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

Abstract

This study focuses on persistence in research productivity over the course of an individual's entire scientific careers. We track "late-career" scientists (N=324,463) in 16 STEMM disciplines (science, technology, engineering, mathematics, and medicine) from 38 OECD countries for up to five decades. We examine the details of their mobility patterns between the top, middle, and bottom productivity classes. Methodologically, we turn a large-scale publication and citation bibliometric dataset into a comprehensive, longitudinal data source for research on careers in science. The global science system emerges as highly immobile: 60% of global top performers continue their careers as top performers and half of global bottom performers as bottom performers. Jumpers-Up and Droppers-Down are extremely rare in science. Our regression analyses show that productivity is highly path dependent: for all disciplines examined, there is a single most important predictor of being a top performer: 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—over the course of an entire scientific career. We are tracking "late-career" scientists (N=324,463) from 38 OECD countries for up to five decades to examine their mobility patterns between the top, middle, and bottom productivity classes.

We turn large-scale publication and citation bibliometric dataset into a global, comprehensive, multidimensional, and longitudinal data source for research on careers in science. 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 scientist in our sample, restricting our research to 16 STEMM disciplines (science, technology, engineering, mathematics, and medicine). In our context, "late-career" scientists are defined as individuals 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[1,2,3]: the same individuals are tracked over time over the 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.,[4,5,6]). We are testing the usefulness of publication and citation metadata by examining the global science profession from a longitudinal perspective; these metadata are in fact digital traces left by scientists throughout their professional lives as long as they keep publishing. Digital traces[7,8] allow for the emergence of the whole new multidisciplinary field of science of science[9,10,11], hence allowing science career studies to radically move beyond traditional small-scale surveys and interviews[12,13]. Digital traces left by scientists in global datasets allow researchers to change their focus from single national science systems to a global science system[5,14,15].

The present study explores mobility between the three individual productivity classes—top, middle, and bottom (constructed according to the 20/60/20 formula)—throughout long academic careers encompassing early, mid-, and late-career periods. Our initial hypotheses, which are based on research productivity literature[16,17,18], especially high research productivity literature focused on "top performers" and "prolific" scientists[19,20,21,22,23] are, first, that scientists are generally locked in within their productivity classes for years[18,24]; second, we argue that the elite strata of highly productive scientists often continue their whole careers as highly productive[25,26]; and, finally, we argue that radical changes in productivity classes, especially upwards, although popular in narratives about academic careers, are highly improbable in practice because of the cumulative nature of advantages and disadvantages in careers, as shown over the decades in the traditional sociology of science[27,28,29].

The current study follows research lines explored in science of science, which provides data-driven insights into the inner workings of science[11]. A shift toward new digitalized data sources allows for the exploration of new questions about scientists[8]. Traditional cross-sectional studies can be complemented with longitudinal studies[30,31] 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 an unprecedented level of detail. As a result, the various aspects of academic careers have recently been examined both globally (gender disparities in careers[4,14]; continuous publishing[32]; collaboration with top scientists[22]; gendered nature of authorship[15]; women in science[33]) and nationally, especially in the US (e.g., productivity across career stages[34]; long-term effects on careers of initial setbacks[35]; careers in elite universities[36]; credit distribution[37]), at a scale unthinkable in career studies before.

Productivity Classes in Single-Nation Studies

We have tested our initial hypotheses in a national-level strand of research under the general labels of "once highly productive, forever highly productive"[38,39]. The patterns found consistently supported our initial intuitions about immobility in the system: the majority of highly productive assistant and associate professors continued their careers as highly productive associate and full professors. In the present research, we develop our methodological approach to study mobility between productivity classes (computed for four major productivity types) of scientists from 38 OECD countries, mostly powerfully involved in the ongoing globalization of science[40,41]. We track 324,463 late-career scientists from a wide variety of research systems, which gives us the potential to test hypotheses about the scientific profession more generally.

There are three small-scale longitudinal single-nation studies that are similar to ours. First, for 497 French physicists, Turner and Mairesse[18] 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 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[24] 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 and showed strong support for an accumulative process, which disadvantages scientists with low initial output and advantages highly productive scientists. Finally, Abramo et al.[42] 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 disciplines; and we use a longitudinal and classificatory approach combined with two-dimensional analyses and logistic regression models.

Persistence in High (and Low) Productivity

The cumulative advantage theory of productivity highlights that "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"[25]. Substantial predetermined differences among scientists may have a powerful impact on careers[17,27]. An "initial success" may lead to increased productivity; in contrast, a "bad start" may lead to leaving science[18]. According to the "sacred spark" hypothesis, "there are substantial, predetermined differences among scientists in their ability and motivation to do creative scientific research"[25]. Scientists with the spark are always productive, and a differential distribution of talent affects inequality in productivity more than the recognition system in science[43].

As a result, stratification 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"[26]. 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[24].

Research Questions

At a global scale, we test whether persistence in research productivity over an academic lifetime is as prevalent as traditional sociology and economics of science (using analyses based on small-scale surveys and limited numbers of interviews) claimed for decades[16,27,28,44]. Tracking the career trajectories of thousands of scientists, we seek otherwise invisible, global mobility patterns (whenever we use the term "global," we in fact refer to 38 OECD countries).

Using a large-scale longitudinal dataset, we also test the usefulness of the two traditional conceptions of science that explain individual successes and failures: the accumulation of advantages and disadvantages over scientific lifetime (the cumulative advantage theory of productivity) and individual-level properties (the sacred spark theory of productivity).

We have posed the following research questions: 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 upwards or downwards globally? Finally, what are cross-disciplinary and cross-productivity differences in mobility patterns between global productivity classes?

Results

Horizontal and Vertical Mobility Patterns Between Productivity Classes

Our sample includes all late-career scientists active in scholarly publishing in 2022 and who come from 16 STEMM disciplines and 38 OECD countries, with nonoccasional publishing status (at least three journal articles or publications in conference proceedings, lifetime). However, by definition, late-career scientists have once been both early career (years 5–14 of publishing experience) and mid-career (years 15–24 of publishing experience) scientists. The early-career scientists from the three productivity classes may retain or change their classes while being mid-career scientists; and then, they may retain or change their classes again while being late-career scientists. We follow 324,643 scientists and compare their individual research productivity with that of their peers of the same academic career stage and within the same discipline (see flowchart in Figure 6 in Methods). The structure of the sample by discipline, by academic age and gender, and by country is shown in Supplementary Tables 1 through 3).

Our analysis has been performed at the micro-level of individuals; we have full data on productivity for every scientist in every stage of their careers. Specifically, we analyze the following mobility types (and scientists' types) by discipline and productivity type: (1) top performers and their top-to-top mobility; (2) bottom performers and their bottom-to-bottom mobility; and (3) Jumpers-Up and Droppers-Down and their extreme upward and extreme downward mobility: bottom-to-top mobility and top-to-bottom mobility.

Mobility between Productivity Classes: All Disciplines Combined

Figure 1 presents the lifetime career trajectories of 324,643 late-career scientists in the combined 16 STEMM disciplines. Scientists can remain in their global productivity classes or move up or down; a Sankey diagram shows the percentage flows between the three productivity classes between early career, mid-career, and late-career periods.

The compelling power of the mobility patterns found comes from the power of the microdata at the level of individuals traced over time across their careers in this longitudinal study design.

Every late-career scientist currently (annual productivity in the five-year period of 2018–2022) allocated to the global top, middle, or bottom productivity classes is also retrospectively allocated to some (the same or different) productivity classes when they were early career and mid-career scientists. We move across time by examining scientists and their annual productivity determined for their three career stages and analyze their individual mobility patterns.

There is a clear pattern of persistence in membership in both the global bottom and top productivity classes for early career scientists as they progress up the professional ladder. In Productivity 1, more than half of the scientists in the top productivity classes remained in the same classes in both the early to mid-career transition and mid- to late-career transition, which is presented by thick left-to-right horizontal flows. Top-to-top mobility reaches 60.2% for top performers in the early career stage and 60.1% for top performers in the mid-career stage. Similarly, bottom-to-bottom mobility reaches 55.9% for bottom performers in the early career stage and 47.0% for bottom performers in the mid-career stage. These transitions differ only slightly for the other three productivity types (see the details in Table 1), as shown in Supplementary Figure 2. No matter which productivity type is used in calculations, the general mobility patterns for all disciplines combined are similar.

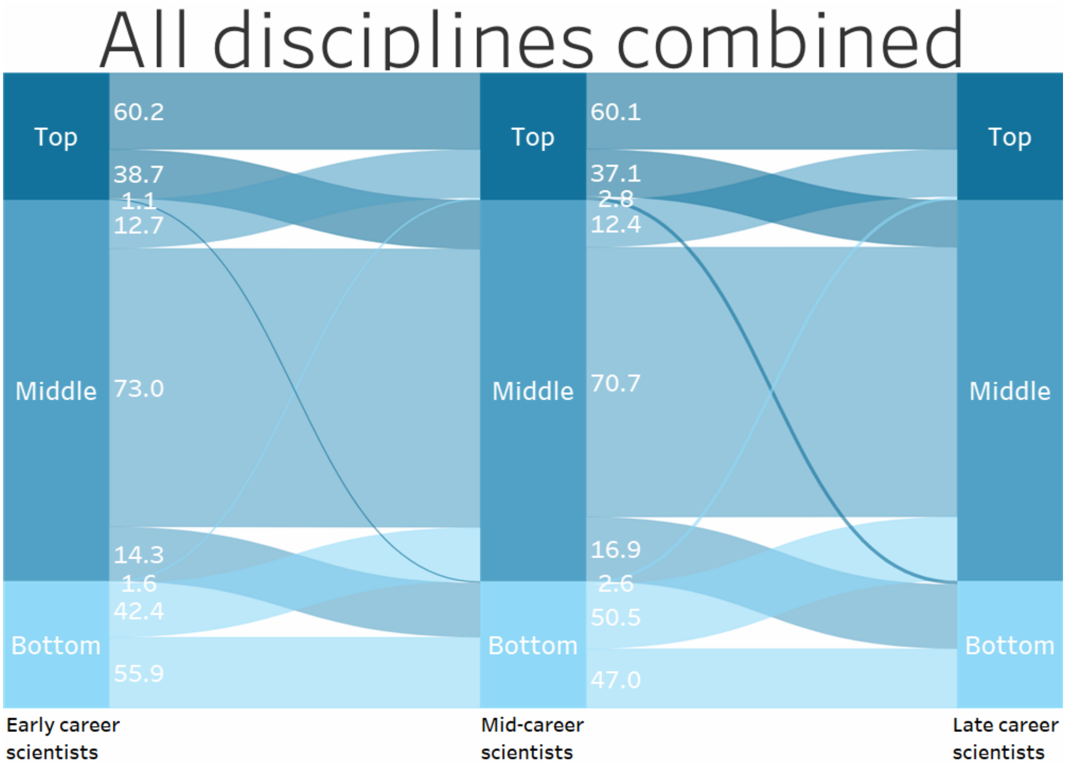


Figure 1. Sankey diagram of retrospectively constructed mobility between productivity classes in the three career stages. The diagram shows mobility from an early to mid- to late-career stage, with the three options of productivity classes (top, middle, and bottom) at each stage. The flows show percentage changes between productivity classes. Productivity 1 (prestige-normalized, full-counting approach). All STEMM disciplines combined, only current nonoccasional OECD late-career scientists (N=324,643)

There is a low likelihood of jumping up (bottom-to-top transition, Jumpers-Up) or of dropping down (top-to-bottom transition, Droppers-Down) over a lifetime, which is presented with thin ascending and descending flows in Figure 1. The percentages of scientists involved in either of these transitions between the three stages (early career to mid-career, mid-career to late career) in Productivity 1 range from 1.1% to 2.8% (1.1–3.1% in Productivity 2, 1.5–3.5% in Productivity 3, and 1.2–3.9% in Productivity 4). Table 1 provides full data on the numbers of scientists in each productivity class and the details of their transitions at a general level of all disciplines combined: the top panel presents the data on mobility from early career to mid-career stage; the middle panel refers to mid-career to late-career stage; and the bottom panel describes the sample distribution of late-career scientists by productivity class. We have full data on productivity for every scientist in every stage of their careers in each of the four productivity type in our dataset.

Specifically, in terms of moving radically upwards in productivity classes in the case of Productivity 1 (prestige-normalized, full counting), only 1.6% of bottom performers in the early career stage became top performers in the mid-career stage (and 2.6% of bottom performers in the mid-career stage became top performers in the late-career stage). This comes out to just 1,057 scientists (out of 65,023) and 1,673 (out of 64,971) in both cases, respectively. At the same time, in terms of moving radically downwards, only 1.1% of top performers in the early career stage became bottom performers in the mid-career stage (and 2.8% of top performers in the mid-career stage became bottom performers in the late-career stage).

Overall, the data suggest (Table 1) that the persistence of global membership in top and bottom productivity classes, as well as the likelihood of jumping up from bottom to top or dropping down from top to bottom classes, varies only slightly across the different productivity types. At the level of all disciplines combined, the patterns of mobility across productivity classes and across the four productivity types remain relatively consistent.

Table 1. Mobility between retrospectively constructed productivity classes in the three career stages. Overview: **Productivity 1–4.** Current nonoccasional OECD late-career scientists only, all disciplines combined, top class transitions shaded (N=324,643). The differences in mobility patterns by productivity type are relatively small: they are in the range of 1–4 percentage points (p.p.) for top-to-top and bottom-to-bottom mobility and in the range of 0.2 p.p. and 0.5 p.p. for bottom-to-top mobility (Jumpers-Up) and 0.4 p.p. and 1.1 p.p. for top-to-bottom mobility (Droppers-Down). For instance, the top-to-top mobility across Productivity 1 through Productivity 4 is 60.2%, 60.6%, 58.8% and 60.0%, respectively (early career to mid-career); and the bottom-to-bottom mobility is 47.0%, 47.8%, 51.1%, and 47.0%, respectively, for the same career stage mobility. The productivity mobility patterns found are stunningly similar across the four productivity types.

