

## Association between productivity and journal impact across disciplines and career age

Andre S. Sunahara <sup>1</sup>, Matjaž Perc <sup>2,3,4,5</sup> and Haroldo V. Ribeiro <sup>1,\*</sup>

<sup>1</sup>*Departamento de Física, Universidade Estadual de Maringá, Maringá, PR 87020-900, Brazil*

<sup>2</sup>*Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia*

<sup>3</sup>*Department of Medical Research, China Medical University Hospital, China Medical University, Taichung 404332, Taiwan*

<sup>4</sup>*Alma Mater Europaea, Slovenska ulica 17, 2000 Maribor, Slovenia*

<sup>5</sup>*Complexity Science Hub Vienna, Josefstädterstraße 39, 1080 Vienna, Austria*



(Received 9 February 2021; accepted 4 August 2021; published 16 August 2021)

The association between productivity and impact of scientific production is a long-standing debate in science that remains controversial and poorly understood. Here we present a large-scale analysis of the association between yearly publication numbers and average journal-impact metrics for the Brazilian scientific elite. We find this association to be discipline specific, career age dependent, and similar among researchers with outlier and nonoutlier performance. Outlier researchers either outperform in productivity or journal prestige, but they rarely do so in both categories. Nonoutliers also follow this trend and display negative correlations between productivity and journal prestige but with discipline-dependent intensity. Our research indicates that academics are averse to simultaneous changes in their productivity and journal-prestige levels over consecutive career years. We also find that career patterns concerning productivity and journal prestige are discipline-specific, having in common a raise of productivity with career age for most disciplines and a higher chance of outperforming in journal impact during early career stages.

DOI: [10.1103/PhysRevResearch.3.033158](https://doi.org/10.1103/PhysRevResearch.3.033158)

### I. INTRODUCTION

The development of knowledge-based economies and the increasing availability of information and knowledge itself have driven transdisciplinary efforts toward a better quantitative understanding of the scientific enterprise: the science of science [1,2]. Beyond the academic question of finding driving mechanisms of science, these initiatives aim to enhance scientific efficiency by identifying successful practices and policies, from the choice of countries' scientific priorities to the selection of research projects and faculty candidates. Scientific progress is nowadays strongly dependent on research evaluation processes, as they regulate the stream of ideas and research projects by means of science funding allocation [2–5]. But while peer review is considered the standard approach for assessing academic performance [6], the process itself is laborious and has several drawbacks, ranging from biases and lack of consistency to fraud [6–9]. In addition, the increasing number of scientific publications [10] and the growth of the scientific workforce [11] impose further limitations on the peer-review method [5]. A direct consequence of these issues is the steady increase (especially after the 2000s [12]) in the use of bibliometric indexes for grading the performance of researchers [5,13].

Bibliometric assessments are considered more objective criteria, but there is no consensus on which indexes are more suitable for evaluating academic performance, and many believe that the intrinsic nature of scientific processes can only be precisely quantified by multidimensional features [14,15]. This data-driven culture of performance evaluation has amassed much criticism [6,16–18], and it also exerts enormous pressure on scholars (particularly on young scientists [19]) for publishing in large quantities, in prestigious journals, and developing highly cited research [5,20,21]. Still, research productivity and impact measures are often and widely used to quantify academic performance, representing essential ingredients for the perception and recognition of academic success. While productivity is defined as the number of research items in a given period, impact has a more subjective character and is usually measured by the number of citations, the share of articles among highly cited papers, and the prestige of the publication venue. Regardless of which metric is used, research evaluation via bibliometrics has raised the “quality versus quantity” debate since its conception [22–37], and there is still no agreement on the association between productivity and impact. For instance, while Larivière and Costas [32] have found a positive association between productivity and number of highly cited articles, Bornmann and Tekles [36] have shown that top productive authors usually have lower fractions of publications among top-cited articles (that is, a negative association between productivity and impact at overly high productivity levels). An important part of these controversial findings reflects the fact that the association between productivity and impact is discipline specific, career age dependent, and scale dependent, and it may be affected by the presence of outlier individuals. However, there is still a lack of works

\*hvr@dfi.uem.br

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

simultaneously addressing all these points to reveal the overall complexity of the “quantity versus quality” relationship.

Here we investigate multifaceted aspects of this association by analyzing the scientific career of more than six thousand scientists from the Brazilian scientific community’s elite from 14 different disciplines. We determine the yearly publication numbers and the respective average value of the journal-impact metrics over the careers of these academics. Although the use of journal-level metrics for assessing the individual performance of researchers is controversial [16,38], this approach remains widespread [39], especially in Brazil where several universities use journal prestige (or derivative indicators) for everything from grading resumes of graduate students to the selection of tenure-track faculty positions and grant applications. Recent works have also demonstrated that journal-level metrics carry information about academic performance [40–44] and that these metrics are correlated with citations, thus indicating that citations and journal-level metrics are partly substitutes. Whether journal-level metrics (or even citations) are suitable or not for research evaluation, these metrics are still important for the scientific community and deserve further investigation.

