kwiek & Szymula, 2023). Within the first 15 years of academic publishing, those who stay on in science are already distributed among decile-based classes of individual productivity, from the top 10% and bottom 10% within their disciplines. Our focus is only on scientists who stay on in science (and keep publishing) for at least 25 years, hence following our definition of late-career scientists in the present research.

What is stunning is the persistence of membership in global top and bottom productivity classes from a life cycle perspective. Later on in their careers, the majority of global top performers (decile 10 in productivity distribution) keep being top performers, and about one-third of global bottom performers (decile 1 in productivity distribution) keep being bottom performers. For them, the probability of staying in global top and bottom productivity classes—horizontal mobility in productivity—over the decades of scientific careers is high; in contrast, the probability of radically vertically changing productivity classes (Jumpers-Up, Droppers-Down) is extremely limited.

The global science system is highly immobile in terms of membership in productivity classes: Jumpers-Up and Droppers-Down are extremely rare (e.g., our micro-level data show that only 0.51% scientists move from early-career bottom class to mid-career top class; and only 0.26% scientists move from early-career top class to mid-career bottom class; there are only 162 and 82 such outliers, respectively, out of 32,063 scientists in all fields combined; Figure 5 and Table 6).

For instance, among all current top-performing late-career economists and psychologists, there are only four Jumpers-Up: two economists (0.52%) and two psychologists (0.36%) who experienced extreme mobility from decile 1 to decile 10 (out of 385 and 548, respectively, Table 5). The strength of this micro-level data associated with individual Scopus IDs lies in the rich insights these IDs offer into the unique trajectories of outlier individuals, such as those who have made significant upward leaps in productivity (Jumpers-Up) and those who have made significant downward leaps in productivity (Droppers-Down).

Large-scale computations based on raw Scopus data (e.g., 1.8 billion cited references used to define unique academic discipline for each scientist in our dataset) allow us to explore not only these unique cases but also any other individuals or groups in our study. For every scientist in our sample, we have access to a comprehensive set of micro-level demographic, institutional, and publishing pattern data derived from a global bibliometric dataset.

By utilizing Scopus IDs, we can examine the publishing and collaboration histories of our four impressive Jumpers-Up in economics and psychology, including their evolving impact on academic science, their gender (determined using gender-detection software, as detailed in Karimi et al. 2016; Santamaria & Mihaljević 2018; Sebo 2021, 2023), their country affiliations at various stages of their careers, the research intensity of their institutions, the year they began (and concluded, if applicable) their scientific careers, their international collaboration rate (both lifetime and during specific periods), their overall collaboration rate, median team size, field-weighted citation impact of each publication within a selected time frame (e.g., a four-year window), their average publication journal percentile rank, and their lifetime scholarly output by publishing outlet type (e.g., top journals, open access journals, etc.).

In essence, we can gain a comprehensive understanding—within the constraints of the database and computing methods—of who these outlier scientists are, how they collaborate, publish, and work, and how their contributions are received by the academic community.

As many as 8 in 10 global top performers classified in decile 10 of productivity distribution come from deciles 8–10 (83.66% of mid-career scientists and 83.39% of late-career scientists, a stunningly similar percentage); analogously, global bottom performers classified in decile 1 of productivity distribution predominantly come from deciles 1–3 (from 75.31% in the first stage to 68.40% in the second stage), with some cross-disciplinary variation.

Individual research productivity emerges from our regression analyses as highly path dependent: For all the examined disciplines, STEMM and SOCIAL clusters alike, there is a single most important predictor of becoming a top productive late-career scientist (and a top productive mid-career scientist): being a top productive scientist at an earlier career stage.

The TOP200 predictor (working in the 200 most research-intensive institutions globally) substantially (by 30–50%) increases the odds of success in half of the disciplines studied, with the highest impact in MED, which is the largest discipline, hence pointing to the role of expensive infrastructure in high performance in medical research. Finally, the team size is especially important for membership in the top productive mid-career class in traditionally sole-authored or small-team disciplines, such as the three social science disciplines (BUS, ECON, and PSYCH) and MATH in the STEMM cluster. In these disciplines, a one-unit increase (i.e., one more coauthor in publications in an early-career period) increases the probability of success by 25–40%, testifying to the importance of collaboration in science (Kwiek, 2021; Wagner, 2018).

A general picture of mobility between productivity classes over the course of entire scientific careers based on discipline-aggregated data (Table 6) hides much more nuanced pictures for different disciplines (Tables 7–8). Some disciplines are much more competitive at the very beginning of their careers, with radical upward mobility being extremely difficult (as in CHEM and COMP, 0.07% and 0.13%, respectively). There are also other disciplines that are much less competitive from the very beginning, in which the presence of Jumpers-Up is much higher (e.g., PHYS, 0.75%, or 10 times more than in CHEM). Our focus is not on globally evolving productivity over time (Rørstad & Aknes, 2015) or evolving productivity from a generational perspective (e.g., the old in science being more productive than the young or the other way round Savage & Olejniczak, 2021; Way et al., 2017) but on interclass mobility of individuals over their entire scientific careers.

Why does prior class memberships (top, bottom), to a large extent, determine later class memberships (top, bottom)? There are two explanations, we can speculate. First, previous research has shown that the distribution of productivity among scientists has always been highly skewed (Abramo et al., 2017; Albarrán et al., 2011; David, 1994) and that a minority of scientists have always been responsible for the vast majority of publications (Allison, 1980; Ruiz-Castillo & Costas, 2014; Xie, 2014). The old research and science policy theme, which can be summarized as "the majority of scientific work is performed by a relatively small number of scientists" (Crane, 1965: 714), has been at the core of these theories of individual research productivity. Esteem comes from peers in science, and the reward system in science is based on publications. In addition, academic promotions and tenure prospects, salary levels, free time for research, and access to research grants are more or less directly all related to publication productivity (Allison et al., 1982; Stephan, 2012). Previous success breeds current and future success, here success being membership in the tiny class of global top research performers.

Second, higher publishing productivity generally leads to new research funding, as the credibility cycle in academic careers shows (Latour & Woolgar, 1986). In this cycle, research published in prestigious journals (quantity, quality) is converted into recognition; successful grant applications are converted into new equipment, arguments, and articles. The credibility cycle may be more consequential, determining career opportunities in the early-career stages: Once funded based on prestigious articles, scientists' probability to be funded again are higher than that of their less-productive colleagues, at least in more meritocratic national research funding systems with substantial individual-level grant funding. Once funded and once they have excellent publications, scientists have better odds to be funded again and to be promoted sooner to higher ranks, reflecting the idea that each element of the credibility cycle in academic careers "is but one part of an endless cycle of investment and conversion" (Latour & Woolgar, 1986: 200).

In terms of shifting productivity classes from a life cycle perspective, scientists who are less successful early on in their careers (and success here requires a combination of productivity, motivation, determination, aspirations, mentorship, resources, quality of training, innate abilities, and luck) will find it difficult, if not impossible, to prove that they are as good as their more successful, more productive, more determined, more able, better trained, luckier, and possibly better funded colleagues. Initial publication success is highly correlated with later publication success, which may be explained by two factors: Scientists from the very beginning are different, with some being much more productive, and scientists happen to experience initial publishing success for reasons unrelated to their exceptionality. In both cases, others may view them more positively, leading to more successes in grant acquisition, publishing acceptance, and so forth. Any of the two major explanations separately or together can work for individuals, strengthening the credibility cycle in academic careers.

Our research has reconfirmed the power of very strong track record as opposed to very weak track record in science (whenever individual scientists are assessed by research funding panels and promotion committees): For a variety of reasons—which we are not able to fully examine using our dataset—the probability of past global top performers becoming global top performers in the future is very high, and their probability of becoming global bottom performers is marginal. At the same time, the chances of global bottom performers to reach the productivity levels achieved by their top-performing colleagues in the very same career stages and within their disciplines (Jumpers-Up) are marginal. Catching up with the top performers just does not happen, except for a few outliers (and in some system, as in Poland, it does not happen at all: The chances we have computed for Poland are 0%; Kwiek & Roszka, 2024b).