				Productivity 1 (prestige-normalized, full counting)			Productivity 2 (prestige-normalized, fractional counting)			Productivity 3 (non-normalized, full counting)			Productivity 4 (non-normalized, fractional counting)		
Transition from source academic position	Transition from productivity class	Transition to target academic position	Transition to productivity class	Number of scientists in transition	Number of scientists in each productivity class	%	Number of scientists in transition	Number of scientists in each productivity class	%	Number of scientists in transition	Number of scientists in each productivity class	%	Number of scientists in transition	Number of scientists in each productivity class	%
Early career	Bottom	Mid-career	Bottom	36,373	65,023	55.9	36,308	64,934	55.9	41,633	72,877	57.1	36,716	65,330	56.2
Early career	Bottom	Mid-career	Middle	27,593	65,023	42.4	27,608	64,934	42.5	29,909	72,877	41.0	27,539	65,330	42.2
Early career	Bottom	Mid-career	Top	1,057	65,023	1.6	1,018	64,934	1.6	1,335	72,877	1.8	1,075	65,330	1.7
Early career	Middle	Mid-career	Bottom	27,867	194,697	14.3	27,929	194,778	14.3	28,496	187,829	15.2	27,490	194,394	14.1
Early career	Middle	Mid-career	Middle	142,042	194,697	73.0	142,302	194,778	73.1	134,542	187,829	71.6	141,982	194,394	73.0
Early career	Middle	Mid-career	Top	24,788	194,697	12.7	24,547	194,778	12.6	24,791	187,829	13.2	24,922	194,394	12.8
Early career	Top	Mid-career	Bottom	731	64,923	1.1	695	64,931	1.1	938	63,937	1.5	751	64,919	1.2
Early career	Top	Mid-career	Middle	25,109	64,923	38.7	24,871	64,931	38.3	25,395	63,937	39.7	25,243	64,919	38.9
Early career	Top	Mid-career	Top	39,083	64,923	60.2	39,365	64,931	60.6	37,604	63,937	58.8	38,925	64,919	60.0
Mid-career	Bottom	Late career	Bottom	30,508	64,971	47.0	31,015	64,932	47.8	36,299	71,067	51.1	30,513	64,957	47.0
Mid-career	Bottom	Late career	Middle	32,790	64,971	50.5	32,131	64,932	49.5	32,998	71,067	46.4	32,468	64,957	50.0
Mid-career	Bottom	Late career	Top	1,673	64,971	2.6	1,786	64,932	2.8	1,770	71,067	2.5	1,976	64,957	3.0
Mid-career	Middle	Late career	Bottom	32,898	194,744	16.9	31,904	194,781	16.4	34,382	189,846	18.1	32,156	194,764	16.5
Mid-career	Middle	Late career	Middle	137,633	194,744	70.7	137,536	194,781	70.6	131,680	189,846	69.4	136,391	194,764	70.0
Mid-career	Middle	Late career	Top	24,213	194,744	12.4	25,341	194,781	13.0	23,784	189,846	12.5	26,217	194,764	13.5
Mid-career	Top	Late career	Bottom	1,787	64,928	2.8	2,027	64,930	3.1	2,256	63,730	3.5	2,507	64,922	3.9
Mid-career	Top	Late career	Middle	24,102	64,928	37.1	25,101	64,930	38.7	23,968	63,730	37.6	25,799	64,922	39.7
Mid-career	Top	Late career	Top	39,039	64,928	60.1	37,802	64,930	58.2	37,506	63,730	58.9	36,616	64,922	56.4
Late career	Bottom			65,193	65,193	100	64,946	64,946	100	72,937	72,937	100	65,176	65,176	100
Late career	Middle			194,525	194,525	100	194,768	194,768	100	188,646	188,646	100	194,658	194,658	100
Late career	Top			64,925	64,925	100	64,929	64,929	100	63,060	63,060	100	64,809	64,809	100

Mobility Between Productivity Classes: Cross-Disciplinary Differences

First, we briefly focus on Medicine (MED), the largest discipline in our sample, with N = 137,748 late-career scientists in 2022, and Materials Science (MATER), a relatively small (N = 5,898) discipline with very high top-to-top and bottom-to-bottom mobility (Figure 2). Although MED is generally in the upper ranges of top-to-top mobility (Figure 5) and mid ranges of bottom-to-bottom mobility (Supplementary Figure 3) for both transition periods, regardless of the productivity type used, MATER is in the top ranges, especially for bottom-to-bottom mobility. In MATER, the share of Jumpers-Up is relatively low, in the 1.10–1.53% range for the first transition period (and 2.97–3.30% for the second transition period). Interestingly, no scientists in MATER belong to Droppers-Down in Productivity 1 (0.76–1.21% in the other productivity types). The productivity patterns of the two contrasted disciplines are much different, with different career opportunities for scientists in top and bottom productivity classes early in their careers.

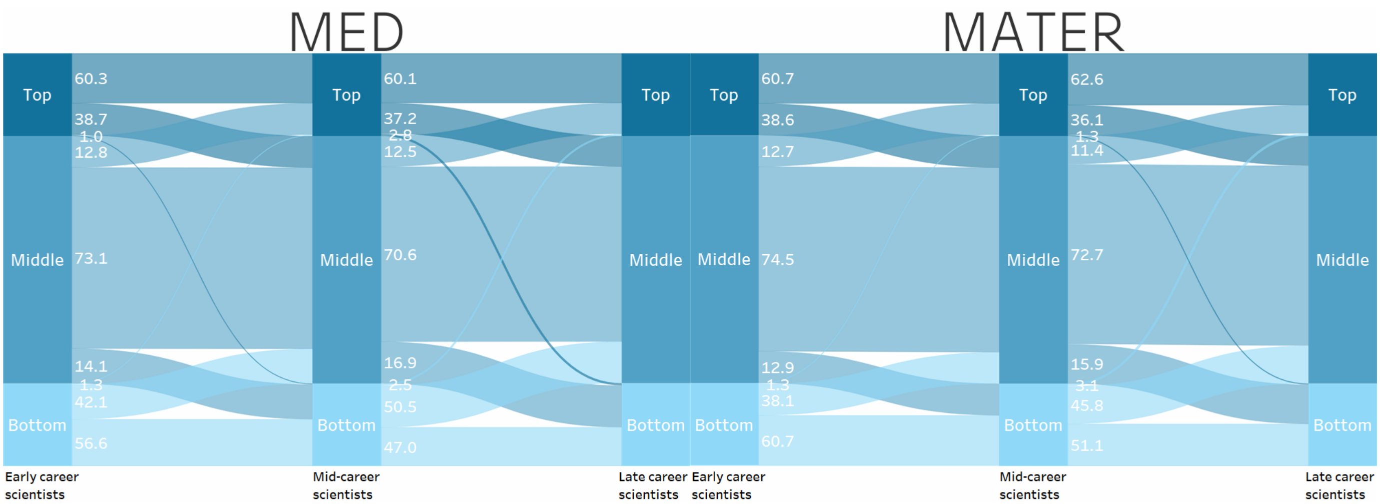


Figure 2. Sankey diagram of retrospectively constructed mobility between productivity classes in the three career stages, Medicine (MED, N = 137,748) and Materials Science (MATER, N = 5,898). Productivity 1 (prestige-normalized, full-counting approach). All STEMM disciplines combined, only current nonoccasional OECD late-career scientists.

Figure 3 presents Sankey diagrams for the three stages of academic careers for Productivity 1 (prestige-normalized, full counting) and visualizes the cross-disciplinary differences for all 16 STEMM disciplines; Supplementary Figure 1 presents mobility patterns for Productivity 2. The transitions across the 16 STEMM disciplines are similar—with some small but noticeable differences. Figure 4 shows the details of cross-disciplinary differences in top-to-top and bottom-to-bottom mobility for each discipline by productivity type for the transition between the early career to mid-career stage, and Figure 5 – for extreme vertical mobility; Supplementary Figure 3 shows the transition between the mid- to late-career stage. Strikingly, for all disciplines combined (Total in Figure 5), the percentage of scientists who moved from top classes to top classes differs marginally between the four productivity types (the 59–61% range), as does the percentage of scientists who moved from bottom classes to bottom classes (the 55–57% range).

The mobility patterns found in lifetime academic careers are complex and vary significantly across disciplines. Although the horizontal mobility patterns are relatively consistent across disciplines, the extreme vertical mobility patterns (bottom to top, top to bottom) show much greater variability.

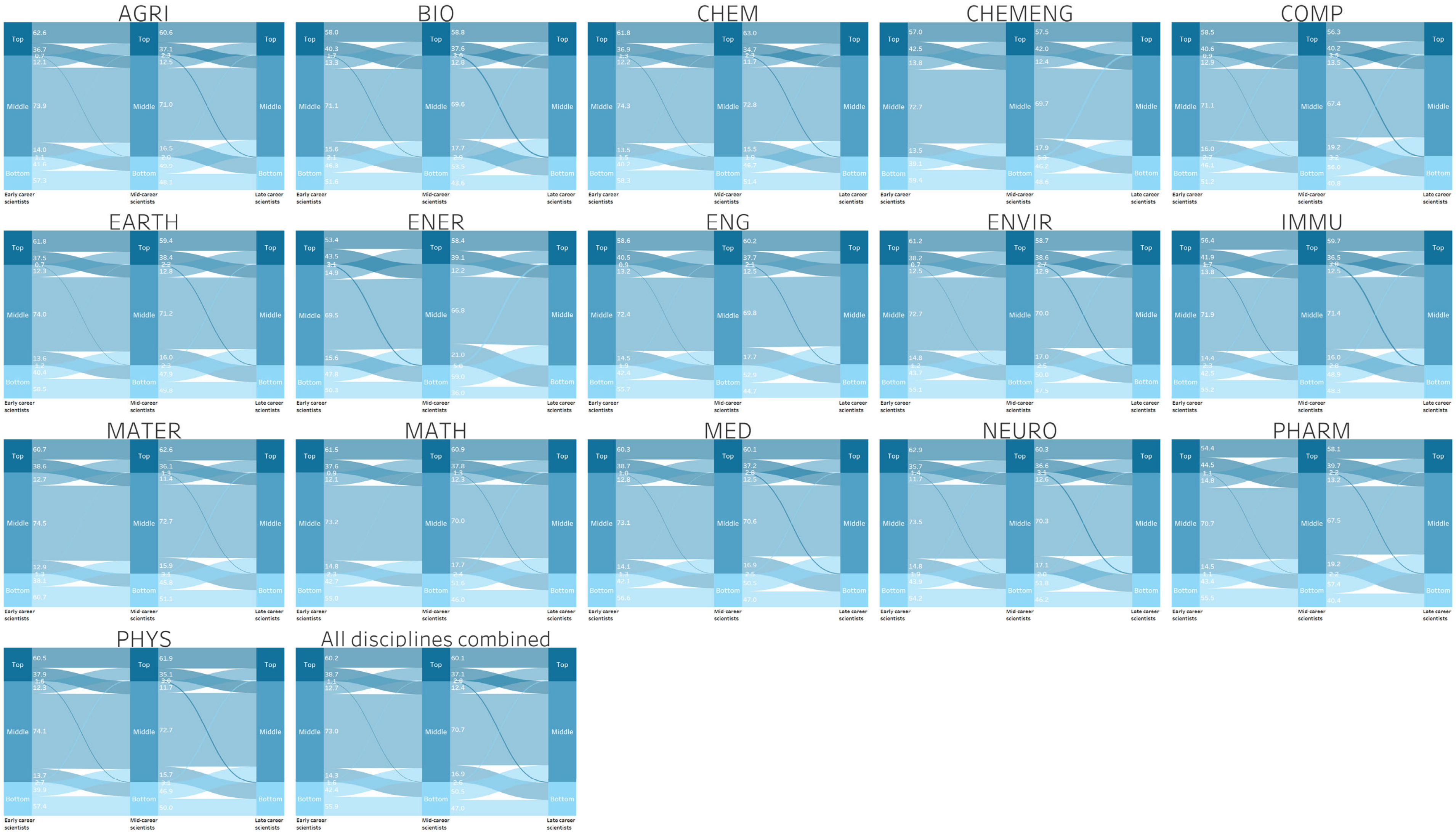


Figure 3. Sankey diagrams of retrospectively constructed mobility between productivity classes in the three career stages. **Productivity 1 (prestige-normalized, full counting)**. Sixteen STEMM disciplines, current nonoccasional OECD late-career scientists only (N=324,643)

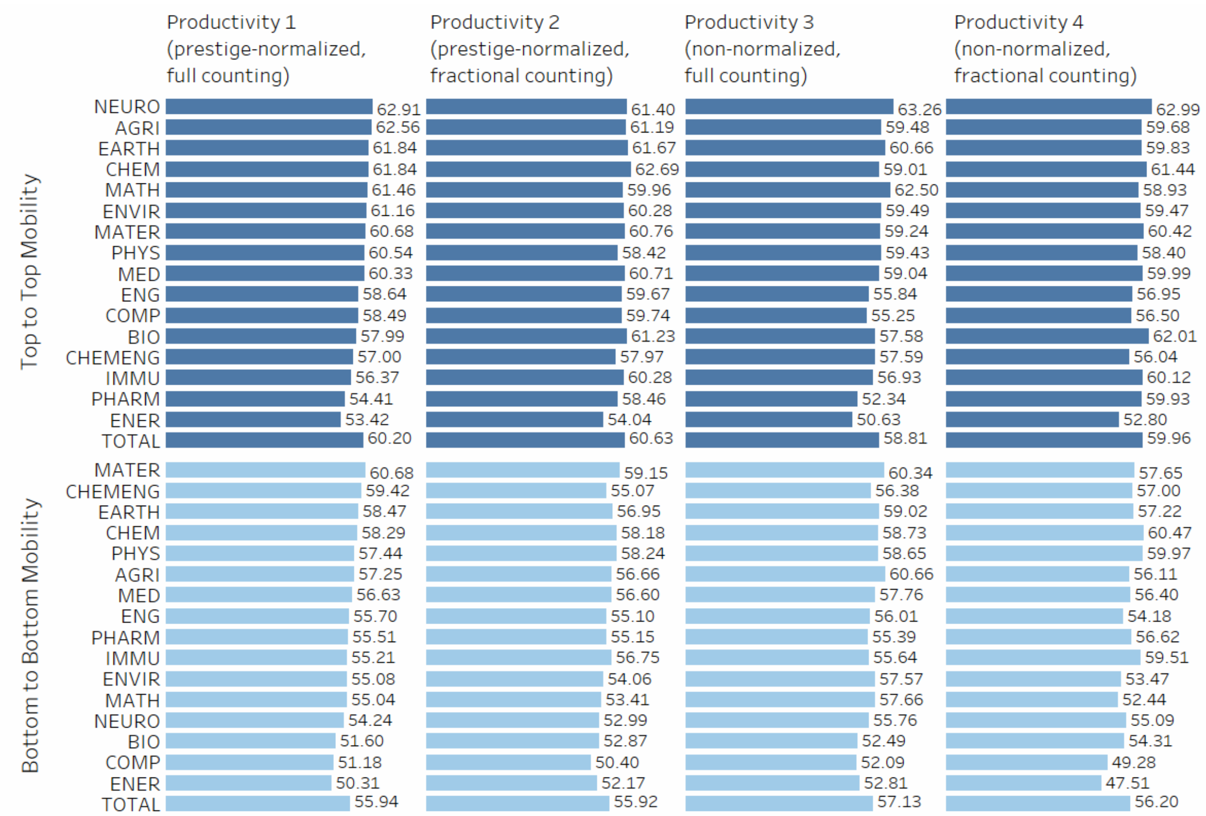


Figure 4. Early career to mid-career stage: horizontal mobility patterns. Top-to-top and bottom-to-bottom mobility by discipline and productivity type, current nonoccasional OECD late-career scientists only (N=324,643). Cross-productivity differences for disciplines are generally small, and cross-disciplinary differences within productivity types are substantial. For instance, top-to-top mobility in Productivity 1 is in the range of 53.42% (ENER) to 62.91% (NEURO) and in Productivity 4, it is in the range of 52.80% to 62.99% for the same disciplines. Bottom-to-bottom mobility in Productivity 1 is in the range of 50.31% (ENER) to 60.68% (MATER), and in Productivity 4, it is in the range of 47.51% (ENER) to 60.47% (CHEM).

Across all disciplines and productivity types, there is a higher proportion of scientists who move from the bottom to top classes compared with those who move from the top to bottom classes in the case of the first transition; there is also a higher proportion of scientists who move from top classes to bottom classes in the second transition. Because of space limitations, we do not present the results for a two-stage mobility: from an early career directly to late-career stage (see a section in Electronic Supplementary Material).