Our research probes patterns of the association between productivity and journal-impact metrics throughout researchers’ careers across different disciplines. In contrast to previous works, we use standard score measures to account for discipline and inflation-like effects and correct for size-dependent biases on average journal prestige. We further identify outlier individuals in productivity and journal impact, finding these academics to either outperform in productivity or journal prestige over their careers but rarely in both categories. We also find that academics are averse to simultaneously changing their levels of productivity and journal prestige and prefer maintaining these levels over consecutive years of their careers. For nonoutlier individuals, our results indicate a negative correlation between productivity and journal prestige for most researchers from most disciplines. However, we show that career patterns of productivity and journal prestige are discipline specific, although they have in common the fact that productivity increases with time for all disciplines. By shedding light on career age and discipline-specific aspects of productivity and journal prestige, we believe our work may contribute significantly to a more comprehensive and fair research evaluation process.

## II. RESULTS

### A. Journal prestige versus productivity plane

To investigate the association between productivity and journal prestige, we have collected the academic curricula of 6028 Brazilian researchers from 14 disciplines (see the Methods Sec. IV A for details) holding the CNPq Research Productivity Fellowship (*Bolsa Produtividade em Pesquisa do CNPq*) as of May 2017. This fellowship has been awarded since the 1970s by the Brazilian National Council for Scientific and Technological Development (CNPq, *Conselho Nacional de Desenvolvimento Científico e Tecnológico*) in recognition of outstanding scientific production. CNPq fellows have significant status among the Brazilian scientific

community and are often considered the elite of Brazilian scientists. We further obtain the journal impact factor (JIF) between 1997 and 2015 from Clarivate’s journal citation reports. We combine these data sets to assign the time-varying values of JIF to the 312 881 articles published by the CNPq fellows between 1997 and 2015. We consider the number of articles published per year as the productivity indicator and the average JIF as a proxy for journal prestige. We have also carried out a comparative analysis when considering the Scopus’s SCImago journal rank (SJR) as an indicator of journal prestige. Despite the substantial differences in the definitions of JIF and SJR, both measures of journal prestige are strongly correlated (Fig. S1 [45]), and yield very similar results. We have opted to present the results for the JIF in the main text, and we refer to the Supplemental Material [45] for comparisons with the SJR.

We start our investigation by noticing that the number of articles and citations have increased over time [46,47]. This produces inflation in productivity and journal-impact measures that needs to be accounted for a fair comparison between different publication years. Our results indicate that the average productivity of the CNPq fellows has increased at a rate of  $\approx 1.57$  papers/year per decade. Similarly, the average JIF of these publications has raised by  $\approx 0.72$  units per decade (Figs. S2 and S3 [45]). This inflation effect is different among disciplines; for instance, the productivity of medical researchers has increased by  $\approx 3.5$  papers/year per decade, while those working in electrical engineering experienced a productivity inflation of  $\approx 0.3$  papers/year per decade. Because of this inflation effect and differences in publication patterns among disciplines, we do not use raw numbers of productivity but instead robust standard scores ( $z$  scores) relative to discipline and year of publication. In addition to discipline and year, the robust standard scores for average journal prestige are also relative to researchers’ productivity levels. This additional normalization accounts for the fact that the more productive a researcher is in a given year, the narrower the range of variation of his or her average journal prestige. A similar size effect has been observed by Antonoyiannakis [48,49] when comparing the impact factor of journals with different sizes, and the approach we use for rescaling the average journal prestige is an adapted version of his method for ranking journals [48,49] (see the Methods Sec. IV B for details).

Figure 1(a) shows a scatter plot of the average journal prestige versus the productivity for all career years of researchers in our data set (see Fig. S4 [45] for a comparison with the SJR data). In this plane, a unit of productivity indicates a performance one standard deviation above (if positive) or below (if negative) the average performance of all scholars of a given discipline in a given year. Similarly, a unit of average journal prestige represents a performance one standard deviation above (if positive) or below (if negative) the average random performance at a given productivity level of a given discipline and year. We divide this plane into four main sectors separating outlier years of researchers ( $z$  scores higher than 3.5) regarding productivity ( $P$ ) and average journal impact ( $I$ ). The sector  $IP++$  contains career years in which researchers were simultaneously outliers in journal prestige and productivity ( $I > 3.5$  and  $P > 3.5$ ). Similarly, sectors  $I++$  and  $P++$  indicate outlier career years only regarding journal