Persistent productivity stratification emerges from our individual micro-level analyses as a powerful feature of global science. Using large numbers of observations, our analyses confirm what traditional productivity theories have been claiming for decades, albeit by using small-scale interviews and surveys (Cole & Cole, 1973; Hermanowicz, 2012; Leišytė & Dee, 2012; Merton, 1973): Success breeds success (as in cumulative advantage theory of productivity), and some scientists will always be globally highly performing while others will always be globally low performing (as in the sacred spark theory of productivity).

More generally, the present research represents trade-offs between what is theoretically desirable and what is practically possible in studying the global patterns of individual research productivity, here based on currently available data from global bibliometric datasets. However, there are trade-offs, biases, and limitations related to our data and methodology.

First, there seem to be no other longitudinal datasets globally available (for 38 OECD countries) than Scopus (or Web of Science – but not OpenAlex; see Priem et al., 2022) that can be meaningfully used to examine changing field-normalized productivity over scientists' lifetimes. In terms of author disambiguation, Scopus is more accurate than Web of Science (Sugimoto & Larivière 2018: 36), but it is certainly not perfect. National datasets are available for selected countries only and with selected parameters only (e.g., CRISTIN dataset for Norway, Academic Analytics dataset for the USA, and RADON dataset for Poland; see Albarrán et al., 2011; Kwiek & Roszka, 2023b; Nygaard et al., 2022; Savage & Olejniczak, 2021). As a result, no longitudinal and discipline-based global (as opposed to selected country-based) approaches to publishing productivity are currently possible without access to the global bibliometric datasets that provide metadata on all indexed publications over time. However, global bibliometric datasets come with their own limitations and biases, as has been widely discussed for at least two decades (see, e.g., Sugimoto & Larivière, 2018: 38–44 on the cultural biases of bibliometric data sources; Sugimoto & Larivière, 2023: 11–12 on published scientific document standing as a proxy for complex interpersonal and cognition processes of knowledge production that precede it).

Second, the character of our dataset determines a reductive understanding of individual productivity in which only Scopus-indexed publications are counted, leaving aside nonindexed publications in English and most publications in local languages. However, our focus on STEMM and the three selected social science disciplines, here generally using English for global scholarly communication, makes the present research less biased (STEMM disciplines are covered in global datasets in much larger percentages than the humanities). Scopus is reported as being the largest quality-controlled citation index, covering substantially more years than Dimensions or the Web of Science Core Collection (Thelwall & Sud, 2022). Additionally, all nonpublishing academic activities do not count toward productivity, and the nonpublishers in science have not been analyzed.

Third, the longitudinal nature of our study makes only the survivors our focus: We leave aside all scientists who are not research active for at least 25 years, which follows our definition of late-career scientists. As a result, being aware of high attrition rates in the STEMM disciplines (in OECD countries as recently analyzed in Kwiek & Szymula, 2024; and in the USA as recently analyzed in Spoon et al., 2023), we acknowledge a "success bias" in our research: The various mobility types between the analyzed productivity classes do not actually refer to all beginning, early-career, and mid-career scientists (active in publishing for less than 5, less than 15, and less than 25 years, respectively) but to survivors in science only (see Wang & Barábasi, 2021: 241–245 on the underexploration of "failures" in science). Our study takes a long-term view in which, necessarily, because of the high attrition rate in science, the majority of currently active scientists are actually not represented.

Fourth, our methodology has clear limitations that are especially evident if we compare the present study to single-nation longitudinal studies of publishing productivity. In single-nation studies, a wealth of national data are used (e.g., individual career histories with academic promotion dates, doctoral and postdoctoral dissertation details, research funding details, national classifications of disciplines, national rankings of institutions, etc.) that are currently not available for the 38 OECD countries. Additionally, our global study examines scientists from systems with different research funding levels and average individual publishing productivity.

Therefore, out of necessity, our analyses have had to rely on several proxies: (1) a commercial journal classification (Scopus All Science Journal Classification system, ASJC) and journal disciplinary classifications rather than a wealth of national disciplinary classifications (Baas et al., 2020); (2) data on individual Scopus IDs rather than data on "real scientists" with their national registry-based IDs (creating a fundamental ontological difference between traditional academic career research and bibliometric and bibliometric data-driven studies); (3) inferred rather than self-declared (or administratively-provided) gender, here based on gender-determining algorithms (probability threshold used: 0.85); (4) a single country affiliation and a single institutional affiliation rather than a plethora of changing country and institutional affiliations over a course of academic careers, at least for some percentage of scientists; (5) academic careers beginning with the first indexed publication and 25 years of research experience counted based on this publication date (i.e., publishing career)—rather than first employment in academia or beyond (i.e., academic or scientific career per se). In our research, the breadth of scientists' activities in academia (e.g., mentoring students, refereeing papers, reviewing grant proposals, and editing journals) is ignored (Liu et al., 2023).

Fifth, our approach to publishing productivity uses journal-level metrics of impact rather than more granular article-level metrics of impact, with all limitations (as discussed, e.g., in Weingart, 2004; an improvement of the present research would be to implement productivity measures based on the citation impact directly achieved by each individual publication, as suggested but not implemented in Carrasco & Ruiz-Castillo, 2012; see Electronic Supplementary Material); however, in logistic regression analyses, we use the field-weighted four-year citation impact assigned to each scientist and based on averaged citations to every article, here using Scopus 4-digit ASJC discipline classification).

Sixth, our dataset does not include environmental variables that bear heavily on individual publishing productivity, especially on the productivity of women scientists: We were not able to examine "work climates" characterizing the basic units in which scientists work, which are reported to be especially important in STEM disciplines for both attrition (Spoon et al., 2023) and productivity (Branch, 2016; Fox & Mohapatra, 2007). We had no data about academic attitudes and behaviors, working time distribution, teaching/research preference, work-life balance, and household and parenting obligations, which have been routinely reported in productivity studies based on survey data (e.g., Kwiek, 2016 on 11 European countries and Kwiek, 2018 on Poland), which leads us to the final limitation of the current study: no gender differentiation in examining publishing careers.

As a result, finally, our research does not tackle an extremely relevant topic of gender differences in mobility between productivity classes over time, except for the regression models. Adding an additional dimension of gender to an examination of late-career scientists—technically perfectly feasible, with full data available—would complicate our global mobility picture substantially; as a result, we have decided to leave a study of gender disparities by discipline for another occasion to make sure this topic can be reflected on with sufficient depth.

Our global and longitudinal approach to mobility between research productivity classes at the micro-level of individual scientists uses various proxies and relies on different trade-offs but hopefully shows new patterns that are, so far, largely underexplored in academic career studies. Our future research involves a more in-depth study of gender-based mobility between productivity classes in the selected disciplines, especially in traditionally male-dominated mathematics and computing.

This paper is accompanied by Supplementary Material *that is available online.*

Acknowledgments

We are grateful to the hosts and audiences of the invited seminars at the University of Oxford (Simon Marginson, CGHE, Center for Global Higher Education, June 2022), Stanford University (John Ioannidis, METRICS, Meta-Research Innovation Center, June 2022), DZHW Berlin (Torger Möller, German Center for Higher Education Research and Science Studies, June 2023), Leiden University (Ludo Waltman, CWTS, Centre for Science and Technology Studies, June 2023), and Sciences Po (Christine Musselin, Sciences Po Paris, October 2023), where Marek Kwiek discussed the strengths and limitations of longitudinal academic career studies using structured Big Data. We gratefully acknowledge the assistance of the International Center for the Studies of Research (ICSR) Lab, with particular gratitude to Kristy James and Alick Bird from the Lab. We also want to thank Dr. Wojciech Roszka from the CPPS Poznan Team for the numerous fruitful discussions and our joint publications about Polish top performers and how to measure their mobility using national datasets.