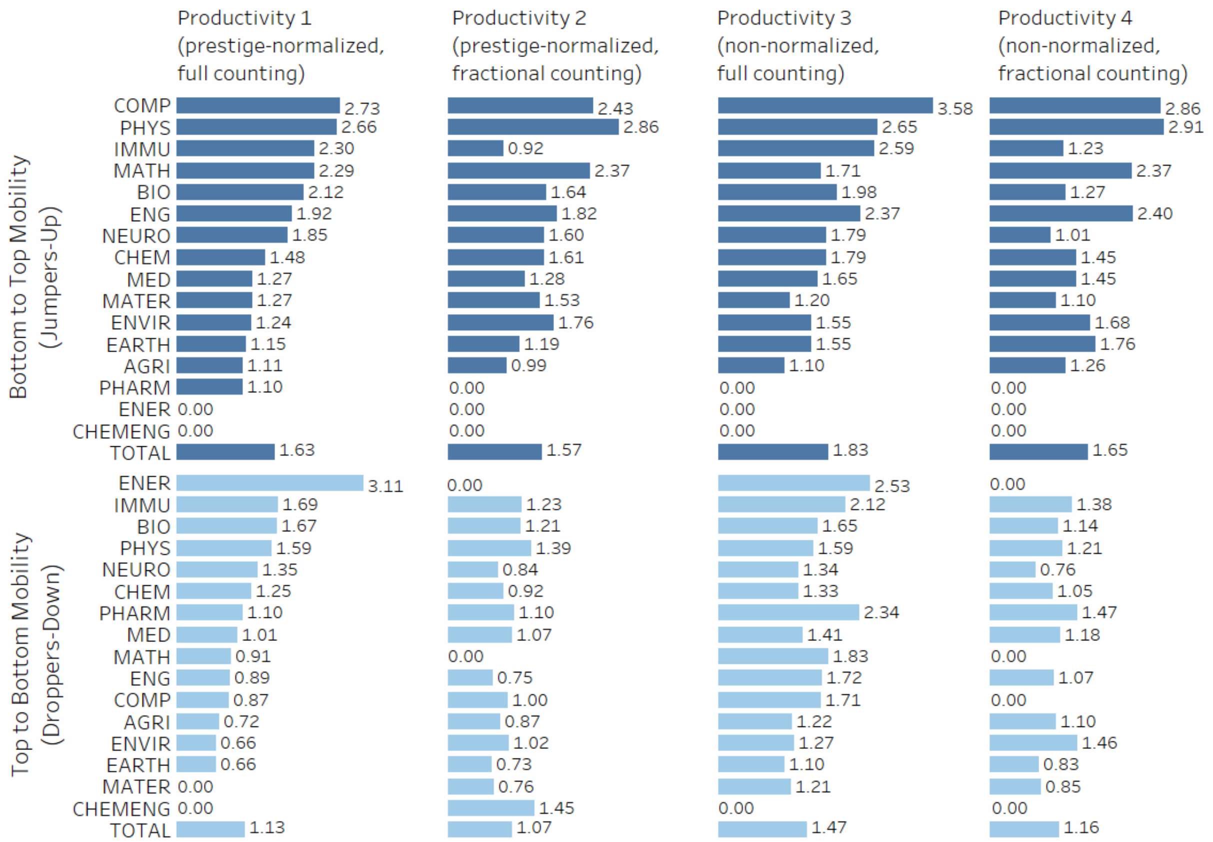


Figure 5. Early career to mid-career stage: extreme vertical mobility patterns. There are two disciplines with no Jumpers-Up (ENER and CHEMENG) and no Droppers-Down (MATER and CHEMENG).

Model Approach: Logistic Regression

In this subsection, we introduce a multidimensional approach and analyze the odds ratio estimates of membership in the classes of 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 20% and the bottom 20%, separately for each discipline, N=324,643).

We use a single demographic variable (gender, binary: male or female) and five variables we have computed using micro-level data on individual scientists. Two variables are related to individual publication quality (citations within fields, received within the first four years after publication; and individual publishing patterns computed with Scopus journals ranks); two other variables are related to individual collaboration patterns (average team size and international collaboration rate); and one variable is related to publication productivity in earlier career stages (prior membership in top and bottom productivity classes).

Publication quality variables and collaboration pattern variables are computed from a lifetime perspective of individual scientists: all journal articles and articles in conference proceedings published throughout one's lifetime. Also all publications (lifetime) and all cited references were used to compute a single discipline to which every scientists was ascribed (see Dataflow in Figure 6, the Methods section). In contrast, membership in productivity classes have been computed for specific periods of careers (early, mid- and late-career periods). Additionally, we also use an institutional variable (TOP 200 institution globally). The variables and their short descriptions are presented in Table 7 in the Methods section.

Extensive previous research on individual productivity has suggested that the most important predictors of high productivity at the individual level are international collaboration[45], collaboration teams[46], gender[4,23], career stage[47,48] and publication productivity earlier in academic careers[49], which is consistent with the Matthew effect in science (the rich get disproportionately richer, and the poor get poorer[29]). As a result—and specifically in the context of our two-dimensional results visualized through Sankey diagrams—we have also added the membership in top and bottom productivity classes.

Logistic Regression, Medicine: Membership in Top and Bottom Productivity Classes

For Medicine (MED), the largest discipline in our sample, regression models have been run for each productivity type separately. The single most important predictor of membership in top (or bottom) productivity classes is membership in top (or bottom) productivity classes at earlier stages of scientific careers (Table 2 and Supplementary Table 10). The influence of this predictor on membership in top productivity classes is powerful: for mid-career scientists, the probability of success increases 11–12 times, depending on productivity type (e.g., in the Model 1, Exp(B)=11.136, with 95% C.I.: 11.109–11.163, and in the Model 4, Exp(B)=12.474, 95% C.I.: 12.447–12.501). For late-career scientists, the models show similarly powerful influence of the class membership predictors: between nine and eleven times. Cross-productivity differences in the power of this predictor are larger in the case of late-career scientists.

The influence of gender is higher for mid-career scientists than for late-career scientists. Being a male scientist increases the odds of success for the former by between 34.3% (Model 1) and 43.5% (Model 4); for the latter, it does so by between 8.4% (Model 2) and 14% (Model 1) only. This may mean that the role of gender in reaching top productivity classes substantially decreases over the course of scientific careers: its influence is more powerful in early years (perhaps when family obligations for women are generally higher) and less powerful in later years.

The influence of the two predictors related to publication quality is consistently higher for mid- than late-career scientists, both the averaged field-weighted four-year citation impact (variable: FWCI 4y) and the average journal percentile rank (variable: AJPR). Although the impact on the global scholarly community viewed through a proxy of individual FWCI 4y is largely irrelevant for probability of success, the impact of (lifetime) individual publishing patterns emerges as an important predictor of success. In Medicine, publishing on average in high-impact rather than low-impact journals matters. We examined one-unit increases in the individual AJPR (Exp(B)=1.012–1.014, depending on the model), but it is more telling to use a 50-unit increase to assess the role of individual publishing patterns.

Table 2. Medicine (MED). Logistic regression statistics: odds ratio estimates of membership in the class of top productive *mid-career* (upper panel) and *late-career* (bottom panel) scientists in Medicine (the top 20%), current nonoccasional OECD late-career scientists in Medicine only (N=137,748).
Odds ratio estimates of membership in the class of bottom productive scientists are shown in Supplementary Table 10.

Model	**Model 1: Productivity 1 - Prestige-normalized full counting**				**Model 2: Productivity 2 - Non-normalized full counting**				**Model 3: Productivity 3 - Prestige-normalized fractional counting**				**Model 4: Productivity 4 - Non-normalized fractional counting**			
	Exp (B)	95% C.I. for Exp(B)		Sig.	Exp(B)	95% C.I. for Exp(B)		Sig.	Exp(B)	95% C.I. for Exp(B)		Sig.	Exp(B)	95% C.I. for Exp(B)		Sig.
		Lower	Upper			Lower	Upper			Lower	Upper			Lower	Upper	
Mid-career scientists																
	$R^2 = 0.231$				$R^2 = 0.232$				$R^2 = 0.223$				$R^2 = 0.223$			
Male	1.343	1.313	1.372	0	1.354	1.325	1.384	0	1.395	1.366	1.425	0	1.435	1.406	1.465	0
FWCI 4y	1.015	1.012	1.019	0	1.02	1.016	1.023	0	1.012	1.008	1.015	0	1.016	1.012	1.019	0
International Collab. Rate	1.007	1.007	1.008	0	1.006	1.005	1.007	0	1.007	1.006	1.007	0	1.005	1.004	1.005	0
AJPR	1.014	1.013	1.014	0	1.014	1.013	1.015	0	1.014	1.013	1.015	0	1.012	1.012	1.013	0
Median Team Size	1.03	1.023	1.036	0	0.962	0.956	0.969	0	1.036	1.029	1.042	0	0.969	0.962	0.975	0
Early Career Top Class	11.136	11.109	11.163	0	11.666	11.639	11.694	0	11.657	11.63	11.683	0	12.474	12.447	12.501	0
Intercept	0.03	-0.032	0.091	0	0.041	-0.016	0.099	0	0.027	-0.032	0.087	0	0.044	-0.011	0.1	0
Late-career scientists																
	$R^2 = 0.224$				$R^2 = 0.201$				$R^2 = 0.215$				$R^2 = 0.18$			
Male	1.14	1.111	1.168	0	1.084	1.056	1.113	0	1.156	1.127	1.185	0	1.089	1.061	1.117	0
FWCI 4y	1.009	1.007	1.012	0	1.006	1.004	1.009	0	1.007	1.005	1.01	0	1.003	1	1.005	0
International Collab. Rate	1.003	1.003	1.004	0	1.005	1.004	1.006	0	1.003	1.002	1.003	0	1.004	1.004	1.005	0
AJPR	1.008	1.007	1.009	0	1.008	1.007	1.009	0	1.009	1.008	1.01	0	1.008	1.007	1.009	0
Median Team Size	1.077	1.07	1.083	0	1.026	1.02	1.032	0	1.079	1.073	1.085	0	1.01	1.004	1.016	0
TOP200	1.357	1.324	1.39	0	1.341	1.309	1.373	0	1.321	1.288	1.354	0	1.283	1.252	1.315	0
Mid-Career Top Class	11.259	11.232	11.286	0	10.562	10.534	10.589	0	11.061	11.035	11.088	0	9.734	9.707	9.76	0
Intercept	0.032	-0.039	0.103	0.049	0.047	-0.019	0.113	0.033	0.029	-0.041	0.099	0.046	0.056	-0.006	0.117	0.023

Note: Sig 0 means p <= 0.001, AJPR is the Average Journal Percentile Rank

Taking publishing on average in journals located in the 90th percentile of Scopus journal ranks (generally more prestigious journals) rather than in those in the 40th percentile (generally less prestigious journals) as an example: the individual average lifetime journal percentile rank is higher by 50 units in a 0–99 continuum of journal ranks in Scopus. For mid-career scientists, the probability of success increases by between 60% (Model 4) and 70% (Models 1, 2, and 3). Journal prestige emerges as a powerful predictor of membership in top productivity classes, especially for mid-career scientists. The TOP200 affiliation increases the odds for late-career scientists by about one-third (all other things being equal).

Second, we have also run four logistic regression models showing odds ratio estimates of membership in bottom productivity classes (Supplementary Table 10), with generally a mirror image of odds ratio estimates of membership in top productivity classes. Being male decreases the odds of success for all productivity types and for both career stages. In addition, the two publication quality predictors decrease the odds of success, as does the median team size. Finally, a lower lifetime international collaboration rate increases the probability of success. There is a single most influential predictor across all productivity types: membership in bottom productivity classes at earlier career stages, which increases the probability of success by eight to nine times for mid-career and by four to five times for late-career scientists. Belonging to bottom productivity classes is more consequential in mid-career than in late-career period which confirms our findings about the share of Jumpers-Up among bottom-class scientists for mid-career (Figure 5) and late-career scientists in Medicine (Supplementary Figure 3).

Logistic Regression, 16 STEMM Disciplines: Membership in Top and Bottom Productivity Classes

First, we analyze top productivity classes (Table 3). In the majority of the 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 8.333 (PHYS) and 14.008 (NEURO) for the first career stage, and 9.410 (CHEMENG) and 13.402 (CHEM) for the second career stage (in all cases: all other things being equal). The direction of the impact of the gender variable is consistent across disciplines; however, the impact is highly differentiated and is much stronger for mid- than late-career scientists. Being male increases the probability of success in 11 disciplines for mid-career scientists by as little as 7% in PHYS and by as much as 105% in IMMU. For late-career scientists, in contrast, the impact of being male is much lower on average (from 4.4% in EARTH to 35.8% in NEURO, with the exception of ENER, where it reaches 99%).

Table 3. Overview of logistic regression models by discipline: odds ratio estimates of membership in the class of top productive *mid-career* (upper panel) and *late-career* (bottom panel) scientists (the upper 20%, separately for each discipline). Productivity 1 (prestige-normalized, full counting), current nonoccasional OECD late-career scientists only (N=324,643)
Odds ratio estimates of membership in the class of bottom productive scientists are shown in Supplementary Table 11.

	AGRI	BIO	CHEM	CHEMENG	COMP	EARTH	ENER	ENG	ENVIR	IMMU	MATER	MATH	MED	NEURO	PHARM	PHYS
Top productive mid-career scientists																
R2	0.265	0.205	0.255	0.214	0.209	0.262	0.207	0.217	0.248	0.192	0.245	0.239	0.231	0.269	0.176	0.261
Male	1.266	1.732	1.387			1.228			1.313	2.050	1.201	1.351	1.343	1.743		1.070
FWCI 4y	1.360	1.131	1.333	1.137	1.031	1.257	1.315		1.209	1.083	1.346	1.162	1.015	1.355	1.025	1.064
International Collab. Rate	1.007	1.004	1.011	1.011	1.003	1.011	1.003	1.008	1.009	1.006	1.007	1.001	1.007	1.004	1.011	1.009
AJPR	1.012	1.001	1.005	1.020	1.013	1.007	1.020	1.011	1.010	1.002	1.005	1.006	1.014	1.000	1.016	1.012
Median Team Size	1.025	0.974	0.939	0.912	1.064	1.040	0.881		1.025	0.997	1.057	1.332	1.030	1.086	0.924	1.117
Early Career Top Class	11.952	10.628	12.531		9.905	11.321			11.159	8.967	10.836	12.485	11.136	14.008		8.833
Intercept	0.023	0.063	0.039			0.023				0.05	0.033		0.03	0.026		
Top productive late-career scientists																
R2	0.231	0.209	0.260	0.202	0.180	0.217	0.227	0.236	0.212	0.229	0.258	0.233	0.224	0.225	0.213	0.267
Male	1.150	1.241	1.148	1.264		1.044	1.990			1.161			1.140	1.358		
FWCI 4y	1.215	1.046	1.268	1.247	1.015	1.099	1.066	1.058		1.270	1.354	1.116	1.009	1.196		1.043
International Collab. Rate	1.005	1.006	1.008	1.000	1.004	1.007	1.011	1.010	1.008	1.011	1.007	1.001	1.003	1.002	1.006	1.014
AJPR	1.009	0.999	1.002	1.004	1.005	1.003	1.012	1.011	1.007	0.992	1.004	1.014	1.008	1.004	1.011	0.996
Median Team Size	0.991	1.039	0.974	1.033		1.003	0.916	0.908	0.955	1.043	0.965	1.211	1.077	1.045	0.976	1.051
Top200	1.360	1.287	1.472	1.490		1.521	1.904			1.403			1.357	1.197		
Mid-Career Top Class	11.015	11.394	13.402	9.410		10.622	9.642			11.372			11.259	11.736		
Intercept	0.034	0.058											0.032			

Only statistically significant results shown in the table. AJPR is the Average Journal Percentile Rank.