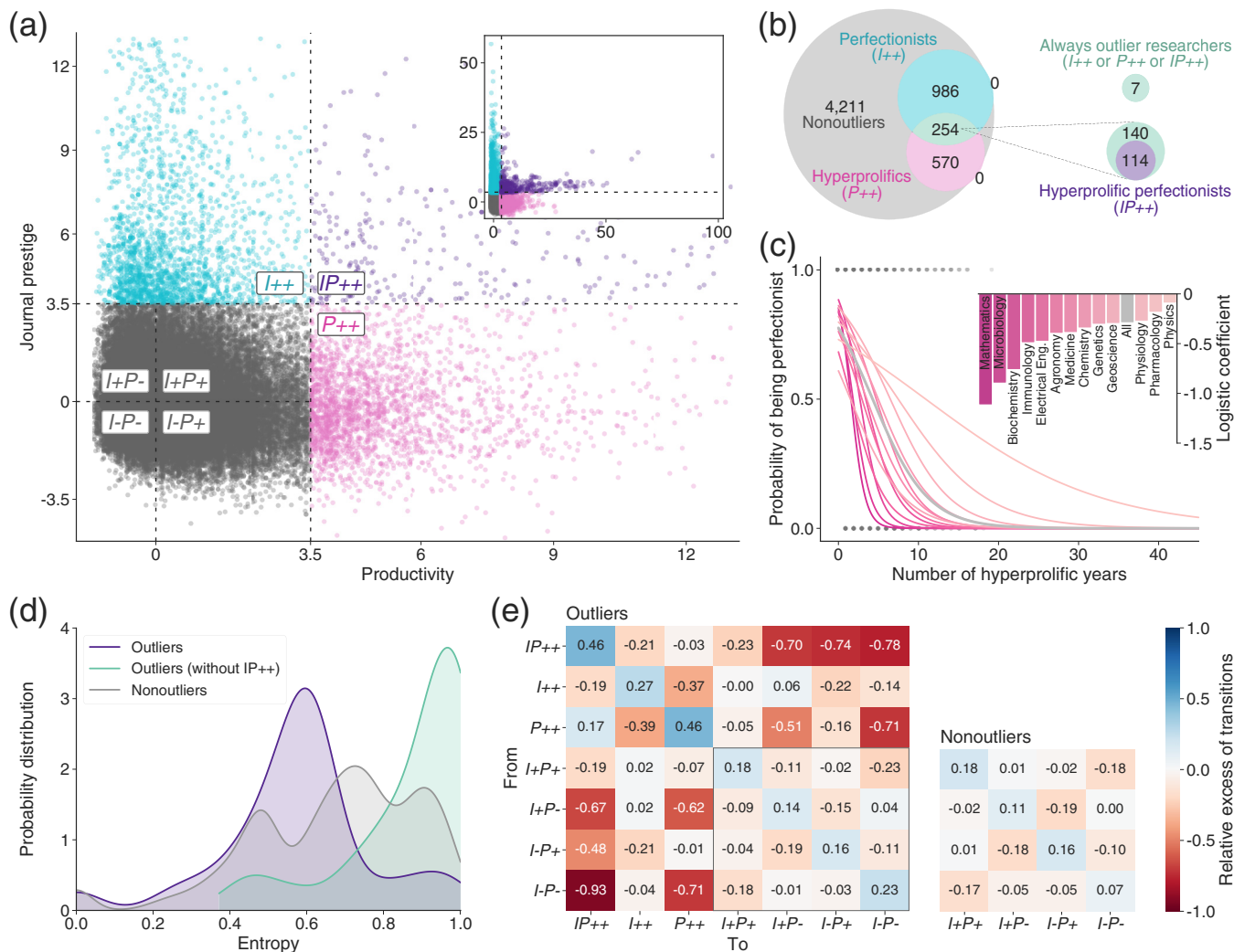


FIG. 1. Journal prestige vs productivity. (a) Relation between average journal impact and productivity in standard score units (the inset shows the full range of the plane). Data points represent career years of researchers from the 14 disciplines in our study. This plane is divided into seven sectors. Three sectors represent career years with overly high performance in journal impact ( $I++$ ), productivity ( $P++$ ), or both quantities ( $IP++$ ). Four nonoutlier sectors represent career years with productivity and journal impact above ( $I+P+$ ) or below ( $I-P-$ ) the average, journal impact below and productivity above ( $I+P-$ ), and journal impact above and productivity below the average ( $I-P+$ ). (b) Venn diagram showing the set relations among the four categories of researchers. Nonoutliers are those with all career years in nonoutlier sectors. Perfectionists and hyperprolifics are researchers with at least one career year in sectors  $I++$  and  $P++$ , respectively. Hyperprolific perfectionists are those having at least one career year within sector  $IP++$ . (c) Probability of being a perfectionist researcher while having a given number of career years in the hyperprolific sector ( $P++$ ), as estimated via logistic regression (the inset shows the logistic coefficients). The colored curves (and bars) refer to different disciplines, while the gray curve represents the aggregated result of all disciplines. Materials engineering (omitted in this panel) is the only discipline that does not display a significant association. (d) Probability distribution of the normalized entropy values associated with the occupation of the plane sectors over researchers' careers. The purple curve shows the results for the occupation of only outlier sectors by outlier researchers and the green curve is the same but after ignoring sector  $IP++$ . The gray curve shows the entropy distribution for nonoutlier researchers. (e) Transition matrix among the plane sectors for outlier (left) and nonoutlier (right) researchers. Each cell represents the relative excess of transitions between two sectors compared with a null model corresponding to shuffled versions of researchers' careers for 10 000 realizations.