Author contributions

Marek Kwiek: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Writing—original draft, Writing—review & editing. Lukasz Szymula: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing—original draft, Writing—review & editing.

Competing interests

The authors have no competing interests.

Funding information

We gratefully acknowledge the support provided by the Ministry of Education and Science through its NDS grant no. NdS/529032/2021/2021.

Data availability

We used data from Scopus, a proprietary scientometric database. For legal reasons, data from Scopus received through collaboration with the ICSR Lab cannot be made openly available.

References

Abramo, G., D'Angelo, C. A., & Caprasecca, A. (2009). The contribution of star scientists to overall sex differences in research productivity. *Scientometrics*, *81*(1), 137–156.

Abramo, G., D'Angelo, C. A., & Soldatenkova, A. (2017). An investigation on the skewness patterns and fractal nature of research productivity distributions at field and discipline level. *Journal of Informetrics*, *11*(1), 324–335.

Aguinis, H., & O'Boyle, E. (2014). Star performers in twenty-first century organizations. *Personnel Psychology*, *67*(2), 313–350.

Albarrán, P., Crespo, J. A., Ortuño, I., & Ruiz-Castillo, J. (2011). The skewness of science in 219 sub-fields and a number of aggregates. *Scientometrics*, *88*(2), 385–397.

Allison, P. D. (1980). Inequality and scientific productivity. *Social Studies of Science*, *10*, 163–179.

Allison, P. D., & Stewart, J. A. (1974). Productivity differences among scientists: Evidence for 1`accumulative advantage. *American Sociological Review*, *39*(4), 596–606.
Allison, P. D., Long, J. S., Krauze, T. K. (1982). Cumulative advantage and inequality in science. *American Sociological Review*, *47*(5), 615–625.
Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*, *1*(1), 377–386. 10.1162/qss_a_00019
Branch, E.H., ed. (2016). *Pathways, potholes, and the persistence of women in science: Reconsidering the pipeline*. Lexington Books.
Carrasco, R., & Ruiz-Castillo, J. (2014). The evolution of the scientific productivity of highly productive economists. *Economic Inquiry, 52*(1), 1–16.
Clauset, A., Larremore, D. B., & Sinatra, R. (2017). Data-driven predictions in the science of science. *Science*, *355*, 477–480.
Cole, J. R., & Cole, S. (1973). *Social stratification in science*. The University of Chicago Press.
Costas, R, Bordons, M. (2007). A classificatory scheme for the analysis of bibliometric profiles at the micro level. In *Proceedings of ISSI 2007: 11th international conference of the ISSI* (Vols I and II, pp. 226–230). CSIC, Madrid.
Costas, R., van Leeuwen, T. N., & Bordons, M. (2010). Self-citations at the meso and individual levels: Effects of different calculation methods. *Scientometrics*, *82*, 517–537.
Crane, D. (1965). Scientists at major and minor universities: A study of productivity and recognition. *American Sociological Review*, *30*(5), 699–714.
David, P. A. (1994). Positive feedbacks and research productivity in science: Reopening another black box. In O. Granstrand (Ed.), *Economics of technology* (pp. 65–89). Elsevier.
DiPrete, T. A., & Eirich, G. M. (2006). Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. *Annual Review of Sociology*, *32*(1), 271–297.
Dusdal, J., & Powell, J. J. W. (2021). Benefits, motivations, and challenges of international collaborative research: A sociology of science case study. *Science and Public Policy*, *48*(1), 235–245.
Fox, M. F. (1983). Publication productivity among scientists: A critical review. *Social Studies of Science*, *13*(2), 285–305.
Fox, M. F., & Mohapatra, S. (2007). Social-organizational characteristics of work and publication productivity among academic scientists in doctoral-granting departments. *The Journal of Higher Education*, *78*(5), 542–571.
Fox, M. F., & Nikivincze, I. (2021). Being highly prolific in academic science: Characteristics of individuals and their departments. *Higher Education*, *81*, 1237–1255.
Geuna, A., & Shibayama, S. (2015). Moving out of academic research: Why do scientists stop doing research? In A. Geuna (Ed.), *Global mobility of research scientists* (pp. 271–297). Elsevier.
Hammarfelt, B. (2017). Recognition and reward in the academy: Valuing publication oeuvres in biomedicine, economics and history. *Aslib Journal of Information Management*, *69*(5), 607–623.
Heckman, J. J., & Moktan, S. (2018). *Publishing and promotion in economics. The tyranny of the Top Five.* NBER Working Paper 25093.
Hermanowicz, J. (2012). The sociology of academic careers: Problems and prospects. In J. C. Smart & M. B. Paulsen (Eds.), *Higher education: Handbook of theory and research* (pp. 207–248). Springer.
Horta, H., & Santos, J. M. (2016). The impact of publishing during PhD studies on career research publication, visibility, and collaborations. *Research in Higher Education*, *57*(1), 28–50.

Huang, J., Gates, A. J., Sinatra, R., & Barabási, A.-L. (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences*, *117*(9), 4609–4616.

Ioannidis, J. P. A., Boyack, K. W., & Klavans, R. (2014). Estimates of the continuously publishing core in the scientific workforce. *PLOS One*, *9*(7), e101698.

Jung, J. (2014). Research productivity by career stage among Korean academics. *Tertiary Education and Management*, *20*(2), 85–105.

Karimi, F., Wagner, C., Lemmerich, F., Jadidi, M., & Strohmaier, M. (2016). Inferring gender from names on the web: A comparative evaluation of gender detection methods. In *Proceedings of the 25th International Conference Companion on World Wide Web* (pp. 53–54). Montreal, Canada.

Kelchtermans, S., & Veugelers, R. (2013). Top research productivity and its persistence: Gender as a double-edged sword. *Review of Economics and Statistics*, *95*(1), 273–285.

King, M. M., Bergstrom, C. T., Correll, S. J., Jacquet, J., & West, J. D. (2017). Men set their own cites high: Gender and self-citation across fields and over time. *Socius*, *3*.

Kwiek, M. (2016). The European research elite: A cross-national study of highly productive academics across 11 European systems. *Higher Education*, *71*(3), 379–397

Kwiek, M. (2018). High research productivity in vertically undifferentiated higher education systems: Who are the top performers? *Scientometrics*, *115*(1), 415–462.

Kwiek, M. (2021). What large-scale publication and citation data tell us about international research collaboration in Europe: Changing national patterns in global contexts. *Studies in Higher Education, 46*(12), 2629–2649.

Kwiek, M., & Roszka, W. (2021a). Gender disparities in international research collaboration: A large-scale bibliometric study of 25,000 university professors. *Journal of Economic Surveys*, *35*(5), 1344–1388.

Kwiek, M., & Roszka, W. (2021b). Gender-based homophily in research: A large-scale study of man-woman collaboration. *Journal of Informetrics*, *15*(3), 1–38.

Kwiek, M., & Roszka, W. (2022). Academic vs. biological age in research on academic careers: A large-scale study with implications for scientifically developing systems. *Scientometrics*, *127*, 3543–3575.

Kwiek, M., & Roszka, W. (2024a). Once highly productive, forever highly productive? Full professors' research productivity from a longitudinal perspective. *Higher Education*, *87*, 519–549.

Kwiek, M., & Roszka, W. (2024b). Are scientists changing their research productivity classes when they move up the academic ladder? *Innovative Higher Education, Online first*, 1–40. https://doi.org/10.1007/s10755-024-09735-3.

Kwiek, M., & Szymula, Ł. (2024). Quantifying attrition in science: A cohort-based, longitudinal study of scientists in 38 OECD countries. *Higher Education*, *Online first*, 1–29. https://doi.org/10.1007/s10734-024-01284-0.

Kwiek, M., & Szymula, L. (2023). Young male and female scientists: A quantitative exploratory study of the changing demographics of the global scientific workforce. *Quantitative Science Studies*, *4*(4), 902–937.