The two variables of publication quality emerge as powerful predictors of success, and they are statistically significant for almost all disciplines. The role of FWCI 4y is more important in some disciplines (e.g., AGRI and CHEM) and less so in others (e.g., COMP and PHYS). Average journal percentile rank (AJPR) is a powerful predictor that increases the odds of success for all disciplines and for both career stages. A one-unit increase in AJPR (i.e., by one percentile rank) increases the odds in the range from 0.1% (BIO) to 2.0% (CHEMENG and ENER) for mid-career scientists and from 0.2% (CHEM) to 1.4% (MATH) for late-career scientists.

In other words, on average, publishing in high-impact rather than low-impact journals (e.g., the 90th vs. the 40th percentile of Scopus journals: AJPR higher by 50 units) for mid-career scientists increases the odds by between 5% (BIO) and 100% (CHEMENG and ENER). However, for late-career scientists, the impact of journal publishing patterns is generally much smaller. International collaboration matters for all disciplines and both career stages. For scientists in the first career stage (in CHEM, CHEMENG, EARTH, and PHARM), a 50-unit increase in the international collaboration rate results in an increase of 55% on average in the probability of success. There is no substantial difference in the role of this predictor for late-career scientists. In contrast, the impact of team size is ambiguous. Finally, the TOP200 affiliation greatly increases the probability of success in the majority of disciplines by between 20% and 90%.

Second, we also analyzed bottom productivity classes (Supplementary Table 11). Contrary to our findings for Medicine, the results are mirror images of those for top productivity classes but only to some extent. Specifically, although the role of membership in 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 classes are statistically significant only for seven (first career period) and five (second career period) disciplines. Membership in bottom productivity classes earlier (publishing years 5–14) increases the probability of membership in this class later (publishing years 15–24) by 7–10 times in the case of the first career period and by 3–5 times in the case of the second period.

Being a female scientist, as expected, substantially increases the odds of success; 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 female in BIO increases the odds of success by one-third on average, but being female in ENG decreases the odds of success by one-third on average (32.0% and 34.9%, respectively).

Discussion

Following the results of our longitudinal study based on micro-level data on 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 and bottom) is already largely settled and that the early global distribution persists over time: for years and decades. Exceptions are very rare: bottom performers almost never become top performers, and top performers almost never become bottom performers.

Some scientists stay on in academic science and keep publishing, while others stop publishing[50,51] about one-third disappears from academic publishing within five years and about a half within a decade[52]. Within the first 15 years of academic publishing, those who stay on in science are already distributed among the classes of global top, middle, and bottom performers within their disciplines (according to the 20/60/20 principle).

What is stunning is the persistence of membership in global top and bottom classes from a life cycle perspective. Later on in their careers, the majority of top performers keep being top performers, and the majority of bottom performers keep being bottom performers. For them, the probability of staying in top and bottom classes—horizontal mobility—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.

Individual research productivity emerges from our regression analyses as highly path dependent: for all the examined disciplines, 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.

For instance, in Medicine, the largest STEMM discipline in our sample, the regression results differ only slightly by productivity type (Models 1 through 4, Table 2). Both for current late-career scientists and for late-career scientists when they were mid-career scientists, the single most important predictor of membership in top productivity classes is prior membership in top productivity classes. The second most powerful predictor of success in Medicine is gender: however, being male is much more influential for mid-career scientists, which increases the odds by 34.3%, than for late-career scientists, which increases the odds by merely 14.0%.

The third predictor of success is average journal percentile rank, which reflects lifetime publishing patterns: consistent publishing in high-impact academic journals substantially increases the odds compared with consistent publishing in low-impact academic journals[53,54,55,56]. The fourth predictor of success is working in the 200 most research-intensive institutions. Finally, the median team size and international collaboration rate (all other things being equal) are much less important than expected based on the literature about collaboration in science[46].

Our large-scale longitudinal analyses show that about 60% of global top productivity scientists continue their careers as top productivity scientists, and about a half of bottom productivity scientists continue their careers as bottom productivity scientists. The global science system is highly immobile: Jumpers-Up and Droppers-Down are extremely rare scholarly species (e.g., our micro-level data show that only 1.6% scientists move from early career bottom class to mid-career top class, and only 1.1% scientists move from early career top class to mid-career bottom class; Productivity 1, Table 1).

A general picture of mobility between productivity classes over the course of entire scientific careers based on discipline-aggregated data (Figure 1) hides much more nuanced pictures for different disciplines (Figure 3). Some disciplines are much more competitive from the very beginning of careers, with radical upward mobility being extremely difficult (or impossible, as in MATER and CHEMENG). There are also other disciplines that are much less competitive from the very beginning, in which the presence of Jumpers-Up is much higher. Our focus is not on globally evolving productivity over time[57] or evolving productivity from a generational perspective (e.g., the old in science being more productive than the young or the other way round[34,58]) but on interclass mobility of individuals over entire scientific careers.

Why does prior class memberships (top, bottom), to a large extent, determine later class memberships (top, bottom)? There are two explanations. First, previous research has shown that distribution of productivity among scientists is highly skewed[42,59] and that the minority of scientists are responsible for the vast majority of publications[60]. Esteem comes from peers in science, and the reward system in science is based on publications. And academic promotions and tenure prospects, salary levels, free time for research, and access to research grants, more or less directly, are all related to publication productivity[16,61].

Second, higher productivity generally leads to new research funding, as the credibility cycle in academic careers shows[62]. 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 early career stages: once funded on the basis of 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. In terms of shifting productivity classes from a life cycle perspective, scientists who are less successful early on in their careers (productivity, luck) will find it difficult, if not impossible, to prove that they are as good as their more successful, more productive, more lucky and possibly better funded colleagues.

Our research reconfirms 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 examine using our dataset—the probability of past global top performers to become global top performers in the future is very high, and their probability to become global bottom performers is marginal. At the same time, the chances of global bottom performers to reach productivity levels achieved by their top-performing colleagues in the very same career stages and within their disciplines are marginal.

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 using small-scale interviews and surveys[12,13,27,28]: 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).

The present research suggests the examination of scientific careers from a comprehensive perspective that requires some trade-offs and limitations (see Methods), if steps in new directions are to be taken. A large-scale, global, longitudinal, publication- and citation-based approach to individual research productivity, which is a cornerstone of scientific careers, uses various proxies and relies on different trade-offs but hopefully shows new interesting patterns that are, so far, largely underexplored.

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 2022 (with at least 25 years of publishing experience) located in 16 STEMM disciplines and coming from 38 OECD countries (N=324,643 scientists with N=16,345,891 research articles, Figure 6). For our calculations, we utilized the Scopus database dated October 21, 2022.

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. The run was 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, 1–6 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 6 hours, and this operation was initiated on January 16, 2023. We obtained the results in CSV format. Subsequently, we employed the SankeyMATIC tool for interactive data visualization in the form of Sankey diagrams.

The academic lives of all late-career scientists from 38 OECD countries research-active in 2022 (N=324,643, with at least 25 years of publishing experience) 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 (2018–2022) and looked back into their past publishing behavior.

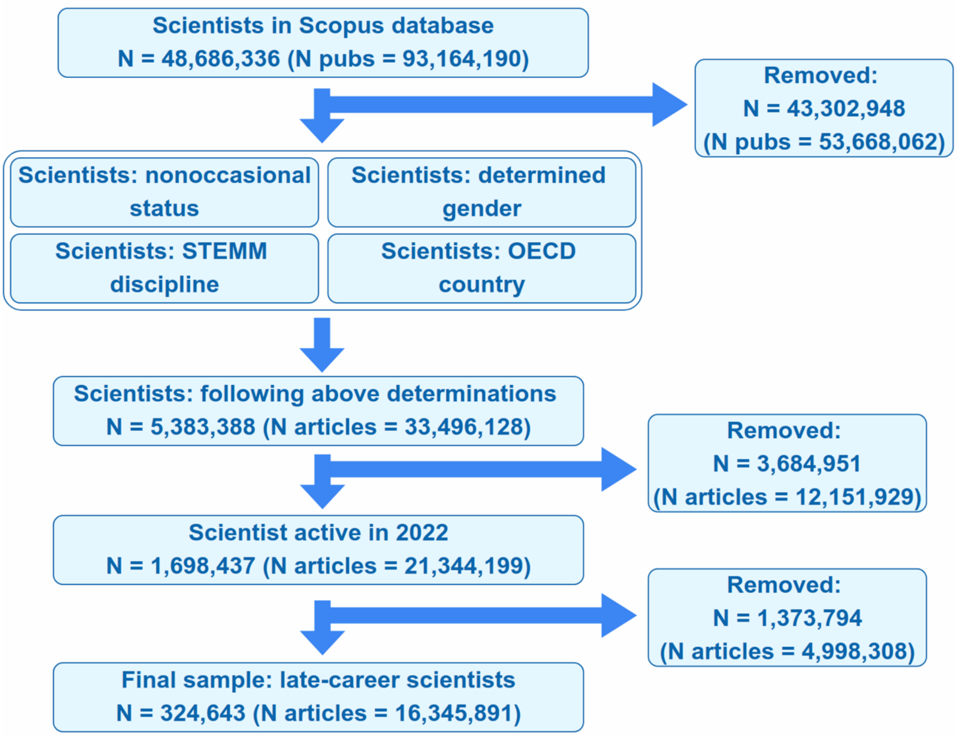


Figure 6. Flowchart and major steps in data preprocessing: from all scientists in the Scopus database to late-career scientists in our sample.

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. We examined the average and median productivity of three age groups among late-career scientists to ensure that their productivity for 2018–2022 would not be skewed by academic age within disciplines (Supplementary Table 8 and Supplementary Table 9). 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 STEMM disciplines) and then, retrospectively, by past productivity classes in the two earlier career periods.

Early career scientists may retain or change their classes (top, middle, bottom) while being mid-career scientists as mid-career scientists may do while being late-career scientists. In the present study, we tracked 324,643 scientists for 25–50 years and compared their productivity with productivity of their peers (the same academic career stage, 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 2022. Within portfolios, all metadata and their constructs were linked to the three career periods (e.g., annual productivity, AJPR), individual publications (e.g., field-weighted four-year citation impact), or to the whole lifetime careers of scientists (e.g., gender, discipline, international collaboration rate, and median team size) (see Variables in Table 7).

Our approach to individual research productivity is longitudinal[1,2,3] and classificatory (or class based)[63,64]. First, we tracked the productivity of late-career scientists as individuals ever since they have become early career scientists, that is, five years after their first globally indexed publication. Second, we did not compare productivity changing 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 productivity classes while scientists get older and move up the professional ladder. Scientists can always be allocated to top and bottom classes so that both terms are used not to judge the level of productivity but rather to classify it. We aimed to verify whether the mobility patterns found depend on major productivity types.

We used four counting methods in examining productivity: two prestige-normalized and two prestige non-normalized: Productivity 1 (prestige-normalized, full counting), Productivity 2 (prestige-normalized, fractional counting), Productivity 3 (non-normalized, full counting), and Productivity 4 (non-normalized, fractional counting). Productivity 1 and 2 refer to quantity and quality of globally indexed publications at the level of individuals; Productivity 3 and 4 refer to quantity only. 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, with different peer review procedures and acceptance rates. Prestige normalization is determined by the number of citations received by the journal (41,474 journals in 2023) 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[39].

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 (20/60/20, top / middle / bottom) were traced over their lifetime—as early, mid- and late-career scientists.

In the present research, we used a global bibliometric dataset to define scientists' individual attributes. The determination of some attributes have 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 three research articles), and determining academic age (using the distance in years between the first publication, of any type, and 2022)[65]. Four other individual attributes were used in individual publication and citation portfolios (their construction is described in Table 7): international collaboration rate (lifetime), field-weighted four-year citation impact (FWCI 4y), median team size (lifetime), and average journal percentile rank (AJPR). The distribution of the sample by academic age (i.e., publishing experience) is shown in Figure 7, with further details in Supplementary Tables 1 through 3.

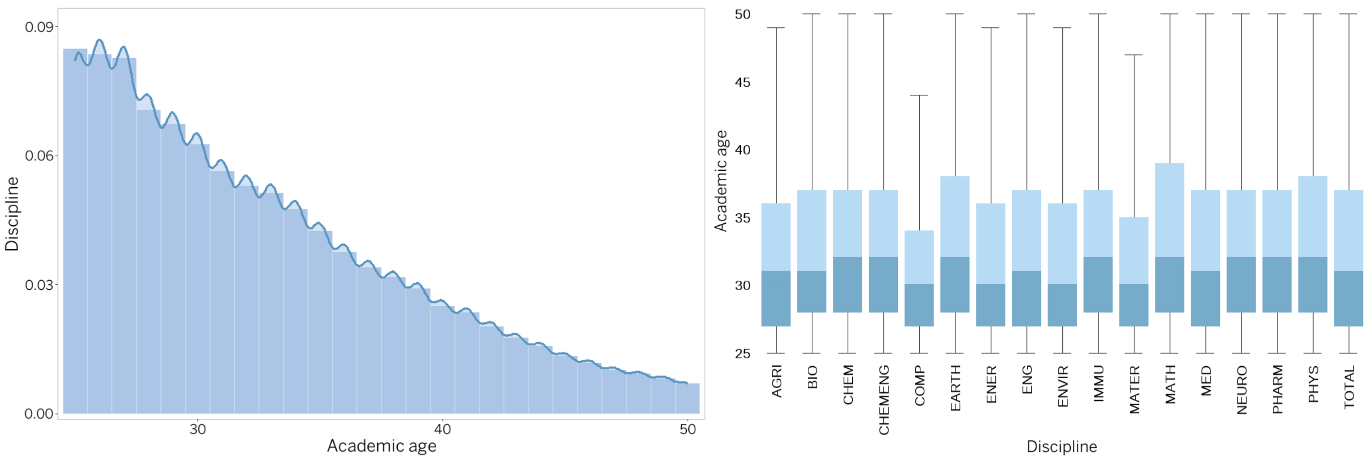


Figure 7. Distribution of academic age: kernel density plot, nonoccasional late-career scientists in 16 STEMM academic disciplines combined (left panel), and their median academic age (age from the first publication), by discipline (right panel) (N=324,643). The kernel density plot shows how the number of late-career scientists decreases with every year of academic age. The median academic age of late-career scientists in most disciplines falls between 30 and 32 years. The disciplines with the highest median age (32 years) are PHYS, PHARM, MATH, and EARTH; the disciplines with the lowest median age (30 years) include MATER, ENVIR, ENG, ENER, and COMP. The overall median academic age of our sample across all disciplines is 31 years.

Trade-offs and Limitations

The present research shows trade-offs between what is theoretically desirable and what is practically possible in studying the global patterns of research productivity, here based on currently available global datasets. The trade-offs and limitations are related to data and methodology.