prestige ( $I > 3.5$  and  $P < 3.5$ ) and productivity ( $I < 3.5$  and  $P > 3.5$ ), respectively. We further divide the nonoutlier sector ( $I < 3.5$  and  $P < 3.5$ ) into four other sectors:  $I+P+$  for career years with journal prestige and productivity above the average ( $I > 0$  and  $P > 0$ );  $I+P-$  for career years with journal prestige above and productivity below the average ( $I > 0$  and  $P < 0$ );  $I-P+$  for career years with journal prestige below and productivity above the average ( $I < 0$  and  $P > 0$ ); and  $I-P-$

for career years with journal prestige and productivity below the average ( $I < 0$  and  $P < 0$ ).

### B. Outlier and nonoutlier researchers

One of the most striking features of the plane shown in Fig. 1(a) is the existence of researchers who, despite belonging to the elite of Brazilian scientists, further differentiate

themselves by exhibiting productivity or average journal prestige (or both) in overly high levels at specific years of their careers. These outlier career years are relatively rare and represent only 7.7% of the 76 454 total career years (Fig. S5A [45]). Among the outlier sectors, the number of career years in  $P++$  and  $I++$  represent 47% and 46% of the total, respectively. Consequently, career years in sector  $IP++$  are much rarer and correspond to only 7% of the total outlier years. Similar results are obtained for the SJR data set (Fig. S5B [45]).

Outlier years also represent only a small fraction of the careers of researchers covered by our data set (Fig. S6A [45]). More than 47.6% of these researchers are outliers in productivity or journal prestige (or both) only in one year, and only 6.7% have more than 50% of their career years in outlier sectors. The Venn diagram of Fig. 1(b) depicts the set relations between researchers categorized as nonoutlier (all career years in nonoutlier sectors), perfectionist (at least one career year in sector  $I++$ ), hyperprolific (at least one career year in sector  $P++$ ), and hyperprolific perfectionist (at least one career year in sector  $IP++$ ). About 30% of all researchers manage to have at least one career year in outlier sectors. There is no researcher with all career years in sector  $IP++$  or in either sectors  $I++$  or  $P++$ . In addition, only seven researchers (a chemist, an agronomist, and five physicists) have all career years covered by our data set in the three outlier sectors. Similar results are found for the SJR data set (Fig. S6B [45]).

Among the 1817 outlier researchers, 1556 (85.6%) are only hyperprolific or only perfectionist over their careers. This result indicates that most outlier researchers have a persistent behavior regarding being hyperprolific or perfectionist. This clear distinction between hyperprolifics and perfectionists is further corroborated by the existence of only 121 researchers (6.7% of the outliers) simultaneously outliers in both categories, that is, in sector  $IP++$ . A similar pattern was recently observed by Bornmann and Tekles [36] for the association between productivity and number of articles in the top 1% most cited. Our result thus indicates that it is extremely hard to frequently publish in very prestigious journals and keep productivity at overly high levels. Intriguingly, we observe that extremely hyperprolific research years ( $P > 27.7$ ) are all in sector  $IP++$ . This result shows that while very rare, there are 16 researchers capable of maintaining extreme performances in productivity and journal prestige.

To reinforce this result, we use logistic regression to estimate the effect of hyperprolific years on the probability of being a perfectionist researcher (see the Methods Sec. IV C for details). Figure 1(c) shows the probability of being a perfectionist researcher as a function of the number of hyperprolific years and the logistic coefficients when considering all disciplines both together and separated. Materials engineering does not show a significant association ( $p$  value  $> 0.05$ ) and has been omitted in Fig. 1(c). For the other 13 disciplines and when aggregating all disciplines, the coefficients are significant and negative, establishing that an increase in the number of hyperprolific years decreases the chances of being a perfectionist researcher. However, this effect varies considerably among the disciplines. For instance, while five hyperprolific years practically prevent the existence of perfectionist researchers in mathematics, there is a probability

of 63.2% of being a perfectionist in physics with the same number of hyperprolific years. For the SJR data set, 23 out of 25 disciplines display a negative and significant association between the number of hyperprolific years and the probability of being a perfectionist (Fig. S4C [45]), reaffirming the negative association between these two behaviors.