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C.R. (2013). Global gender disparities in science. *Nature*, *504*, 211–213.

Latour B. & Woolgar S. (1986) *Laboratory life. The construction of scientific facts*. Princeton University Press.

Leišytė, L., & Dee, J. R. (2012). Understanding academic work in changing institutional environment. *Higher Education: Handbook of Theory and Research*, *27*, 123–206.

Li, W., Aste, T., Caccioli, F., & Livan, G. (2019). Early coauthorship with top scientists predicts success in academic careers. *Nature Communications, 10*, 5170.

Lindahl, J. (2018). Predicting research excellence at the individual level: The importance of publication rate, top journal publications, and top 10% publications in the case of early career mathematicians. *Journal of Informetrics*, *12*(2), 518–533.
Liu, L., Jones, B.F., Uzzi, B., & Wang, D.. (2023). Data, measurement and empirical methods in the science of science. *Nature Human Behaviour*, 7, 1046–1058.
Lutter, M., & Schröder, M. (2016). Who becomes a tenured professor, and why? Panel data evidence from German sociology, 1980–2013. *Research Policy*, *45*(5), 999–1013.
Ma, Y., Mukherjee, S., & Uzzi, B. (2020). Mentorship and protégé success in STEM fields. *Proceedings of the National Academy of Sciences*, *117*, 14077–14083.
Marginson, S. (2022) What drives global science? The four competing narratives. *Studies in Higher Education*, *47*(8), 1566–1584. 10.1080/03075079.2021.1942822
Menard, S. (2002). *Longitudinal research*. Sage.
Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago Press.
Ni, C., Smith, E., Yuan, H., Larivière, V., & Sugimoto, C. R. (2021). The gendered nature of authorship. *Science Advances*, 7, eabe4639.
Nielsen, M. W., & Andersen, J. P. (2021). Global citation inequality is on the rise. *Proceedings of the National Academy of Sciences*, *118*(7), e2012208118.
Nygaard, L. P., Aksnes, D. W., & Piro, F. N. (2022). Identifying gender disparities in research performance: The importance of comparing apples with apples. *Higher Education*, *84*, 1127–1142.
Preston, A. E. (2004). *Leaving science. Occupational exit from scientific careers*. Russell Sage Foundation.
Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *ArXiv*. https://arxiv.org/abs/2205.01833
Rørstad, K., & Aksnes, D. W. (2015). Publication rate expressed by age, gender and academic position – A large-scale analysis of Norwegian academic staff. *Journal of Informetrics*, *9*, 317–333.
Ross, M. B., Glennon, B.M., Murciano-Goroff, R., Berkes, E. G., Weinberg, B. A., & Lane, J.I. (2022). Women are credited less in science than men. *Nature*, *608*, 135–145.
Rowland, D. T. (2014). *Demographic methods and concepts*. Oxford University Press.
Ruiz-Castillo, J., & Costas, R. (2014). The skewness of scientific productivity. *Journal of Informetrics*, *8*(4), 917–934.
Ruspini, E. (1999). Longitudinal research and the analysis of social change. *Quality and Quantity, 33*(3), 219–227.
Salganik, M. J. (2018). *Bit by bit. Social research in a digital age*. Princeton University Press.
Santamaría, L., & Mihaljević, H. (2018). Comparison and benchmark of name-to-gender inference services. *PeerJ Computer Science, 4*, e156. https://doi.org/10.7717/ peerj- cs.156
Savage, W. E., & Olejniczak, A. J. (2021). Do senior faculty members produce fewer research publications than their younger colleagues? Evidence from Ph.D. granting institutions in the United States. Scientometrics, *126*, 4659–4686.
Sebo, P. (2021). Performance of gender detection tools: a comparative study of name-to-gender inference services. *Journal of the Medical Library Association*, *109*(3), 414–421.
Sebo, P. (2023). How well does NamSor perform in predicting the country of origin and ethnicity of individuals based on their first and last names? *PLOS One*, November 16, 2023, https://doi.org/10.1371/journal.pone.0294562
Shibayama, S., & Baba, Y. (2015). Impact-oriented science policies and scientific publication practices: The case of life sciences in Japan. *Research Policy, 44*(4), 936–950.
Shin, J. C., & Cummings, W. K. (2010). Multilevel analysis of academic publishing across disciplines: Research preference, collaboration, and time on research. *Scientometrics, 85*, 581–594.

Spoon, K. LaBerge, N., Wapman, K. H. , Zhang. S., Morgan. A. C., Galesic. M., Fosdick. B. K., Larremore, D. B., & Clauset. A. (2023). Gender and retention patterns among U.S. faculty. *Science Advances, 9*, eadi2205. 10.1126/sciadv.adi2205

Stephan, P. (2012). *How economics shapes science*. Harvard University Press.

Stephan, P. E., & Levin, S. G. (1992). *Striking the mother lode in science: The importance of age, place, and time.* Oxford University Press.

Sugimoto, C., & Larivière, V. (2018). *Measuring research: What everyone needs to know*. Oxford University Press.

Sugimoto, C., & Larivière, V. (2023). *Equity for women in science. Dismantling systemic barriers to advancement.* Harvard University Press.

Thelwall, M., & Sud, P. (2022). Scopus 1900–2020: Growth in articles, abstracts, countries, fields, and journals. *Quantitative Science Studies, 3*(1), 37–50.

Turner, L., & Mairesse, J. (2005). *Individual productivity differences in public research: How important are non-individual determinants? An econometric study of French physicists' publications and citations (1986–1997)*. CNRS.

Wagner, C. S. (2018). *The collaborative era in science. Governing the network*. Palgrave Macmillan.

Wang, D., & Barabási, A.-L. (2021). *The science of science*. Cambridge University Press.

Wang, Y., Jones, B. F., & Wang, D. (2019). Early career setback and future career impact. *Nature Communications, 10*, 4331.

Way, S. F., Morgan, A. C., Clauset, A., & Larremore, D. B. (2017). The misleading narrative of the canonical faculty productivity trajectory. *Proceedings of the National Academy of Sciences, 114*(44), E9216–E9223. 10.1073/pnas.1702121114

Weingart, P. (2004). Impact of bibliometrics upon the science system: inadvertent consequences? In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook on quantitative science and technology research* (pp. 117–131). Kluwer Academic Publishers.

Xie, Y. (2014). 'Undemocracy': Inequalities in science. *Science*, *344*(6186), 809–810.

Zeng, A., Shen, Z., Zhou, J., Wu, J., Fan, Y., Wang, Y., & Stanley, H. E. (2017). The science of science: from the perspective of complex systems. *Physics Reports, 714*, 1–73

Zhang, S., Wapman, K. H., Larremore, D. B., & Clauset, A. (2022). Labor advantages drive the greater productivity of faculty at elite universities. *Science Advances, 8*, eabq7056.