First, there seem to be no other longitudinal datasets globally available (38 countries) than Scopus (or Web of Science) that can be meaningfully used to examine changing field-normalized productivity over scientists' lifetimes. These datasets are available for selected countries only and with selected parameters only (e.g., CRISTIN for Norway, Academic Analytics for the USA, RADON for Poland[39,59,66]); as a result, no longitudinal and discipline-based (as opposed to selected country-based) approaches are possible without access to global bibliometric datasets. However, bibliometric datasets come with their own limitations, as discussed for at least two decades[67].

Table 7: Variables used in regression analysis.

No.	Variable	Description
1.	Gender	Gender (binary: female/male) provided by 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 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.	Average journal percentile rank (AJPR)	Average of the journals' percentiles assigned to each publication in author's publishing portfolio: computed separately for early career, mid-career, and late-career periods. The percentile value has been taken from the 2022 Journal CiteScore metric for discipline with the highest percentile value.
6.	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.
7.	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 between 2019 and 2022. 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).
8. 9.	Early career / mid-career top class	Membership in the top 20% of scientists among early career /mid- career scientists in terms of research productivity, separately within 16 STEMM disciplines, separately for each of the four productivity types.
10. 11.	Early career / mid-career bottom class	Membership in the bottom 20% of scientists among early career / mid-career scientists in terms of research productivity (disciplines and productivity: as above).

Second, the character of our dataset determines a reductive understanding of productivity in which only Scopus-indexed publications are counted, leaving aside nonindexed publications in local languages. However, our focus on STEMM disciplines, generally using English for global scholarly communication, makes this research less biased. Additionally, all nonpublishing academic activities do not count toward productivity.

Third, the longitudinal nature of our study makes only survivors in science our focus: we leave aside all scientists who are not research active for at least 25 years. As a result, being aware of high attrition rates in STEMM disciplines[52,68] we are aware of a "success bias" in our research: the various mobility types between productivity classes analyzed do not actually refer to beginning, early and mid-career scientists (active in publishing for less than 5, less than 15, and less than 25 years, respectively). Our study takes a long-term view in which, necessarily, because of high attrition in science, the majority of currently active scientists are not represented.

Finally, our methodology has clear limitations that are especially evident if we compare the present study to single-nation studies of productivity. In single-nation studies a wealth of national data are used (e.g., individual career histories with promotion dates, doctoral and postdoctoral dissertation details, research funding details, national classifications of disciplines, national rankings of institutions, etc.), not unavailable at a global level. Additionally, our global study examines scientists from systems with different research funding levels and average individual productivity. Out of necessity, our analyses have to rely on several proxies: on a commercial journal classification (Scopus All Science Journal Classification system, ASJC) and journal disciplinary classifications rather than on a wealth of national disciplinary classifications[69]; data on individual Scopus IDs rather than data on "real scientists" with their national registry-based IDs; on inferred rather than self-declared, administrative gender, based on gender-determining algorithms (probability threshold: 0.85); on a single country affiliation and single institutional affiliation rather than on a plethora of changing country and institutional affiliations, at least for some scientists.

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

Acknowledgments

We are grateful to the hosts and audiences of invited seminars at the University of Oxford (CGHE, Center for Global Higher Education, June 2022), Stanford University (METRICS, Meta-Research Innovation Center, June 2022), DZHW Berlin (German Center for Higher Education Research and Science Studies, June 2023), and Leiden University (CWTS, Centre for Science and Technology Studies, June 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, Senior Data Scientist in the Lab.

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:

1. Ruspini, E. (1999). Longitudinal research and the analysis of social change. Quality and Quantity 33(3), 219–227.
2. Menard, S. (2002). Longitudinal research. Thousand Oaks: Sage.
3. Rowland, D. T. (2014). Demographic methods and concepts. Oxford University Press.
4. Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C.R. (2013). Global gender disparities in science. Nature, 504, 211–213.
5. 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.
6. Nielsen, M. W., & Andersen, J. P. (2021). Global citation inequality is on the rise. Proceedings of the National Academy of Sciences, 118(7), e2012208118.
7. Salganik, M. J. (2018). Bit by bit. Social research in a digital age. Princeton University Press.
8. Liu, L., Jones, B.F., Uzzi, B., et al. (2023). Data, measurement and empirical methods in the science of science. Nature Human Behaviour, 7, 1046–1058.
9. Clauset, A., Larremore, D. B., & Sinatra, R. (2017). Data-driven predictions in the science of science. Science, 355, 477–480.
10. Zeng, A. et al. (2017). The science of science: from the perspective of complex systems. Phys. Rep., 714, 1–73
11. Wang, D., & Barabási, A.-L. (2021). The science of science. Cambridge: Cambridge University Press.
12. Leišytė, L., & Dee, J. R. (2012). Understanding academic work in changing institutional environment. Higher Education: Handbook of Theory and Research, 27, 123–206.
13. 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 27 (pp. 207–248). Dordrecht: Springer.

14. 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.
15. Ni, C., Smith, E., Yuan, H., Larivière, V., & Sugimoto, C. R. (2021). The gendered nature of authorship. Sci. Adv., 7, eabe4639.
16. Allison, P. D., Long, J. S., Krauze, T. K. (1982). Cumulative advantage and inequality in science. American Sociological Review, 47(5), 615–625.
17. Fox, M. F. (1983). Publication productivity among scientists: A critical review. Social Studies of Science, 13(2), 285–305.
18. 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). Paris: CNRS.
19. 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.
20. Aguinis, H., & O'Boyle, E. (2014). Star performers in twenty-first century organizations. Personnel Psychology, 67(2), 313–350.
21. 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
22. Li, W., Aste, T., Caccioli, F., & Livan, G. (2019). Early coauthorship with top scientists predicts success in academic careers. Nature Communications 10, 5170.
23. Fox, M. F., & Nikivincze, I. (2021). Being highly prolific in academic science: Characteristics of individuals and their departments. Higher Education, 81, 1237–1255.
24. 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.
25. Allison, P. D., & Stewart, J. A. (1974). Productivity differences among scientists: Evidence for accumulative advantage. American Sociological Review, 39(4), 596–606.
26. 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). Amsterdam: Elsevier.
27. Cole, J. R., & Cole, S. (1973). Social stratification in science. Chicago: The University of Chicago Press.
28. Merton, R. K. (1973). The sociology of science: Theoretical and empirical investigations. University of Chicago Press.
29. 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.
30. 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.
31. Ma, Y., Mukherjee, S., & Uzzi, B. (2020). Mentorship and protégé success in STEM fields. Proceedings of the National Academy of Sciences, 117, 14077–14083.
32. 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.
33. Sugimoto, C., & Larivière, V. (2023). Equity for Women in Science. Dismantling Systemic Barriers to Advancement. Boston: Harvard University Press.
34. 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
35. Wang, Y., Jones, B. F., & Wang, D. (2019). Early career setback and future career impact. Nature Communications, 10, 4331.
36. 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.
37. Ross, M. B. et al. (2022). Women are credited less in science than men. Nature, 608, 135–145.

38. Kwiek, M., & Roszka, W. (2023). Once highly productive, forever highly productive? Full professors' research productivity from a longitudinal perspective. Higher Education. Online first. https://doi.org/10.1007/s10734-023-01022-y
39. Kwiek, M., & Roszka, W. (2023). Moving up the academic ladder, changing productivity? The case of Polish high and low performers. Preprint at SocArXiv: https://osf.io/preprints/socarxiv/tajs4
40. Marginson, S. (2022) What drives global science? The four competing narratives. Studies in Higher Education, 47(8), 1566–1584. DOI: 10.1080/03075079.2021.1942822
41. Kwiek, M. (2023). The globalization of science: The increasing power of individual scientists. In P. Mattei, X. Dumay, E. Mangez & Ja. Behrend (Eds.), The Oxford handbook of education and globalization (online ed.). Oxford Academic. https://doi.org/10.1093/oxfordhb/9780197570685.013.16.
42. 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.
43. Stephan, P. E., & Levin, S. G. (1992). Striking the mother lode in science: The importance of age, place, and time. Oxford: Oxford University Press.
44. Allison, P. D. (1980). Inequality and scientific productivity. Social Studies of Science, 10, 163–179.
45. 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.
46. Wagner, C. S. (2018). The collaborative era in science. Governing the network. Cham: Palgrave Macmillan.
47. 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.
48. Jung, J. (2014). Research productivity by career stage among Korean academics. Tertiary Education and Management, 20(2), 85–105.
49. 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.
50. Preston, A. E. (2004). Leaving science. Occupational exit from scientific careers. New York: Russell Sage Foundation.
51. 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). Amsterdam: Elsevier.
52. Kwiek, M., & Szymula, Ł. (2023). Quantifying attrition in science: A cohort-based, longitudinal study of scientists in 38 OECD countries. Preprint at SocArXiv: https://doi.org/10.31235/osf.io/8kzb7
53. 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.
54. 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.
55. 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.
56. Heckman, J. J., & Moktan, S. (2018). Publishing and promotion in economics. The tyranny of the Top Five. NBER Working Paper 25093.
57. 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.
58. 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.
59. 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.
60. Ruiz-Castillo, J., & Costas, R. (2014). The skewness of scientific productivity. Journal of Informetrics, 8(4), 917–934.
61. Stephan, P. (2012). How economics shapes science. Cambridge, MA: Harvard University Press.

62. Latour B. & Woolgar S. (1986) Laboratory life. The construction of scientific facts. Princeton University Press.
63. 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, 226–230.
64. 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.
65. 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. Online first. https://doi.org/10.1162/qss_e_00276
66. 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.
67. Sugimoto, C., & Larivière, V. (2018). Measuring research: What everyone needs to know. Oxford: Oxford University Press.
68. Spoon, K. et al. (2023). Gender and retention patterns among U.S. faculty. Science Advances, 9, eadi2205. DOI:10.1126/sciadv.adi2205
69. 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

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 OECD late-career scientists by gender, STEMM discipline, and country (N=324,643)

	Female scientists			**Male scientists**			**Total**		
	N	**% row**	**% col**	**N**	**% row**	**% col**	**N**	**% row**	**% col**
AGRI	6,227	25.72	7.44	17,984	74.28	7.46	24,211	100.00	7.46
BIO	15,034	31.49	17.97	32,709	68.51	13.57	47,743	100.00	14.71
CHEMENG	168	16.22	0.20	868	83.78	0.36	1,036	100.00	0.32
CHEM	3,229	21.27	3.86	11,955	78.73	4.96	15,184	100.00	4.68
COMP	1,185	14.75	1.42	6,849	85.25	2.84	8,034	100.00	2.47
EARTH	2,623	17.32	3.14	12,522	82.68	5.20	15,145	100.00	4.67
ENER	87	10.82	0.10	717	89.18	0.30	804	100.00	0.25
ENG	1,222	8.72	1.46	12,788	91.28	5.31	14,010	100.00	4.32
ENVI	1,631	23.87	1.95	5,201	76.13	2.16	6,832	100.00	2.10
IMMU	1,079	33.09	1.29	2,182	66.91	0.91	3,261	100.00	1.00
MATER	1,089	18.46	1.30	4,809	81.54	2.00	5,898	100.00	1.82
MATH	1,128	15.71	1.35	6,051	84.29	2.51	7,179	100.00	2.21
MED	43,258	31.40	51.71	94,490	68.60	39.21	137,748	100.00	42.43
NEURO	1,670	28.10	2.00	4,273	71.90	1.77	5,943	100.00	1.83
PHARM	472	34.71	0.56	888	65.29	0.37	1,360	100.00	0.42
PHYS	3,560	11.77	4.26	26,695	88.23	11.08	30,255	100.00	9.32
TOTAL	83,662	25.77	100	240,981	74.23	100	324,643	100.00	100

Supplementary Table 2. Structure of the sample of all nonoccasional OECD late-career scientists by academic age (publishing experience) and gender (N=324,643)

Academic age	Female scientists	Male scientists	% Female scientists	% Male scientists	Total	Male to Female Rate (MFR)
25	8,692	18,922	31.48	68.52	27,614	2.18
26	8,543	18,644	31.42	68.58	27,187	2.18
27	8,123	18,787	30.19	69.81	26,910	2.31
28	6,700	16,270	29.17	70.83	22,970	2.43
29	6,317	15,605	28.82	71.18	21,922	2.47
30	5,817	14,566	28.54	71.46	20,383	2.50
31	5,107	13,262	27.80	72.20	18,369	2.60
32	4,494	12,753	26.06	73.94	17,247	2.84
33	4,408	12,297	26.39	73.61	16,705	2.79
34	3,749	11,715	24.24	75.76	15,464	3.12
35	3,315	10,520	23.96	76.04	13,835	3.17
36	2,791	9,433	22.83	77.17	12,224	3.38
37	2,510	8,547	22.70	77.30	11,057	3.41
38	2,151	8,180	20.82	79.18	10,331	3.80
39	1,986	7,504	20.93	79.07	9,490	3.78
40	1,645	6,508	20.18	79.82	8,153	3.96
41	1,456	6,235	18.93	81.07	7,691	4.28
42	1,289	5,310	19.53	80.47	6,599	4.12
43	1,033	4,732	17.92	82.08	5,765	4.58
44	856	4,283	16.66	83.34	5,139	5.00
45	677	3,685	15.52	84.48	4,362	5.44
46	607	3,239	15.78	84.22	3,846	5.34
47	439	2,878	13.23	86.77	3,317	6.56
48	389	2,640	12.84	87.16	3,029	6.79
49	325	2,377	12.03	87.97	2,702	7.31
50	243	2,089	10.42	89.58	2,332	8.60

Supplementary Table 3. Structure of the sample of all nonoccasional OECD late-career scientists by gender and country, top 20 countries only (N=324,643).