The group of 261 researchers who manage to publish both as perfectionist and hyperprolific (simultaneously or not) is significantly more productive than the exclusively hyperprolific ones (average  $z$ -score productivity of  $2.71 \pm 0.08$  versus  $2.06 \pm 0.03$ ;  $p$  value  $< 10^{-16}$ , permutation test) and the exclusively perfectionist ones (average  $z$ -score productivity of  $2.71 \pm 0.08$  versus  $0.54 \pm 0.02$ ;  $p$  value  $< 10^{-16}$ , permutation test). Furthermore, this former group of researchers publish in journals with higher prestige than hyperprolific (average  $z$ -score JIF of  $1.89 \pm 0.05$  versus  $0.23 \pm 0.02$ ;  $p$  value  $< 10^{-16}$ , permutation test) and perfectionist researchers (average  $z$ -score JIF of  $1.89 \pm 0.05$  versus  $1.45 \pm 0.02$ ;  $p$  value  $< 10^{-16}$ , permutation test). We find similar results for the SJR data.

We have also quantified whether outlier researchers have a preference for a particular outlier sector. To do so, we consider only career years in outlier sectors to estimate their corresponding fractions and calculate the normalized Shannon entropy for every outlier researcher outperforming in more than one category. Entropy values close to unity represent more alternating behaviors, while values around zero indicate that these researchers prefer a given outlier sector. Figure 1(d) shows that the distribution of these entropy values has a peak around 0.6 (purple curve), suggesting a preference for particular outlier sectors. However, if we do not consider sector  $IP++$  (the most underpopulated sector), the entropy shifts to higher values and its distribution peaks around 1 (green curve), indicating that there is no preference between sectors  $I++$  and  $P++$  for researchers publishing in both sectors. In this aspect, these atypical researchers are not so different from those present only in nonoutlier sectors. As shown in Fig. 1(d) (gray curve), nonoutlier researchers also do not exhibit a strong preference for any sector over their careers. The same patterns are observed for the SJR data set (Fig. S4D [45]).

Another intriguing question is whether there are more frequent transitions among sectors of the journal prestige versus productivity plane over the researchers' careers. To investigate this hypothesis, we estimate the number of transitions among all plane sectors and compare them with a null model defined by the average number of transitions estimated after randomly shuffling researchers' careers (over 10 000 realizations). This process allows us to estimate the relative excess for all possible transitions (that is, the number of transitions among sectors during consecutive career years minus the average value of this quantity estimated from the shuffled careers further divided by this same average value). Figure 1(e) shows these transition matrices when grouping researchers in outlier and nonoutlier categories. Both matrices are almost symmetric and have positive diagonal elements among the highest absolute values, indicating that most transitions have no preferential direction and a short-term trend to remain in the same sector. For outlier researchers, the transitions  $IP++ \bullet \rightarrow IP++$ ,  $I++ \bullet \rightarrow I++$ , and  $P++ \bullet \rightarrow P++$  have the largest excesses among all self-transitions. For nonoutliers,  $I+P+ \bullet \rightarrow I+P+$  and  $I-P+ \bullet \rightarrow I-P+$  are