Electronic Supplementary Material to:

Quantifying Lifetime Productivity Changes: A Longitudinal Study of 325,000 Late-Career Scientists

Marek Kwiek
(1) Center for Public Policy Studies (CPPS), Adam Mickiewicz University of Poznan, Poznan, Poland, and
(2) German Center for Higher Education Research and Science Studies (DZHW), Berlin, Germany
kwiekm@amu.edu.pl, ORCID: orcid.org/0000-0001-7953-1063, corresponding author

Lukasz Szymula
(1) Faculty of Mathematics and Computer Science, Adam Mickiewicz University of Poznan, Poznan, Poland, and

(2) Department of Computer Science, University of Colorado Boulder, USA
ORCID: orcid.org/0000-0001-8714-096X

Supplementary Table 1. Structure of the sample of all nonoccasional (with at least 10 research articles in chapters in conference proceedings) OECD late-career (at least 25 years of publishing experience) scientists by gender, discipline, and country (N = 320,564) (frequencies and percentages)

		Total		Female scientists			Male scientists		
		N	col %	N	col %	row %	N	col %	row %
Total	**TOTAL**	**320,564**	**100**	**84,422**	**100**	**26.34**	**236,142**	**100**	**73.66**
	SOCIAL	**12,585**	**3,93**	**3,582**	**4.24**	**28.46**	**9,003**	**3.81**	**71.54**
	STEMM	**307,979**	**96,07**	**80,840**	**95.76**	**26.25**	**227,139**	**96.19**	**73.75**
Disciplines	AGRI	23,724	7.40	6,269	7.43	26.42	17,455	7.39	73.58
	BIO	45,813	14.29	14,526	17.21	31.71	31,287	13.25	68.29
	BUS	3,259	1.02	813	0.96	24.95	2,446	1.04	75.05
	CHEM	14,898	4.65	3,251	3.85	21.82	11,647	4.93	78.18
	COMP	7,644	2.38	1,187	1.41	15.53	6,457	2.73	84.47
	EARTH	14,370	4.48	2,536	3.00	17.65	11,834	5.01	82.35
	ECON	3,846	1.20	498	0.59	12.95	3,348	1.42	87.05
	ENG	12,814	4.00	1,166	1.38	9.10	11,648	4.93	90.90
	ENVIR	6,519	2.03	1,636	1.94	25.10	4,883	2.07	74.90
	IMMU	3,142	0.98	1,055	1.25	33.58	2,087	0.88	66.42
	MATER	5,839	1.82	1,139	1.35	19.51	4,700	1.99	80.49
	MATH	7,003	2.18	1,139	1.35	16.26	5,864	2.48	83.74
	MED	131,075	40.89	41,636	49.32	31.77	89,439	37.88	68.23
	NEURO	5,863	1.83	1,677	1.99	28.60	4,186	1.77	71.40
	PHYS	29,275	9.13	3,623	4.29	12.38	25,652	10.86	87.62
	PSYCH	5,480	1.71	2,271	2.69	41.44	3,209	1.36	58.56
Countries	United States	95,718	29.86	26,583	31.49	27.77	69,135	29.28	72.23
	Japan	29,358	9.16	2,953	3.50	10.06	26,405	11.18	89.94
	Italy	28,354	8.85	10,606	12.56	37.41	17,748	7.52	62.59
	UK	21,822	6.81	5,512	6.53	25.26	16,310	6.91	74.74
	France	21,129	6.59	6,313	7.48	29.88	14,816	6.27	70.12

	Germany	20,551	6.41	3,437	4.07	16.72	17,114	7.25	83.28
	Spain	12,978	4.05	4,436	5.25	34.18	8,542	3.62	65.82
	Canada	12,605	3.93	3,665	4.34	29.08	8,940	3.79	70.92
	Australia	10,374	3.24	3,118	3.69	30.06	7,256	3.07	69.94
	Netherlands	8,055	2.51	1,995	2.36	24.77	6,060	2.57	75.23
	Poland	5,619	1.75	1,901	2.25	33.83	3,718	1.57	66.17
	Sweden	4,894	1.53	1,372	1.63	28.03	3,522	1.49	71.97
	South Korea	4,847	1.51	627	0.74	12.94	4,220	1.79	87.06
	Switzerland	4,126	1.29	746	0.88	18.08	3,380	1.43	81.92
	Belgium	3,582	1.12	920	1.09	25.68	2,662	1.13	74.32
	Turkey	3,413	1.06	850	1.01	24.90	2,563	1.09	75.10
	Greece	3,412	1.06	873	1.03	25.59	2,539	1.08	74.41
	Israel	3,352	1.05	935	1.11	27.89	2,417	1.02	72.11
	Denmark	2,871	0.90	759	0.90	26.44	2,112	0.89	73.56
	Austria	2,808	0.88	561	0.66	19.98	2,247	0.95	80.02
	Rest	20,696	6.45	6,260	7.42	30.25	14,436	6.12	69.75

Supplementary Table 2. Structure of the sample of all nonoccasional (with at least 10 research articles in chapters in conference proceedings) OECD late-career (at least 25 years of publishing experience) scientists by academic age (publishing experience) and gender (N=320,564)

Academic age	Female scientists	Male scientists	% Female scientists	% Male scientist	Total
25	8,692	17,836	32.77	67.23	26,528
26	8,072	17,640	31.39	68.61	25,712
27	7,964	17,513	31.26	68.74	25,477
28	7,591	17,780	29.92	70.08	25,371
29	6,289	15,160	29.32	70.68	21,449
30	5,857	14,623	28.60	71.40	20,480
31	5,431	13,617	28.51	71.49	19,048
32	4,792	12,507	27.70	72.30	17,299
33	4,152	12,017	25.68	74.32	16,169
34	4,052	11,580	25.92	74.08	15,632
35	3,540	10,971	24.40	75.60	14,511
36	2,994	9,816	23.37	76.63	12,810
37	2,559	8,805	22.52	77.48	11,364
38	2,349	7,927	22.86	77.14	10,276
39	1,942	7,483	20.60	79.40	9,425
40	1,812	6,905	20.79	79.21	8,717
41	1,485	5,970	19.92	80.08	7,455
42	1,340	5,764	18.86	81.14	7,104
43	1,159	4,817	19.39	80.61	5,976
44	957	4,267	18.32	81.68	5,224
45	805	3,941	16.96	83.04	4,746
46	617	3,460	15.13	84.87	4,077
47	524	2,981	14.95	85.05	3,505
48	447	2,581	14.76	85.24	3,028
49	365	2,373	13.33	86.67	2,738
50	321	2,166	12.91	87.09	2,487

Supplementary Table 3. Cut-off points (publication numbers: articles and chapters in conference proceedings) for membership in productivity deciles (N = 320,564), late-career scientists at an early-career stage, by discipline

Discipline	Min	1	2	3	4	5	6	7	8	9	Max
AGRI	0.00	0.81	1.79	2.82	3.96	5.29	6.81	8.71	11.42	16.21	107.48
BIO	0.00	2.28	3.71	5.04	6.32	7.74	9.41	11.49	14.50	19.99	187.60
BUS	0.00	1.01	2.06	3.06	4.11	5.28	6.45	8.11	10.18	13.80	75.41
CHEM	0.00	1.78	3.40	5.10	6.83	8.71	10.99	14.12	18.25	25.49	175.49
COMP	0.00	0.45	1.12	1.82	2.61	3.49	4.59	6.07	8.27	12.06	84.03
EARTH	0.00	1.08	2.22	3.33	4.56	5.98	7.62	9.68	12.41	17.37	106.94
ECON	0.00	1.06	1.96	2.72	3.54	4.44	5.52	6.80	8.66	11.71	61.44
ENG	0.00	0.29	1.09	1.97	2.93	4.09	5.46	7.28	9.90	14.49	113.96
ENVIR	0.00	0.65	1.63	2.77	3.95	5.28	7.02	9.08	12.23	17.73	126.98
IMMU	0.00	2.10	3.70	5.15	6.58	8.14	9.83	11.80	14.94	19.99	107.99
MATER	0.00	0.79	2.06	3.53	5.25	7.08	9.20	12.07	16.03	23.42	114.32
MATH	0.00	0.97	1.81	2.62	3.48	4.36	5.54	7.09	9.06	12.76	74.59
MED	0.00	0.86	2.08	3.46	5.00	6.86	9.11	12.01	16.24	23.87	264.17
NEURO	0.00	2.32	3.63	4.80	6.01	7.31	8.86	10.88	13.91	19.50	151.95
PHYS	0.00	1.72	3.51	5.34	7.26	9.45	11.90	15.23	20.32	30.68	323.72
PSYCH	0.00	0.97	2.19	3.39	4.53	5.96	7.68	9.71	12.25	17.14	128.01
SOCIAL	**0.00**	**0.97**	**1.96**	**2.72**	**3.54**	**4.44**	**5.52**	**6.80**	**8.66**	**11.71**	**61.44**
STEMM	**0.00**	**0.29**	**1.09**	**1.82**	**2.61**	**3.49**	**4.59**	**6.07**	**8.27**	**12.06**	**74.59**
TOTAL	**0.00**	**0.29**	**1.09**	**1.82**	**2.61**	**3.49**	**4.59**	**6.07**	**8.27**	**11.71**	**61.44**