	Female scientists			Male scientists			Total		
	N	% row	% col	N	% row	% col	N	% row	% col
United States	26,193	26.96	31.31	70,960	73.04	29.45	97,153	100	29.93
Japan	2,897	9.57	3.46	27,370	90.43	11.36	30,267	100	9.32
Italy	10,774	37.03	12.88	18,318	62.97	7.60	29,092	100	8.96
France	6,372	29.41	7.62	15,293	70.59	6.35	21,665	100	6.67
United Kingdom	5,243	24.69	6.27	15,991	75.31	6.64	21,234	100	6.54
Germany	3,417	16.57	4.08	17,201	83.43	7.14	20,618	100	6.35
Spain	4,701	34.40	5.62	8,966	65.60	3.72	13,667	100	4.21
Canada	3,551	28.32	4.24	8,987	71.68	3.73	12,538	100	3.86
Australia	2,896	28.83	3.46	7,149	71.17	2.97	10,045	100	3.09
Netherlands	1,886	23.58	2.25	6,114	76.43	2.54	8,000	100	2.46
Poland	2,014	33.62	2.41	3,977	66.38	1.65	5,991	100	1.85
South Korea	589	12.26	0.70	4,214	87.74	1.75	4,803	100	1.48
Sweden	1,290	27.15	1.54	3,462	72.85	1.44	4,752	100	1.46
Switzerland	698	17.38	0.83	3,317	82.62	1.38	4,015	100	1.24
Belgium	924	25.92	1.10	2,641	74.08	1.10	3,565	100	1.10
Turkey	869	25.08	1.04	2,596	74.92	1.08	3,465	100	1.07
Greece	877	25.45	1.05	2,569	74.55	1.07	3,446	100	1.06
Israel	918	27.02	1.10	2,480	72.98	1.03	3,398	100	1.05
Mexico	925	30.57	1.11	2,101	69.43	0.87	3,026	100	0.93
Denmark	787	26.62	0.94	2,169	73.38	0.90	2,956	100	0.91
Other	5,841	27.88	6.99	15,106	72.12	6.23	20,947	100	6.46
TOTAL	83,662	25.77	100	240,981	74.23	100	324,643	100	100

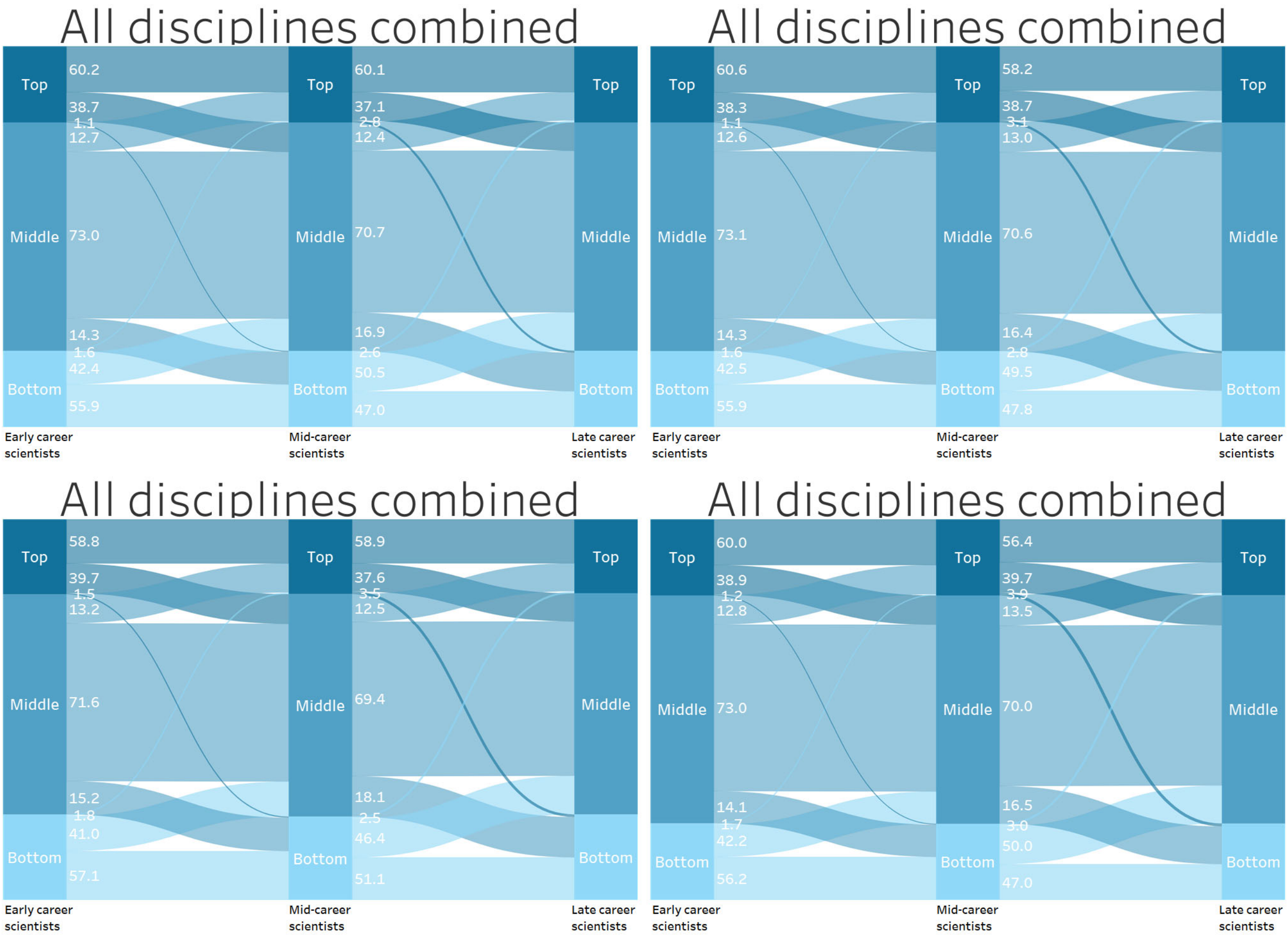
Supplementary Table 4. Productivity 2 (prestige-normalized, fractional counting), three stages. Mobility between retrospectively constructed productivity classes in the three stages of academic careers, current nonoccasional OECD late-career scientists only, all disciplines combined (N=324,643).

Tables for Productivity 3 and Productivity 4 available upon request (space limitations).

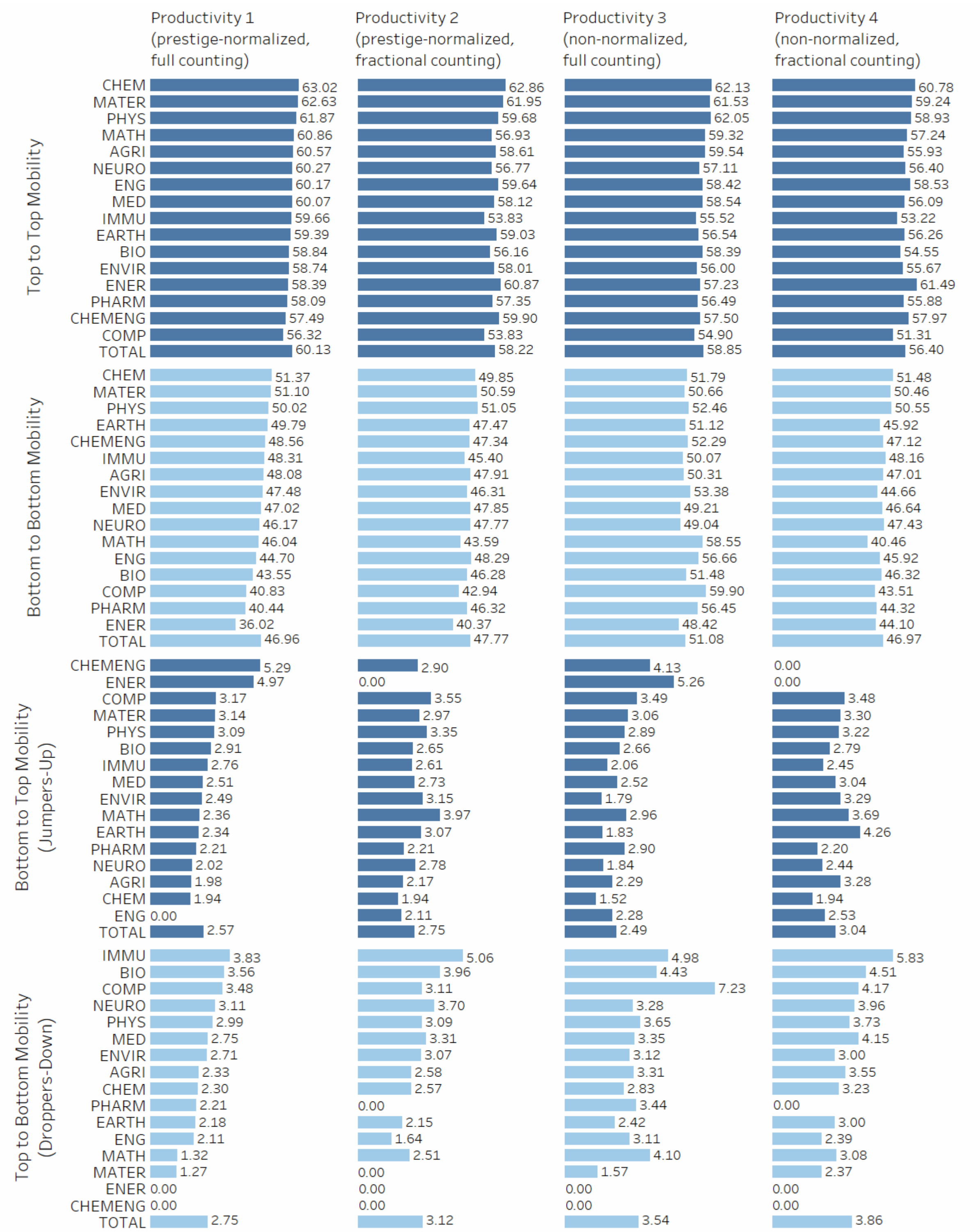
				Female scientists			Male scientists			Total		
Transition from source academic position	Transition from productivity class	Transition to target academic position	Transition to productivity class	Number of scientists in transition	Number of scientists in each productivity class	%	Number of scientists in transition	Number of scientists in each productivity class	%	Number of scientists in transition	Number of scientists in each productivity class	%
Early career	Bottom	Mid-career	Bottom	11,326	20,423	55.5	24,982	44,511	56.1	36,308	64,934	55.9
Early career	Bottom	Mid-career	Middle	8,868	20,423	43.4	18,740	44,511	42.1	27,608	64,934	42.5
Early career	Bottom	Mid-career	Top	229	20,423	1.1	789	44,511	1.8	1,018	64,934	1.6
Early career	Middle	Mid-career	Bottom	8,013	52,353	15.3	19,916	142,425	14.0	27,929	194,778	14.3
Early career	Middle	Mid-career	Middle	39,032	52,353	74.6	103,270	142,425	72.5	142,302	194,778	73.1
Early career	Middle	Mid-career	Top	5,308	52,353	10.1	19,239	142,425	13.5	24,547	194,778	12.6
Early career	Top	Mid-career	Bottom	109	10,886	1.0	586	54,045	1.1	695	64,931	1.1
Early career	Top	Mid-career	Middle	4,540	10,886	41.7	20,331	54,045	37.6	24,871	64,931	38.3
Early career	Top	Mid-career	Top	6,237	10,886	57.3	33,128	54,045	61.3	39,365	64,931	60.6
Mid-career	Bottom	Late career	Bottom	9,292	19,448	47.8	21,723	45,484	47.8	31,015	64,932	47.8
Mid-career	Bottom	Late career	Middle	9,801	19,448	50.4	22,330	45,484	49.1	32,131	64,932	49.5
Mid-career	Bottom	Late career	Top	355	19,448	1.8	1,431	45,484	3.2	1,786	64,932	2.8
Mid-career	Middle	Late career	Bottom	8,300	52,440	15.8	23,604	142,341	16.6	31,904	194,781	16.4
Mid-career	Middle	Late career	Middle	37,804	52,440	72.1	99,732	142,341	70.1	137,536	194,781	70.6
Mid-career	Middle	Late career	Top	6,336	52,440	12.1	19,005	142,341	13.4	25,341	194,781	13.0
Mid-career	Top	Late career	Bottom	291	11,774	2.5	1,736	53,156	3.3	2,027	64,930	3.1
Mid-career	Top	Late career	Middle	4,369	11,774	37.1	20,732	53,156	39.0	25,101	64,930	38.7
Mid-career	Top	Late career	Top	7,114	11,774	60.4	30,688	53,156	57.7	37,802	64,930	58.2
Late career	Bottom			17,883	17,883	100	47,063	47,063	100	64,946	64,946	100
Late career	Middle			51,974	51,974	100	142,794	142,794	100	194,768	194,768	100
Late career	Top			13,805	13,805	100	51,124	51,124	100	64,929	64,929	100

Supplementary Figure 1. Sankey diagrams of retrospectively constructed mobility between productivity classes in the three career stages. **Productivity 2 (prestige-normalized, full counting)**. Sixteen STEMM disciplines, current nonoccasional OECD late-career scientists only (N=324,643)

Figures for Productivity 3 and Productivity 4 available upon request (space limitations).



Supplementary Figure 2. Productivity 1-4, overview, transitions between three career stages. Sankey diagrams of retrospectively constructed mobility between productivity classes in the three career stages. All STEMM disciplines combined, only current nonoccasional OECD late-career scientists. Four productivity types: Productivity 1 (prestige-normalized, full counting) (left upper), Productivity 2 (prestige-normalized, fractional counting) (right upper), Productivity 3 (non-normalized, full counting) (left down), and Productivity 4 (non-normalized, fractional counting) (right down) (N=324,643).



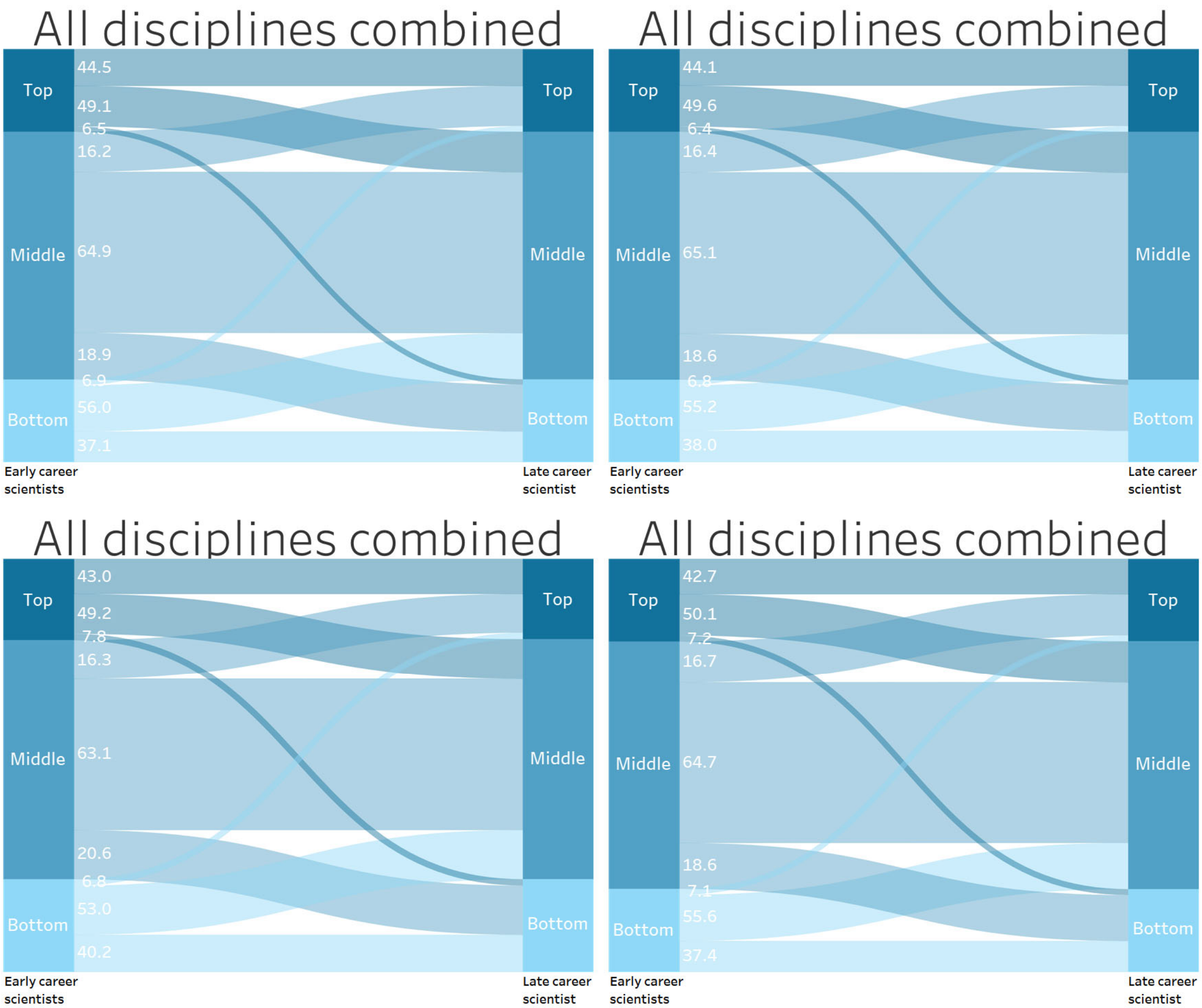
Supplementary Figure 3. Mid-career to late-career stage: horizontal and extreme vertical mobility patterns. Top-to-top and bottom-to-bottom mobility, Jumpers-Up (bottom-to-top mobility) and Droppers-Down (top-to-bottom mobility) by discipline and productivity type, current nonoccasional OECD late-career scientists only (N=324,643). For all disciplines combined, the percentage of scientists who moved from top to top classes differs more between the four productivity types than for the first transition (Total: the 56.40%-60.13% range), as does the percentage for scientists who moved from bottom to bottom classes (Total: the 46.96%-51.09% range). Both cross-productivity differences for disciplines and cross-disciplinary differences within productivity types are substantial.