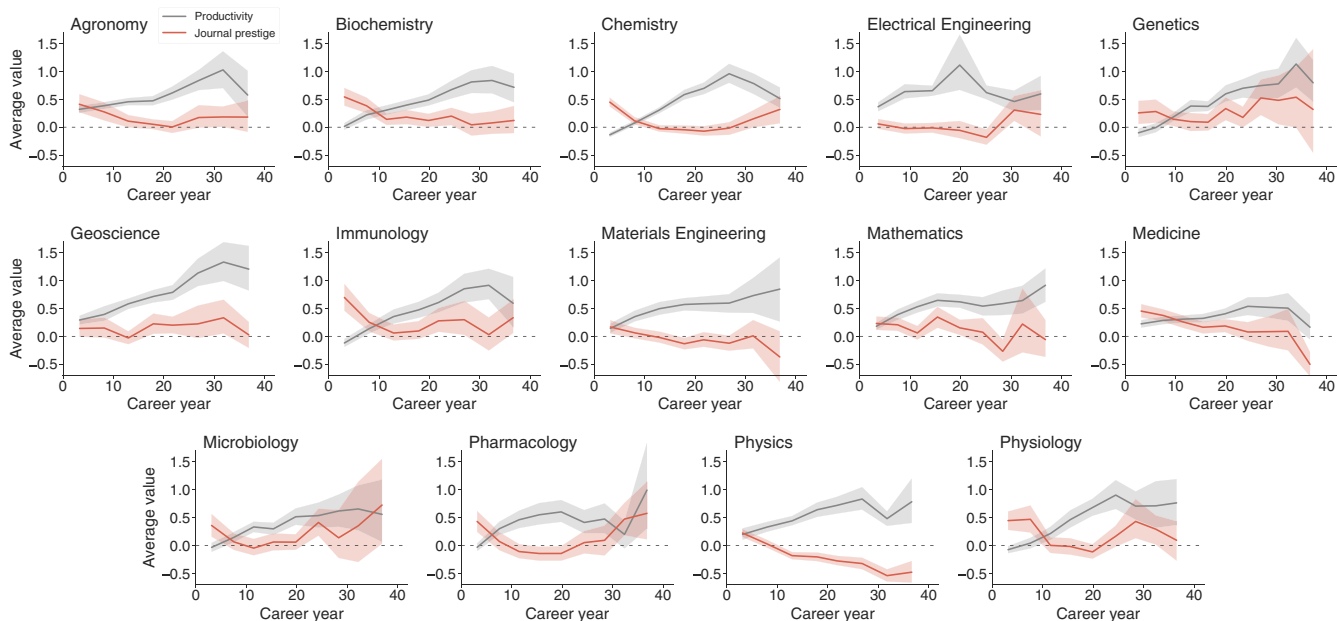


FIG. 2. Average productivity and journal impact over researchers' careers for different disciplines. These visualizations show the average productivity (gray curves) and the average journal prestige (red curves) calculated within 5-yr sliding windows over career years for each discipline in the JIF data set. Shaded areas correspond to bootstrapping 95% confidence intervals. Average productivity increases with career progression for all disciplines (Fig. S8A [45]) and shows a plateau or small decrease in later career stages for most disciplines. Although some disciplines display more complex patterns, average journal prestige has a subtle downward trend and is usually larger in initial career stages for most disciplines (Fig. S8B [45]).

the self-transitions with largest excesses. Intriguingly, the self-transition  $I-P- \bullet \rightarrow I-P-$  (lowest prestige and lowest productivity sector) has an excess that is larger for outlier (23%) than for nonoutlier (7%) researchers.

The transitions among nonoutlier sectors are marked by a negative excess when simultaneously changing levels of productivity and journal impact ( $I+P\pm \leftrightarrow I-P\mp$ ). These transitions represented by the antidiagonal elements in the nonoutlier matrix are less frequent over the careers of outlier and nonoutlier researchers. A similar pattern is observed for transitions involving the outlier sectors  $I++$  and  $P++$ ; that is, the transitions  $I++ \leftrightarrow P++$ ,  $P++ \bullet \rightarrow I+P-$ , and  $P++ \bullet \rightarrow I-P-$  are also less frequent over the careers of outlier researchers. Conversely, transitions among different sectors with similar productivity or journal prestige (for instance,  $I+P+ \leftrightarrow I-P+$  and  $I-P- \leftrightarrow I+P-$ ) usually have excesses close to zero and are thus about as frequent as those occurring in the null model. Together with the excess of self-transitions, these results suggest an aversion to simultaneously changing productivity and journal prestige levels and a preference for maintaining these levels over consecutive years of researchers' careers.

We further notice that most transitions between outlier and nonoutlier sectors occur much less frequently than by chance (negative or close to zero excesses). Career years in sector  $P++$  are usually not preceded nor followed by years in low productivity sectors ( $I+P-$  and  $I-P-$ ). Conversely, career years in sector  $I++$  are less followed and less preceded by years in low journal prestige sectors ( $I-P+$  and  $I-P-$ ). It is also worth noticing that career years in the sector  $IP++$  are more often preceded by years in sector  $P++$  than  $I++$ , suggesting that it is easier for hyperprolifics

to become hyperprolific perfectionists than for perfectionist researchers.

We find overall similar results for the SJR data set (Fig. S4E [45]). The main differences emerge for transitions involving sector  $IP++$ . Outside the diagonal, the two largest transitions for outlier researchers are  $IP++ \rightarrow I++$  and  $IP++ \rightarrow P++$  with 14% and 12% excesses, respectively. This result suggests  $I++$  and  $P++$  years are more commonly preceded by  $IP++$  years when considering SJR as the measure of journal prestige. In addition, although sector  $IP++$  is still more often preceded by hyperprolific years ( $P++$ ) than by perfectionist ones ( $I++$ ), the difference is not as substantial as it is for the JIF data set. All other transitions display about the same behavior. We also verify that the SJR results are robust when considering only the disciplines present in the JIF data set (Fig. S7 [45]).