Supplementary Table 4. Cut-off points (publication numbers: articles and chapters in conference proceedings) for membership in productivity deciles (N = 320,564), late-career scientists at a mid-career stage, by discipline

Discipline	Min	1	2	3	4	5	6	7	8	9	Max
AGRI	0.00	1.90	3.57	5.44	7.60	10.08	12.95	16.82	22.52	32.50	293.78
BIO	0.00	3.19	5.40	7.48	9.71	12.22	15.25	19.11	24.91	35.51	403.08
BUS	0.00	1.86	3.09	4.43	5.82	7.26	8.97	11.13	14.17	19.54	89.39
CHEM	0.00	2.91	5.63	8.72	11.71	15.32	19.41	24.52	32.62	48.49	460.35
COMP	0.00	1.37	2.53	3.69	4.93	6.57	8.48	10.97	14.69	21.58	221.27
EARTH	0.00	2.27	4.45	6.63	9.00	11.62	14.74	18.88	24.74	35.22	261.89
ECON	0.00	1.29	2.35	3.29	4.37	5.54	6.83	8.58	11.11	15.72	123.88
ENG	0.00	1.06	2.46	4.03	5.82	7.86	10.57	14.14	19.25	28.71	227.00
ENVIR	0.00	1.98	3.85	5.87	8.38	10.89	14.20	18.40	24.99	37.74	409.99
IMMU	0.00	3.31	5.89	8.16	10.76	13.72	17.00	20.56	26.43	36.58	257.07
MATER	0.00	2.09	4.77	7.53	10.61	14.17	18.73	24.64	32.77	48.78	344.97
MATH	0.00	1.30	2.34	3.43	4.54	5.78	7.36	9.20	12.18	17.81	274.53
MED	0.00	1.69	3.62	5.80	8.39	11.61	15.54	20.86	28.91	44.38	940.30
NEURO	0.00	3.48	5.60	7.54	9.66	11.88	14.47	18.14	23.70	33.97	257.43
PHYS	0.00	2.77	5.35	8.04	11.00	14.56	18.93	24.89	34.58	57.35	1,012.33
PSYCH	0.00	1.68	3.42	5.17	7.24	9.46	11.94	15.57	20.63	29.22	150.48
SOCIAL	**0.00**	**1.29**	**2.35**	**3.29**	**4.37**	**5.54**	**6.83**	**8.58**	**11.11**	**15.72**	**89.39**
STEMM	**0.00**	**1.06**	**2.34**	**3.43**	**4.54**	**5.78**	**7.36**	**9.20**	**12.18**	**17.81**	**221.27**
TOTAL	**0.00**	**1.06**	**2.34**	**3.29**	**4.37**	**5.54**	**6.83**	**8.58**	**11.11**	**15.72**	**89.39**

Journal prestige–normalized (exponential) approach to individual research productivity

If a linear function was used, the value of an article that is published in the journal in the 99th journal percentile rank would be 0.99, which is slightly less than two articles published in the journals in the 50th journal percentile rank—which does not seem to properly reflect academic workload. The value of a paper used here increases slowly in lower-ranked journals (range 1–50) and steeply for top-ranked journals (range 90–99) so that the difference in values between articles published in the journals in the 99th and 50th percentile ranks is closer to fivefold (4.67). Within the four-year periods studied, we do not distinguish between changing percentile location of the journals in Scopus (the historic locations in the past are not available). We use the most recent (2023) Scopus percentile ranks as a proxy. For the vast majority of journals in STEMM, changes in percentile ranks are moderate. In Scopus, the journal ranking system that uses percentile ranks is based on the citations received by all publications from a journal in the previous four years. Hence, although journal percentile ranks are a proxy of quality (representing the overall impact of a journal rather than of a specific paper on the academic community), the individual articles in these journals are, on average, more or less cited.

In a non-normalized approach to productivity (full counting method), an article published in any journal traditionally receives a value of 1, whereas, in our prestige-normalized approach (exponential) using a full counting method, articles in journals with a percentile rank of 90 receive a value of 0.77; articles published in journals with percentile ranks of 50 receive a value of 0.18, and so forth.

The formula is as follows:

$$p_{adj(exp)}=(perc/100)^{2.5}$$

where $p_{adj(exp)}$ is the prestige adjusted article equivalent (exponential) and perc is the percentile of the journal in which the article was published, as assigned by Scopus. So with an output of four articles in the 50th percentile in a period of four years, productivity is 0.177 times 4 and divided by 4; with a single article in the 99th percentile, it is 0.975 divided by 4, that is, 0.248. The exponential approach imposes a “penalty” that is particularly severe for lower percentile ranks and diminishes as the journal’s percentile rank increases. The formula was first used in a previous study (Kwiek & Roszka, 2024b).

Supplementary Table 5. Mobility of bottom performers between two career stages: early career (initial stage) and mid-career (target stage): From which initial productivity deciles (at an early-career stage) do bottom-performing scientists at a mid-career stage come from? Late-career scientists who were bottom performers at a mid-career stage (N=32,063) by academic discipline and initial productivity decile (frequencies and percentages)

		Total	Bottom 10%	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Top 10%
Late-career scientists who were bottom performers at a mid-career stage												
TOTAL	**N**	**32,063**	**11,996**	**7,325**	**4,825**	**3,091**	**2,036**	**1,277**	**785**	**425**	**221**	**82**
	%	**100**	**37.41**	**22.85**	**15.05**	**9.64**	**6.35**	**3.98**	**2.45**	**1.33**	**0.69**	**0.26**
SOCIAL	**N**	**1,259**	**433**	**294**	**186**	**125**	**95**	**60**	**30**	**18**	**15**	**3**
	%	**100**	**34.39**	**23.35**	**14.77**	**9.93**	**7.55**	**4.77**	**2.38**	**1.43**	**1.19**	**0.24**
STEMM	**N**	**30,804**	**11,563**	**7,031**	**4,639**	**2,966**	**1,941**	**1,217**	**755**	**407**	**206**	**79**
	%	**100**	**37.54**	**22.82**	**15.06**	**9.63**	**6.30**	**3.95**	**2.45**	**1.32**	**0.67**	**0.26**
AGRI	N	2,373	916	556	365	239	143	77	42	18	13	4
	%	100	38.60	23.43	15.38	10.07	6.03	3.24	1.77	0.76	0.55	0.17
BIO	N	4,582	1,675	1,008	680	427	312	209	145	66	45	15
	%	100	36.56	22.00	14.84	9.32	6.81	4.56	3.16	1.44	0.98	0.33
BUS	N	326	99	73	48	30	30	18	15	5	7	1
	%	100	30.37	22.39	14.72	9.20	9.20	5.52	4.60	1.53	2.15	0.31
CHEM	N	1,490	611	361	226	129	78	47	12	17	8	1
	%	100	41.01	24.23	15.17	8.66	5.23	3.15	0.81	1.14	0.54	0.07
COMP	N	765	248	159	123	77	57	46	36	13	5	1
	%	100	32.42	20.78	16.08	10.07	7.45	6.01	4.71	1.70	0.65	0.13
EARTH	N	1,437	580	333	205	138	81	48	35	9	5	3
	%	100	40.36	23.17	14.27	9.60	5.64	3.34	2.44	0.63	0.35	0.21
ECON	N	385	115	98	64	39	28	18	10	8	4	1
	%	100	29.87	25.45	16.62	10.13	7.27	4.68	2.60	2.08	1.04	0.26
ENG	N	1,282	474	267	202	120	94	61	30	21	10	3
	%	100	36.97	20.83	15.76	9.36	7.33	4.76	2.34	1.64	0.78	0.23
ENVIR	N	652	232	135	107	72	44	23	23	9	4	3
	%	100	35.58	20.71	16.41	11.04	6.75	3.53	3.53	1.38	0.61	0.46
IMMU	N	315	113	81	39	27	22	9	12	6	4	2
	%	100	35.87	25.71	12.38	8.57	6.98	2.86	3.81	1.90	1.27	0.63
MATER	N	584	226	142	102	42	40	21	6	4	1	
	%	100	38.70	24.32	17.47	7.19	6.85	3.60	1.03	0.68	0.17	
MATH	N	701	254	179	97	66	47	31	14	7	4	2
	%	100	36.23	25.53	13.84	9.42	6.70	4.42	2.00	1.00	0.57	0.29
MED	N	13,108	4,767	2,989	2,039	1,322	856	525	325	188	76	21
	%	100	36.37	22.80	15.56	10.09	6.53	4.01	2.48	1.43	0.58	0.16
NEURO	N	587	230	125	91	61	22	26	18	8	4	2
	%	100	39.18	21.29	15.50	10.39	3.75	4.43	3.07	1.36	0.68	0.34
PHYS	N	2,928	1,237	696	363	246	145	94	57	41	27	22
	%	100	42.25	23.77	12.40	8.40	4.95	3.21	1.95	1.40	0.92	0.75
PSYCH	N	548	219	123	74	56	37	24	5	5	4	1
	%	100	39.96	22.45	13.50	10.22	6.75	4.38	0.91	0.91	0.73	0.18