Mobility Between Productivity Classes: Two Stages of Academic Careers

Apart from examining the three stages of academic careers separately (transitions early to mid-careers, and mid- to late-careers), we have also examined two stages only, that is, the direct transitions from the early career stage to late-career stage. Some scientists have been top performers and bottom performers both in their early career and now in their late-career stage; also, some scientists changed their productivity classes radically.

Figure 12 shows the Sankey diagrams of mobility of the productivity classes between the two stages of an academic career for all STEMM disciplines combined by productivity type. The percentage of scientists who moved from the top productivity class to the top productivity class differs marginally between the four productivity types (42.7–44.5%; see an overview in Table 6). In addition, the percentage for scientists who moved from the bottom class to the bottom class also differs marginally between the four three productivity types (42.7–44.5%).

The pattern for transitions between two stages only is weaker than in the case of the three stages discussed above. The two radical vertical transitions between productivity classes are at a much higher level: the percentage of Jumpers-Up is in the range of 6.9–7.1%, and of Droppers-Down, it is in the range of 6.4–7.8%



Supplementary Figure 4. Productivity 1-4, overview, two career stages. Sankey diagrams of mobility between productivity classes in the two stages (early-career directly to late-career stage) of an academic career. All STEMM disciplines are combined, and only current late-career scientists are shown. Four productivity types: Productivity 1 (prestige-normalized, full counting) (left upper), Productivity 2 (prestige-normalized, fractional counting) (right upper), Productivity 3 (non-normalized, full counting) (left down), and Productivity 4 (non-normalized, fractional counting) (right down), current nonoccasional OECD late-career scientists only (N=324,643).

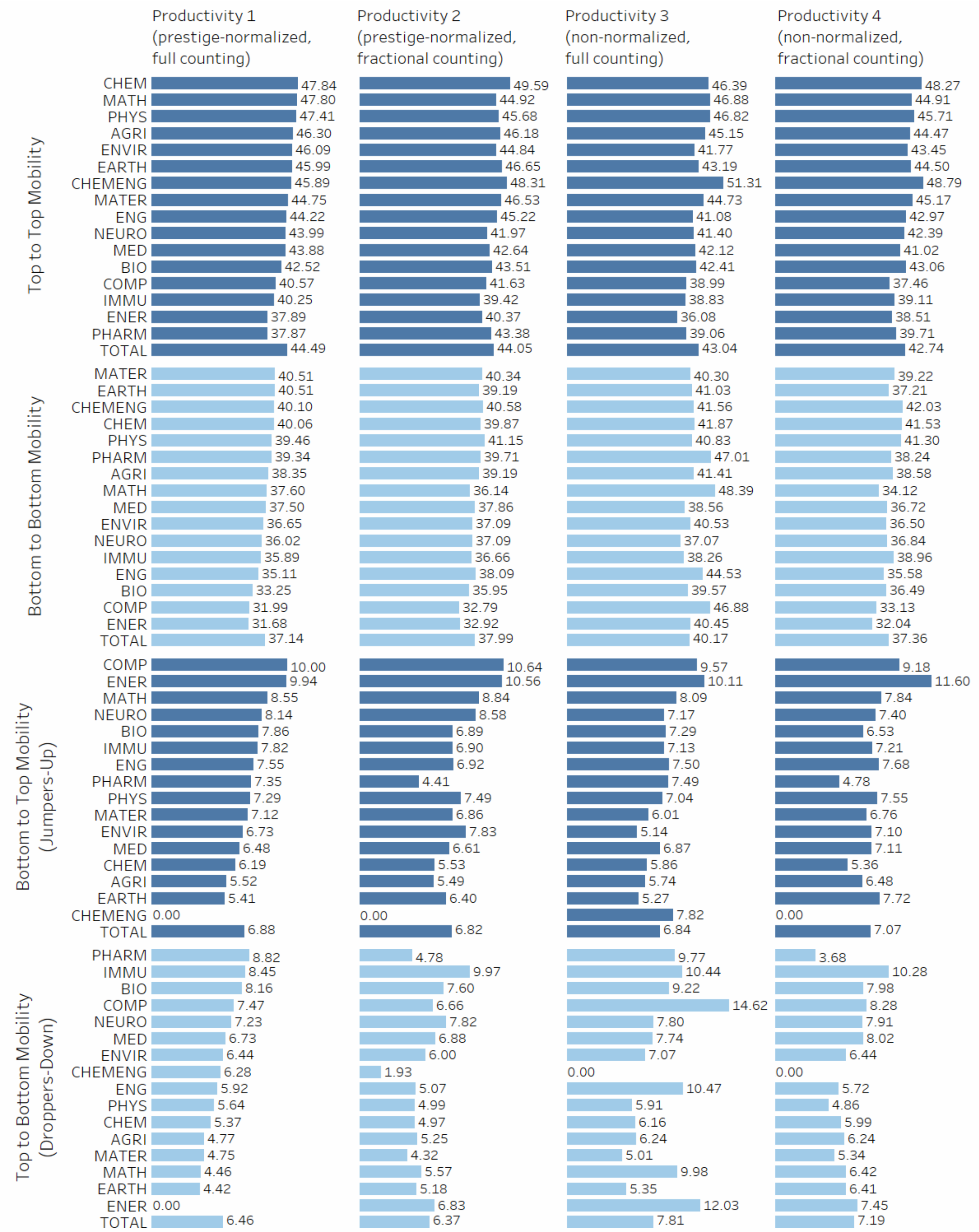
Supplementary Table 5. Overview: **productivity 1-4, two career stages.** Mobility between productivity classes in the two stages (early-career directly to late-career stage) of academic careers, current nonoccasional OECD late-career scientists only, all disciplines (N=324,643)

				Productivity 1 (prestige-normalized, full counting)			Productivity 2 (prestige-normalized, fractional counting)			Productivity 3 (non-normalized, full counting)			Productivity 4 (non-normalized, fractional counting)		
Transition from source academic position	Transition from productivity class	Transition to target academic position	Transition to productivity class	Number of scientists in transition	Number of scientists in each productivity class	%	Number of scientists in transition	Number of scientists in each productivity class	%	Number of scientists in transition	Number of scientists in each productivity class	%	Number of scientists in transition	Number of scientists in each productivity class	%
Early career	Bottom	Late career	Bottom	24,148	65,023	37.1	24,671	64,934	38.0	29,272	72,877	40.2	24,410	65,330	37.4
Early career	Bottom	Late career	Middle	36,402	65,023	56.0	35,834	64,934	55.2	38,619	72,877	53.0	36,303	65,330	55.6
Early career	Bottom	Late career	Top	4,473	65,023	6.9	4,429	64,934	6.8	4,986	72,877	6.8	4,617	65,330	7.1
Early career	Middle	Late career	Bottom	36,854	194,697	18.9	36,141	194,778	18.6	38,672	187,829	20.6	36,096	194,394	18.6
Early career	Middle	Late career	Middle	126,275	194,697	64.9	126,740	194,778	65.1	118,602	187,829	63.1	125,855	194,394	64.7
Early career	Middle	Late career	Top	31,568	194,697	16.2	31,897	194,778	16.4	30,555	187,829	16.3	32,443	194,394	16.7
Early career	Top	Late career	Bottom	4,191	64,923	6.5	4,134	64,931	6.4	4,993	63,937	7.8	4,670	64,919	7.2
Early career	Top	Late career	Middle	31,848	64,923	49.1	32,194	64,931	49.6	31,425	63,937	49.2	32,500	64,919	50.1
Early career	Top	Late career	Top	28,884	64,923	44.5	28,603	64,931	44.1	27,519	63,937	43.0	27,749	64,919	42.7
Late career	Bottom			65,193	65,193	100	64,946	64,946	100	72,937	72,937	100	65,176	65,176	100
Late career	Middle			194,525	194,525	100	194,768	194,768	100	188,646	188,646	100	194,658	194,658	100
Late career	Top			64,925	64,925	100	64,929	64,929	100	63,060	63,060	100	64,809	64,809	100

Mobility Between Productivity Classes: Two Stages of Academic Careers, Cross-disciplinary Differences

The cross-productivity differences between disciplines in the top-to-top mobility are generally in the range of 10 p.p.; they are lower in the bottom-to-bottom mobility. Figure 13 shows the differences in mobility patterns for each discipline by productivity type. Cross-disciplinary differences within productivity types are substantial for all four transition patterns examined, and they are much higher for transitions between two stages than for the three-stage transitions studied above. For instance, in Productivity 1, top-to-top mobility ranges from 37.87% (PHARM) to 47.84% (CHEM). Cross-disciplinary differences are higher for bottom-to-bottom mobility than for top-to-top mobility; in Productivity 1, bottom-to-bottom mobility ranges from 31.68% to 40.51%. In addition, in the case of Jumpers-Up and Droppers-Down, the differences are much higher than in the case of three-stage transitions, reaching the 0.0–10.00% range for the former and 0.0–8.82% range for the latter mobilities.

There are disciplines in which extreme upward mobility opportunities are higher and others, in which they are very limited (leading to different distribution of Jumpers-Up by discipline). There are also disciplines in which extreme downward mobility is a viable option and others in which it is marginal phenomenon (leading to different distribution of Droppers-Down by discipline). There is an interesting special case: in chemical engineering (CHEMENG), the chances for being a Jumper-Up are zero (in three out of four productivity types).



Supplementary Figure 5. Horizontal and extreme vertical mobility patterns (two career stages: early-career directly to late-career stage). Top to top and bottom to bottom mobility, Jumpers-Up (bottom to top mobility) and Droppers-Down (top to bottom mobility) by discipline and productivity type, current nonoccasional OECD late-career scientists only (N=324,643)

Collinearity: In order to check the combined multidimensional influence of dependent variables in the model, an analysis of inverse correlation matrices was performed and main diagonals were analyzed: none of the variables in any of the models is characterized by significantly larger values than the others; no collinearity is reported.

Supplementary Table 6. Inverse correlation matrix main diagonal, top productivity model, scientists in mid-career class.
Table for bottom productivity model available upon request (space limitations).

Variable	AGRI	BIO	CHEM	CHEMENG	COMP	EARTH	ENER	ENG	ENVIR	IMMU	MATER	MATH	MED	NEURO	PHARM	PHYS
Male	1.012	1.032	1.014	1.019	1.003	1.009	1.002	1.003	1.010	1.034	1.010	1.020	1.019	1.027	1.017	1.007
FWCI 4y	1.345	1.147	1.138	1.161	1.105	1.180	1.279	1.177	1.286	1.211	1.186	1.203	1.070	1.182	1.247	1.086
International Collab. Rate	1.114	1.159	1.095	1.117	1.062	1.303	1.205	1.094	1.129	1.161	1.168	1.046	1.139	1.124	1.103	1.447
AJPR	1.420	1.230	1.299	1.216	1.228	1.252	1.503	1.293	1.392	1.251	1.359	1.249	1.294	1.292	1.394	1.119
Median Team Size	1.192	1.234	1.142	1.140	1.089	1.310	1.368	1.094	1.246	1.258	1.277	1.112	1.208	1.209	1.196	1.440
Early Career Top Class	1.120	1.044	1.083	1.094	1.097	1.121	1.134	1.098	1.114	1.082	1.114	1.106	1.084	1.071	1.071	1.178

AJPR = Average Journal Percentile Rank

Supplementary Table 7. Inverse correlation matrix main diagonal, top productivity model, scientists in late-career class.
Table for bottom productivity model available upon request (space limitations).

Variable	AGRI	BIO	CHEM	CHEMENG	COMP	EARTH	ENER	ENG	ENVIR	IMMU	MATER	MATH	MED	NEURO	PHARM	PHYS
Male	1.019	1.028	1.008	1.018	1.005	1.009	1.012	1.005	1.012	1.038	1.011	1.018	1.017	1.025	1.038	1.007
FWCI 4y	1.343	1.131	1.170	1.186	1.086	1.128	1.230	1.173	1.252	1.291	1.282	1.209	1.067	1.241	1.276	1.124
International Collab. Rate	1.180	1.214	1.116	1.171	1.121	1.355	1.154	1.157	1.211	1.201	1.219	1.049	1.217	1.169	1.180	1.488
AJPR	1.482	1.325	1.453	1.356	1.279	1.273	1.500	1.329	1.445	1.333	1.479	1.280	1.404	1.397	1.434	1.133
Median Team Size	1.284	1.364	1.300	1.283	1.144	1.379	1.322	1.136	1.297	1.356	1.389	1.137	1.338	1.309	1.358	1.510
TOP200	1.022	1.018	1.044	1.030	1.024	1.017	1.022	1.055	1.033	1.014	1.038	1.034	1.033	1.017	1.040	1.015
Mid Career Top Class	1.122	1.053	1.117	1.117	1.110	1.117	1.107	1.114	1.118	1.088	1.130	1.103	1.088	1.079	1.089	1.200

AJPR = Average Journal Percentile Rank

Average and Median Productivity, by Discipline and Career Stage

Supplementary Table 8. Average annual productivity, by discipline and career stage (Productivity 1: prestige-normalized, full counting).
Tables for Productivity 2, Productivity 3, and Productivity 4 available upon request (space limitations).

Career Stage	Academic Age Group	AGRI	BIO	CHEM	CHEMEN G	COMP	EARTH	ENER	ENG	ENVIR	IMMU	MATER	MATH	MED	NEURO	PHARM	PHYS
Early	25-35	1.028	1.284	1.773	1.204	0.704	1.111	0.743	0.858	1.011	1.314	1.517	0.969	1.308	1.369	1.252	2.232
	36-40	0.906	1.388	1.424	0.956	0.705	0.846	0.487	0.782	0.813	1.381	1.060	0.810	1.391	1.295	1.307	1.830
	41-50	0.778	1.426	1.301	0.836	0.641	0.737	0.468	0.722	0.770	1.374	0.817	0.740	1.362	1.399	1.360	1.300
Middle	25-35	1.980	2.102	3.086	2.329	1.280	2.141	1.864	1.717	2.092	2.281	2.974	1.332	2.357	2.206	1.986	4.511
	36-40	1.674	2.182	2.688	1.980	1.043	1.696	1.026	1.334	1.571	2.448	2.377	1.153	2.229	2.049	1.827	2.965
	41-50	1.461	2.522	2.586	1.628	1.012	1.397	0.857	1.251	1.394	2.346	2.036	1.002	2.417	2.286	2.064	2.643
Late	25-35	2.702	2.531	3.536	3.157	2.146	2.909	3.029	2.629	2.919	3.054	3.758	1.739	3.426	2.605	2.591	5.538
	36-40	2.713	2.663	3.524	3.212	1.943	2.802	2.380	2.530	2.848	3.404	3.683	1.642	3.633	2.517	2.438	5.688
	41-50	2.451	2.540	3.366	3.272	1.803	2.471	3.176	2.244	2.401	2.959	3.431	1.450	3.526	2.497	2.266	5.660

Supplementary Table 9. Median annual productivity, by discipline and career stage (Productivity 1: prestige-normalized, full counting).
Tables for Productivity 2, Productivity 3, and Productivity 4 available upon request (space limitations).