### C. Effects of career age

We have also investigated the effects of researchers' career age on the average journal prestige and productivity. To do so, we consider the year after Ph.D. graduation as the first career year. Next, we calculate the average productivity and average journal prestige within 5-yr sliding windows of career years for all disciplines. Figure 2 shows these average values as a function of researchers' career years. We observe a significantly increasing trend of average productivity over the years for all disciplines (Fig. S8A [45]), followed by a plateau or slight decrease in the latest career years. For average journal prestige, while some disciplines show complex patterns, we observe these values are slightly larger during first career years and present a subtle downward trend for



FIG. 3. Occupation trends in the journal prestige vs productivity plane over researchers' careers. These visualizations show the fraction of career years in each nonoutlier sector and in outlier sectors  $I++$  and  $P++$  as a function of researchers' career age for the 14 disciplines in the JIF data set. Columns indicate 5-yr intervals and lines represent the different sectors. The same color code indicates the fractions for the nonoutlier sectors (gray shades), and the other two color codes are used for the outlier sectors  $I++$  (blue shades) and  $P++$  (pink shades). Sector  $IP++$  is omitted because career years in this sector are very rare. We observe that low-productivity sectors are more populated during initial career years and a shifting trend toward high-productivity sectors in later career stages for most disciplines. Only 5-yr intervals having at least 20 researchers are shown in these visualizations.

most disciplines (Fig. S8B [45]). Figures S9 and S10 [45] show similar results for the SJR data set. We remark, however, that these average trends for disciplines may not represent the individual behavior of researchers, as we shall discuss in the next section.

To further characterize the effects of researchers' career age on productivity and journal prestige, we divide scholarly careers into 5-yr intervals and estimate the average fraction of career years in each sector of the journal prestige versus productivity plane as a career age function. Figure 3 shows these fractions for all disciplines in our study. In this matrix representation, columns stand for career-year intervals, lines indicate different plane sectors, and color codes stand for the fraction values. Because our data comprise researchers at different career stages, this analysis spans a time interval in career years larger than the number of years in the JIF data set (19 years).

Figure 3 indicates that occupation trends in the journal prestige versus productivity plane vary among disciplines (see Fig. S11 [45] for results based on the SJR data set). However, some evolution patterns are common. By analyzing the nonoutlier sectors, we observe a concentration in low-productivity sectors ( $I+P-$  and  $I-P-$ ) during initial career years and a shifting trend to high-productivity sectors ( $I+P+$  and  $I-P+$ ) in later career stages of researchers from most disciplines. This trend is particularly evident in physics and chemistry, for which we observe a more pronounced growth in sector  $I-P+$ . For the outlier sectors, we notice a low prevalence in sector  $P++$  during initial career stages and an increasing trend over time for all disciplines. This rise in productivity levels over the years may reflect the consolidation of researchers' careers and the likely increase of their scientific collaborations. Furthermore, these patterns for nonoutlier and outlier researchers agree with the overall increasing