Supplementary Table 6. Mobility of bottom performers between two career stages: mid-career (initial stage) and late career (target stage): From which initial productivity deciles (at a mid-career stage) do bottom-performing scientists at a late career stage come from? Late-career scientists who were bottom performers at a late-career stage (N=32,075) by academic discipline and initial productivity decile (frequencies and percentages)

		Total	Bottom 10%	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Top 10%
Late-career scientists who are bottom performers at a late career stage												
TOTAL	**N**	**32,075**	**9,836**	**7,051**	**5,052**	**3,473**	**2,450**	**1,691**	**1,127**	**726**	**447**	**222**
	%	**100**	**30.67**	**21.98**	**15.75**	**10.83**	**7.64**	**5.27**	**3.51**	**2.26**	**1.39**	**0.69**
SOCIAL	**N**	**1,259**	**400**	**254**	**201**	**154**	**89**	**70**	**47**	**21**	**14**	**9**
	%	**100**	**31.77**	**20.17**	**15.97**	**12.23**	**7.07**	**5.56**	**3.73**	**1.67**	**1.11**	**0.71**
STEMM	**N**	**30,816**	**9,436**	**6,797**	**4,851**	**3,319**	**2,361**	**1,621**	**1,080**	**705**	**433**	**213**
	%	**100**	**30.62**	**22.06**	**15.74**	**10.77**	**7.66**	**5.26**	**3.50**	**2.29**	**1.41**	**0.69**
AGRI	N	2,373	764	579	372	245	164	106	80	39	18	6
	%	100	32.20	24.40	15.68	10.32	6.91	4.47	3.37	1.64	0.76	0.25
BIO	N	4,582	1,295	905	751	486	392	287	194	126	106	40
	%	100	28.26	19.75	16.39	10.61	8.56	6.26	4.23	2.75	2.31	0.87
BUS	N	326	86	74	56	34	23	25	16	6	4	2
	%	100	26.38	22.70	17.18	10.43	7.06	7.67	4.91	1.84	1.23	0.61
CHEM	N	1,490	498	349	222	157	107	69	46	21	15	6
	%	100	33.42	23.42	14.90	10.54	7.18	4.63	3.09	1.41	1.01	0.40
COMP	N	767	217	177	121	77	66	51	30	17	9	2
	%	100	28.29	23.08	15.78	10.04	8.60	6.65	3.91	2.22	1.17	0.26
EARTH	N	1,437	503	323	211	150	87	73	46	24	16	4
	%	100	35.00	22.48	14.68	10.44	6.05	5.08	3.20	1.67	1.11	0.28
ECON	N	385	112	68	65	54	32	24	14	10	3	3
	%	100	29.09	17.66	16.88	14.03	8.31	6.23	3.64	2.60	0.78	0.78
ENG	N	1,282	383	279	206	168	111	65	34	22	8	6
	%	100	29.88	21.76	16.07	13.10	8.66	5.07	2.65	1.72	0.62	0.47
ENVIR	N	652	187	140	103	79	54	41	23	16	6	3
	%	100	28.68	21.47	15.80	12.12	8.28	6.29	3.53	2.45	0.92	0.46
IMMU	N	319	113	59	38	34	27	18	16	7	7	
	%	100	35.42	18.50	11.91	10.66	8.46	5.64	5.02	2.19	2.19	
MATER	N	590	190	137	88	74	38	30	19	8	4	2
	%	100	32.20	23.22	14.92	12.54	6.44	5.08	3.22	1.36	0.68	0.34
MATH	N	701	227	158	121	75	52	30	21	8	7	2
	%	100	32.38	22.54	17.26	10.70	7.42	4.28	3.00	1.14	1.00	0.29
MED	N	13,108	3,864	2,900	2,084	1,451	1,027	685	466	332	192	107
	%	100	29.48	22.12	15.90	11.07	7.83	5.23	3.56	2.53	1.46	0.82
NEURO	N	587	184	123	84	70	39	29	25	19	12	2
	%	100	31.35	20.95	14.31	11.93	6.64	4.94	4.26	3.24	2.04	0.34
PHYS	N	2,928	1,011	668	450	253	197	137	80	66	33	33
	%	100	34.53	22.81	15.37	8.64	6.73	4.68	2.73	2.25	1.13	1.13
PSYCH	N	548	202	112	80	66	34	21	17	5	7	4
	%	100	36.86	20.44	14.60	12.04	6.20	3.83	3.10	0.91	1.28	0.73

Supplementary Table 7. Inverted correlation matrix, main diagonal, model of odds ratio estimates of membership in the class of top productive mid-career

Variable	AGRI	BIO	BUS	CHEM	COMP	EARTH	ECON	ENG	ENVIR	IMMU	MATER	MATH	MED	NEURO	PHYS	PSYCH
Male	1.012	1.021	1.013	1.015	1.000	1.009	1.010	1.000	1.012	1.028	1.008	1.016	1.011	1.018	1.005	1.026
Avg. FWCI 4y Early	1.145	1.062	1.041	1.058	1.031	1.097	1.058	1.068	1.135	1.103	1.101	1.096	1.040	1.050	1.065	1.040
Inter. Collab. Rate Early	1.102	1.133	1.038	1.105	1.029	1.245	1.021	1.052	1.112	1.153	1.152	1.026	1.103	1.103	1.446	1.040
Median Team Size Early	1.110	1.163	1.054	1.101	1.024	1.284	1.047	1.042	1.142	1.198	1.168	1.060	1.097	1.145	1.435	1.059
Top Early	1.070	1.031	1.030	1.064	1.019	1.070	1.053	1.035	1.058	1.081	1.073	1.072	1.028	1.061	1.163	1.042

Supplementary Table 8. Inverted correlation matrix, main diagonal, model of odds ratio estimates of membership in the class of bottom productive mid-career

Variable	AGRI	BIO	BUS	CHEM	COMP	EARTH	ECON	ENG	ENVIR	IMMU	MATER	MATH	MED	NEURO	PHYS	PSYCH
Male	1.007	1.016	1.005	1.012	1.000	1.004	1.009	1.000	1.008	1.015	1.005	1.016	1.005	1.008	1.005	1.013
Avg. FWCI 4y Early	1.160	1.058	1.070	1.053	1.042	1.130	1.065	1.085	1.161	1.091	1.111	1.091	1.042	1.068	1.063	1.073
Inter. Collab. Rate Early	1.101	1.137	1.048	1.103	1.046	1.250	1.034	1.067	1.111	1.161	1.159	1.050	1.108	1.108	1.416	1.048
Median Team Size Early	1.133	1.168	1.085	1.111	1.056	1.283	1.054	1.086	1.196	1.202	1.224	1.065	1.152	1.131	1.400	1.088
Bottom Early	1.123	1.036	1.112	1.066	1.094	1.113	1.086	1.141	1.169	1.074	1.179	1.108	1.098	1.065	1.047	1.116