Career Stage	Academic Age Group	AGRI	BIO	CHEM	CHEMENG	COMP	EARTH	ENER	ENG	ENVIR	IMMU	MATER	MATH	MED	NEURO	PHARM	PHYS
Early	25-35	0.769	1.007	1.385	0.885	0.488	0.867	0.530	0.586	0.705	1.063	1.096	0.755	0.884	1.077	0.897	1.506
	36-40	0.704	1.094	1.098	0.769	0.508	0.655	0.283	0.542	0.588	1.158	0.731	0.626	0.965	1.039	1.007	1.260
	41-50	0.583	1.121	0.992	0.580	0.445	0.584	0.232	0.466	0.555	1.169	0.532	0.583	0.937	1.070	1.034	0.945
Middle	25-35	1.403	1.530	2.202	1.562	0.864	1.597	1.098	1.111	1.435	1.713	2.030	0.975	1.433	1.631	1.283	2.091
	36-40	1.258	1.632	2.005	1.571	0.718	1.286	0.533	0.907	1.103	2.015	1.668	0.875	1.452	1.567	1.402	2.036
	41-50	1.104	1.954	1.927	1.099	0.705	1.097	0.381	0.857	1.028	1.991	1.451	0.792	1.630	1.708	1.621	1.741
Late	25-35	1.825	1.773	2.258	1.919	1.404	2.027	1.818	1.611	1.914	1.989	2.278	1.182	1.934	1.818	1.553	2.301
	36-40	1.843	1.813	2.187	2.117	1.300	1.962	1.510	1.558	1.816	2.124	2.421	1.114	2.066	1.664	1.635	2.263
	41-50	1.606	1.737	2.096	1.944	1.108	1.653	1.419	1.337	1.581	1.919	2.096	1.000	1.932	1.735	1.606	2.005

Supplementary Table 10. Medicine. Logistic regression statistics: odds ratio estimates of membership in the class of bottom productive *mid-career* scientists (upper panel) and *late-career* scientists (bottom panel) in Medicine (the bottom 20%), current nonoccasional OECD late-career scientists in Medicine only (N=137,748)

Model	**Model 1: Productivity 1 - Prestige-normalized full counting**				**Model 2: Productivity 2 - Non-normalized full counting**				**Model 3: Productivity 3 - Prestige-normalized fractional counting**				**Model 4: Productivity 4 - Non-normalized fractional counting**			
	Exp (B)	95% C.I. for Exp(B)		Sig.	Exp(B)	95% C.I. for Exp(B)		Sig.	Exp(B)	95% C.I. for Exp(B)		Sig.	Exp(B)	95% C.I. for Exp(B)		Sig.
		Lower	Upper			Lower	Upper			Lower	Upper			Lower	Upper	
Mid-career scientists																
	$R^2 = 0.186$				$R^2 = 0.184$				$R^2 = 0.194$				$R^2 = 0.186$			
Male	0.917	0.89	0.943	0	0.882	0.855	0.909	0	0.884	0.858	0.909	0	0.837	0.81	0.863	0
FWCI 4y	0.998	0.994	1.002	0	0.997	0.993	1.001	0	1.002	0.999	1.005	0	1.002	0.999	1.005	0
International Collab. Rate	0.992	0.991	0.993	0	0.992	0.992	0.993	0	0.994	0.993	0.995	0	0.995	0.994	0.995	0
AJPR	0.993	0.993	0.994	0	0.993	0.993	0.994	0	0.99	0.989	0.99	0	0.99	0.99	0.991	0
Median Team Size	0.988	0.982	0.993	0	1.02	1.015	1.026	0	1.001	0.995	1.006	0	1.023	1.017	1.028	0
Early Career Bottom Class	8.267	8.238	8.297	0	8.634	8.605	8.664	0	8.735	8.71	8.761	0	9.111	9.085	9.137	0
Intercept	0.237	0.191	0.284	0	0.204	0.16	0.249	0	0.296	0.254	0.338	0	0.239	0.198	0.281	0
Late-career scientists																
	$R^2 = 0.111$				$R^2 = 0.114$				$R^2 = 0.121$				$R^2 = 0.101$			
Male	0.985	0.96	1.011	0	0.959	0.934	0.985	0	0.917	0.892	0.942	0	0.894	0.869	0.919	0
FWCI 4y	0.99	0.986	0.994	0	1.001	0.999	1.004	0	1.001	0.999	1.004	0	1.003	1.001	1.005	0
International Collab. Rate	0.994	0.994	0.995	0	0.995	0.994	0.995	0	0.997	0.997	0.998	0	0.997	0.996	0.998	0
AJPR	0.993	0.992	0.994	0	0.991	0.991	0.992	0	0.995	0.994	0.995	0	0.993	0.993	0.994	0
Median Team Size	0.967	0.962	0.972	0	1.062	1.057	1.068	0	0.973	0.968	0.978	0	1.059	1.053	1.064	0
TOP200	0.719	0.684	0.755	0	0.706	0.67	0.743	0	0.784	0.749	0.818	0	0.791	0.756	0.825	0
Mid Career Bottom Class	4.329	4.302	4.357	0	5.055	5.027	5.082	0	5.507	5.481	5.532	0	5.174	5.148	5.2	0
Intercept	0.395	0.345	0.444	0	0.224	0.176	0.272	0	0.311	0.264	0.358	0	0.196	0.15	0.242	0

Note: Sig 0 means p <= 0.001, AJPR = Average Journal Percentile Rank

Supplementary Table 11. Overview of logistic regression models by discipline: odds ratio estimates of membership in the class of bottom productive *mid-career* scientists (upper panel) and *late-career* (bottom panel) (the lower 20%, separately for each discipline). Productivity 1 (prestige-normalized, full counting), current nonoccasional OECD late-career scientists only (N=324,643)
Tables for Productivity 2, Productivity 3, and Productivity 4 available upon request (space limitations).

	AGRI	BIO	CHEM	CHEMENG	COMP	EARTH	ENER	ENG	ENVIR	IMMU	MATER	MATH	MED	NEURO	PHARM	PHYS
Bottom productive mid-career scientists																
R2	0.200	0.141	0.202	0.206	0.138	0.218	0.137	0.180	0.180	0.172	0.234	0.181	0.186	0.178	0.174	0.196
Male	0.982	0.680	0.972					1.349		0.712			0.917	0.706		
FWCI 4y	0.726	0.940	0.828	1.009	0.915	0.754	0.849	0.893		0.890	0.798		0.998	0.691		0.968
International Collab. Rate	0.995	0.997	0.992	0.999	0.997	0.993	0.995	0.996	0.994	0.998	0.994	0.995	0.992	0.997	1.000	0.992
AJPR	0.996	0.997	1.003	0.995	0.995	0.996	0.993	0.993	0.995	0.994	0.999	0.998	0.993	0.999	0.988	0.997
Median Team Size	0.966	1.033	0.982	0.929	0.984	0.982	1.012	1.094	0.995	1.063	0.920		0.988	0.978	1.040	0.999
Early Career Bottom Class	7.718	6.985	10.721					7.172		8.077			8.267	7.297		
Intercept	0.302	0.221											0.237	0.375		
Bottom productive late-career scientists																
R2	0.125	0.087	0.147	0.133	0.075	0.130	0.095	0.114	0.114	0.119	0.146	0.113	0.111	0.102	0.076	0.139
Male		0.889								0.867		1.152	0.985	0.873		
FWCI 4y		0.891			0.962		0.781			0.873		0.761	0.990	0.829	1.019	0.972
International Collab. Rate	0.998	0.995	0.993	0.997	1.000	0.995	0.994	0.993	0.995	0.996	0.992	0.999	0.994	0.995	0.997	0.990
AJPR	0.990	0.997	0.997	0.996	0.985	0.996	0.985	0.987	0.994	0.996	1.000	0.985	0.993	0.998	0.989	1.001
Median Team Size	1.008	0.980			0.987	1.019	1.294	1.113		0.968	1.061	0.961	0.967	0.975	1.028	0.990
Top200		0.787										0.852	0.719	0.793		
Mid-Career Bottom Class		4.125								5.379		3.669	4.329	4.596		
Intercept		0.385											0.395			

Only statistically significant results shown in the table. AJPR = average Journal Percentile Rank.

[1] Ruspini, E. (1999). Longitudinal research and the analysis of social change. Quality and Quantity 33(3), 219–227.
[2] Menard, S. (2002). Longitudinal research. Thousand Oaks: Sage.
[3] Rowland, D. T. (2014). Demographic methods and concepts. Oxford University Press.
[4] Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C.R. (2013). Global gender disparities in science. Nature, 504, 211–213.
[5] 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.
[6] Nielsen, M. W., & Andersen, J. P. (2021). Global citation inequality is on the rise. Proceedings of the National Academy of Sciences, 118(7), e2012208118.
[7] Salganik, M. J. (2018). Bit by bit. Social research in a digital age. Princeton University Press.
[8] Liu, L., Jones, B.F., Uzzi, B., et al. (2023). Data, measurement and empirical methods in the science of science. Nature Human Behaviour, 7, 1046–1058.
[9] Clauset, A., Larremore, D. B., & Sinatra, R. (2017). Data-driven predictions in the science of science. Science, 355, 477–480.
[10] Zeng, A. et al. (2017). The science of science: from the perspective of complex systems. Phys. Rep., 714, 1–73
[11] Wang, D., & Barabási, A.-L. (2021). The science of science. Cambridge: Cambridge University Press.
[12] Leišytė, L., & Dee, J. R. (2012). Understanding academic work in changing institutional environment. Higher Education: Handbook of Theory and Research, 27, 123–206.
[13] 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 27 (pp. 207–248). Dordrecht: Springer.
[14] 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.
[15] Ni, C., Smith, E., Yuan, H., Larivière, V., & Sugimoto, C. R. (2021). The gendered nature of authorship. Sci. Adv., 7, eabe4639.
[16] Allison, P. D., Long, J. S., Krauze, T. K. (1982). Cumulative advantage and inequality in science. American Sociological Review, 47(5), 615–625.
[17] Fox, M. F. (1983). Publication productivity among scientists: A critical review. Social Studies of Science, 13(2), 285–305.
[18] 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). Paris: CNRS.
[19] 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.
[20] Aguinis, H., & O'Boyle, E. (2014). Star performers in twenty-first century organizations. Personnel Psychology, 67(2), 313–350.
[21] 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
[22] Li, W., Aste, T., Caccioli, F., & Livan, G. (2019). Early coauthorship with top scientists predicts success in academic careers. Nature Communications 10, 5170.
[23] Fox, M. F., & Nikivincze, I. (2021). Being highly prolific in academic science: Characteristics of individuals and their departments. Higher Education, 81, 1237–1255.

[24] 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.
[25] Allison, P. D., & Stewart, J. A. (1974). Productivity differences among scientists: Evidence for accumulative advantage. American Sociological Review, 39(4), 596–606.
[26] 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). Amsterdam: Elsevier.
[27] Cole, J. R., & Cole, S. (1973). Social stratification in science. Chicago: The University of Chicago Press.
[28] Merton, R. K. (1973). The sociology of science: Theoretical and empirical investigations. University of Chicago Press.
[29] 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.
[30] 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.
[31] Ma, Y., Mukherjee, S., & Uzzi, B. (2020). Mentorship and protégé success in STEM fields. Proceedings of the National Academy of Sciences, 117, 14077–14083.
[32] 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.
[33] Sugimoto, C., & Larivière, V. (2023). Equity for Women in Science. Dismantling Systemic Barriers to Advancement. Boston: Harvard University Press.
[34] 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
[35] Wang, Y., Jones, B. F., & Wang, D. (2019). Early career setback and future career impact. Nature Communications, 10, 4331.
[36] 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.
[37] Ross, M. B. et al. (2022). Women are credited less in science than men. Nature, 608, 135–145.
[38] Kwiek, M., & Roszka, W. (2023). Once highly productive, forever highly productive? Full professors' research productivity from a longitudinal perspective. Higher Education. Online first. https://doi.org/10.1007/s10734-023-01022-y
[39] Kwiek, M., & Roszka, W. (2023). Moving up the academic ladder, changing productivity? The case of Polish high and low performers. Preprint at SocArXiv: https://osf.io/preprints/socarxiv/tajs4
[40] Marginson, S. (2022) What drives global science? The four competing narratives. Studies in Higher Education, 47(8), 1566–1584. DOI: 10.1080/03075079.2021.1942822
[41] Kwiek, M. (2023). The globalization of science: The increasing power of individual scientists. In P. Mattei, X. Dumay, E. Mangez & Ja. Behrend (Eds.), The Oxford handbook of education and globalization (online ed.). Oxford Academic. https://doi.org/10.1093/oxfordhb/9780197570685.013.16.
[42] 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.

[43] Stephan, P. E., & Levin, S. G. (1992). Striking the mother lode in science: The importance of age, place, and time. Oxford: Oxford University Press.
[44] Allison, P. D. (1980). Inequality and scientific productivity. Social Studies of Science, 10, 163–179.
[45] 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.
[46] Wagner, C. S. (2018). The collaborative era in science. Governing the network. Cham: Palgrave Macmillan.
[47] 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.
[48] Jung, J. (2014). Research productivity by career stage among Korean academics. Tertiary Education and Management, 20(2), 85–105.
[49] 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.
[50] Preston, A. E. (2004). Leaving science. Occupational exit from scientific careers. New York: Russell Sage Foundation.
[51] 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). Amsterdam: Elsevier.
[52] Kwiek, M., & Szymula, Ł. (2023). Quantifying attrition in science: A cohort-based, longitudinal study of scientists in 38 OECD countries. Preprint at SocArXiv: https://doi.org/10.31235/osf.io/8kzb7
[53] 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.
[54] 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.
[55] 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.
[56] Heckman, J. J., & Moktan, S. (2018). Publishing and promotion in economics. The tyranny of the Top Five. NBER Working Paper 25093.
[57] 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.
[58] 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.
[59] 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.
[60] Ruiz-Castillo, J., & Costas, R. (2014). The skewness of scientific productivity. Journal of Informetrics, 8(4), 917–934.
[61] Stephan, P. (2012). How economics shapes science. Cambridge, MA: Harvard University Press.
[62] Latour B. & Woolgar S. (1986) Laboratory life. The construction of scientific facts. Princeton University Press.
[63] 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, 226–230.

[64] 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.
[65] 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. Online first. https://doi.org/10.1162/qss_e_00276
[66] 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.
[67] Sugimoto, C., & Larivière, V. (2018). Measuring research: What everyone needs to know. Oxford: Oxford University Press.
[68] Spoon, K. et al. (2023). Gender and retention patterns among U.S. faculty. Science Advances, 9, eadi2205. DOI:10.1126/sciadv.adi2205
[69] 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