Supplementary Table 9. Inverted correlation matrix, main diagonal, main diagonal, and model of odds ratio estimates of membership in the class of top productive late career

Variable	AGRI	BIO	BUS	CHEM	COMP	EARTH	ECON	ENG	ENVIR	IMMU	MATER	MATH	MED	NEURO	PHYS	PSYCH
Male	1.016	1.019	1.012	1.015	1.002	1.010	1.011	1.003	1.015	1.022	1.011	1.015	1.010	1.021	1.005	1.030
Avg. FWCI 4y Mid	1.165	1.058	1.036	1.079	1.045	1.080	1.054	1.056	1.105	1.135	1.135	1.086	1.059	1.111	1.177	1.075
Inter. Collab. Rate Mid	1.152	1.176	1.076	1.153	1.069	1.264	1.056	1.085	1.185	1.177	1.175	1.034	1.169	1.142	1.499	1.087
Median Team Size Mid	1.149	1.194	1.062	1.136	1.053	1.320	1.081	1.050	1.142	1.243	1.197	1.054	1.153	1.183	1.492	1.096
TOP200	1.015	1.009	1.001	1.013	1.008	1.013	1.010	1.020	1.015	1.010	1.017	1.014	1.016	1.014	1.005	1.021
Top Mid	1.083	1.041	1.039	1.080	1.034	1.075	1.052	1.053	1.067	1.077	1.106	1.059	1.048	1.066	1.207	1.074

Supplementary Table 10. Inverted correlation matrix, main diagonal, main diagonal, model of odds ratio estimates of membership in the class of bottom productive late career

Variable	AGRI	BIO	BUS	CHEM	COMP	EARTH	ECON	ENG	ENVIR	IMMU	MATER	MATH	MED	NEURO	PHYS	PSYCH
Male	1.013	1.013	1.006	1.012	1.002	1.007	1.011	1.003	1.013	1.011	1.008	1.014	1.004	1.010	1.005	1.023
Avg. FWCI 4y Mid	1.167	1.051	1.047	1.060	1.046	1.088	1.063	1.064	1.122	1.123	1.114	1.086	1.058	1.112	1.151	1.118
Inter. Collab. Rate Mid	1.143	1.176	1.085	1.150	1.073	1.273	1.060	1.085	1.173	1.181	1.172	1.055	1.167	1.141	1.484	1.090
Median Team Size Mid	1.160	1.198	1.076	1.141	1.059	1.314	1.089	1.058	1.150	1.241	1.211	1.055	1.187	1.176	1.460	1.109
TOP200	1.011	1.007	1.001	1.011	1.008	1.011	1.011	1.015	1.011	1.010	1.015	1.015	1.011	1.014	1.006	1.018
Bottom Mid	1.080	1.029	1.078	1.056	1.051	1.085	1.079	1.075	1.084	1.050	1.093	1.095	1.077	1.041	1.055	1.138

Supplementary Table 11. Logistic regression statistics: odds ratio estimates of membership in the class of global bottom productive mid-career (the bottom 10%, separately for each academic discipline) (N = 320,564)

Model	AGRI	BIO	BUS	CHEM	COMP	EARTH	ECON	ENG	ENVIR	IMMU	MATER	MATH	MED	NEURO	PHYS	PSYCH
R^2	0.14	0.10	0.08	0.13	0.08	0.14	0.08	0.11	0.11	0.10	0.12	0.12	0.10	0.14	0.14	0.16
Male	1.20 (2)	0.72 (3)	1.32	1.21	1.04	1.15	1.08	1.48 (1)	1.31	0.68 (2)	1.22	1.01	1.02 (3)	0.80 (3)	1.14	1.01 (2)
Avg. FWCI 4y Early	0.50 (3)	0.88 (3)	0.84 (2)	0.77	0.80	0.70	0.61	0.73 (3)	0.69	0.79 (3)	0.75 (1)	0.55	0.98 (3)	0.59 (3)	0.97 (3)	0.56 (3)
Inter. Collab. Rate Early	1.00 (3)	0.99 (3)	0.99 (3)	0.99 (3)	1.00 (3)	0.99 (3)	1.00 (3)	0.99 (3)	0.99 (3)	0.99 (3)	0.99 (3)	0.99 (3)	0.99 (3)	0.99 (3)	0.99 (3)	0.99 (3)
Median Team Size Early	0.95 (3)	1.02 (3)	0.85 (1)	1.02 (1)	0.98	0.99 (1)	0.86 (1)	1.13 (3)	0.99 (2)	1.09 (3)	0.96 (3)	0.89	0.98 (3)	0.99 (3)	1.01 (3)	0.89 (3)
Bottom Early	4.70 (3)	6.49 (3)	3.39	7.97	4.23	5.71	3.12	5.72 (3)	4.17	5.86 (2)	5.56	4.24	6.28 (3)	5.76 (3)	8.67	4.31 (1)
Constant	0.18 (1)	0.13 (3)	0.16	0.10	0.13	0.15	0.22	0.06	0.14	0.11	0.13	0.24	0.10 (3)	0.24 (2)	0.09	0.26

Note: (1) = p-value ≤ 0.05; (2) = p-value ≤ 0.01; (3) = p-value ≤ 0.001

Supplementary Table 12. Logistic regression statistics: odds ratio estimates of membership in the class of global bottom productive late career (the bottom 10%, separately for each academic discipline) (N = 320,564)

Model	AGRI	BIO	BUS	CHEM	COMP	EARTH	ECON	ENG	ENVIR	IMMU	MATER	MATH	MED	NEURO	PHYS	PSYCH
R^2	0.11	0.06	0.08	0.10	0.06	0.11	0.09	0.09	0.07	0.11	0.10	0.10	0.08	0.09	0.11	0.13
Male	1.36	1.00 (3)	1.35	1.23	1.15	1.13	1.09	1.35	1.29	0.72	1.07	1.02	1.17 (3)	0.90	1.21	1.18
Avg. FWCI 4y Mid	0.40	0.84 (3)	0.71	0.62 (1)	0.86	0.68 (3)	0.63 (2)	0.65 (2)	0.65	0.68 (1)	0.54	0.48 (1)	0.91 (3)	0.62 (1)	0.94 (1)	0.52
Inter. Collab. Rate Mid	1.00 (3)	0.99 (3)	0.99	0.99 (3)	1.00 (3)	1.00 (3)	1.00 (3)	0.99 (3)	0.99 (3)	1.00 (3)	0.99 (3)	1.00 (3)	0.99 (3)	0.99 (3)	0.99 (3)	0.99 (3)
Median Team Size Mid	0.97 (2)	0.97 (3)	0.82	0.98 (3)	0.89	0.96 (3)	0.71 (2)	1.16 (3)	1.01 (2)	1.01 (3)	1.00 (1)	0.89 (3)	0.95 (3)	0.96 (3)	1.04 (3)	0.94 (2)
TOP200	0.66	0.76 (2)	0.83	0.68	0.93	0.69	1.06	0.45	0.72	0.63	0.67	0.67	0.59 (3)	0.77	0.66	0.87
Bottom Mid	3.14	3.68 (2)	2.28	4.34	3.80	4.51	2.80	3.36	3.04	5.24	3.40	3.54	3.46 (3)	3.85	4.96	3.95
Constant	0.23	0.20 (1)	0.30	0.17	0.15	0.19	0.33	0.09	0.16	0.21	0.22	0.26	0.17 (2)	0.30	0.12	0.23

Note: (1) = p-value ≤ 0.05; (2) = p-value ≤ 0.01; (3) = p-value ≤ 0.001