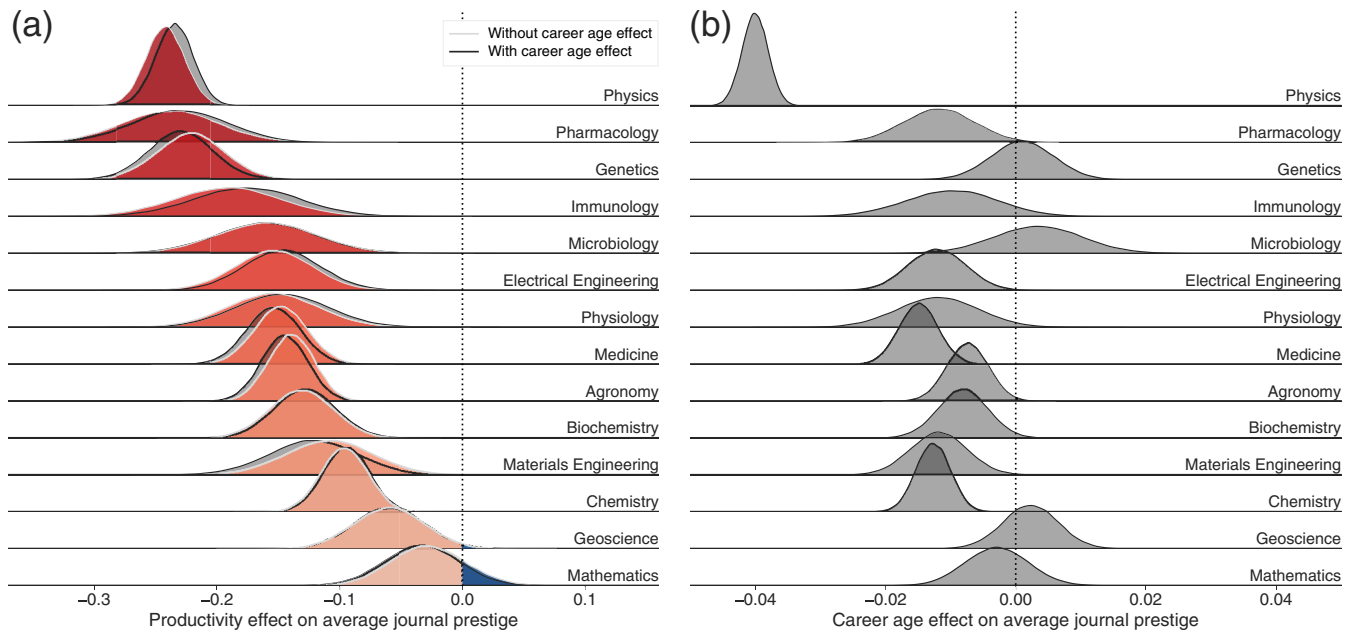


FIG. 4. Effect of productivity on journal prestige for nonoutlier researchers. (a) Posterior probability distributions of the average value of the linear coefficient ( $\mu_P$ ) when considering the association between productivity and journal impact for nonoutlier researchers of each discipline. The colored-filled curves represent the results without accounting for the effects of career age, while the gray-filled curves show the distributions of  $\mu_P$  after including career age as a confounding factor in the hierarchical Bayesian model (see the Methods Sec. IV D for details). (b) Posterior probability distributions of the average value of the linear coefficient ( $\mu_A$ ) related to the effect of career age on journal impact for nonoutlier researchers of each discipline.

trend in average productivity for all disciplines observed in Fig. 2.

Conversely, it is intriguing to observe that sector  $I++$  tends to be more populated during the initial stages of researchers' careers—a result that partially explains the slightly larger average journal prestige during first career years for most disciplines observed in Fig. 2. This behavior not only indicates that it is more likely to become an impact outlier in initial career years, but also that younger researchers (those having shorter career paths) may outperform more often in this category. Indeed, among the outlier researchers, the chance of finding perfectionist researchers decreases from 79% to 58% when career length increases from 10 to 30 years (Fig. S12A [45]). It is worth mentioning this trend of exhibiting high journal prestige in initial career stages may also reflect a selection effect as our data set only includes researchers belonging to the scientific elite of Brazil. Results for the SJR data set corroborate this finding (Fig. S12B [45]) and indicate very similar trends not only for disciplines present in both data sets but also for disciplines exclusive of the SJR data set.

#### D. Quantifying the effect of productivity on journal prestige

While our findings indicate a negative association between productivity and journal prestige at very high levels of both quantities for most researchers, we have not yet explored this relationship for researchers who never accessed outlier sectors. These nonoutlier academics represent 70% of the researchers in our data set and may exhibit heterogeneous behaviors, limiting the emergence of a clear aggregated relationship at discipline level. To account for these individual

behaviors, we select only the productive years of nonoutlier researchers with careers longer than 5 years (see Tables S1 and S2 [45] for details regarding this data set) and use a hierarchical Bayesian model (see the Methods Sec. IV D for details) for probing the association between productivity and average journal prestige. We assume a linear relationship between journal prestige and productivity, where the distribution of the linear coefficient related to each researcher has a mean drawn from another distribution with average value  $\mu_P$ .

By fitting this model to data with the Bayesian approach, we estimate the posterior probability distribution of the linear coefficient of each researcher and the posterior distribution of  $\mu_P$  for each area. Thus, the distribution of  $\mu_P$  represents the aggregated effect of productivity on journal impact for nonoutlier researchers in each discipline. Distributions of  $\mu_P$  shifted toward positive values represent disciplines where most researchers display a positive association between productivity and journal impact. In contrast, distributions more concentrated in negative values characterize disciplines where an increase in productivity correlates with a decline in journal impact for most researchers.

Figure 4(a) shows that the distribution of  $\mu_P$  (colored-filled curves) varies significantly among disciplines. All disciplines but mathematics have distributions entirely located in values of  $\mu_P$  lower than zero, suggesting an overall negative association between productivity and average journal impact for most nonoutlier researchers. In the most extreme case, a rise in one unit in the productivity of physicists associates with  $\approx 0.242$  decrease in average journal impact of their publications (in  $z$ -score units). On the other extreme, we have mathematics with distribution located near zero. This result indicates that

productivity usually plays a small role on journal impact for most mathematicians, while some may display more intense associations (positive or negative).

The results of Figs. 2 and 3 have already shown that career age affects the average productivity and journal prestige when aggregating researchers by their respective disciplines. Thus, we can also expect career age to affect the association between journal prestige and productivity at individual level. This is a critical aspect as the overall negative association reported in Fig. 4(a) may reflect a change from an early-career stage marked by low productivity and high impact to higher productivity and lower impact over the years.

To account for the possible confounding effect of career age on the association between journal prestige and productivity, we have included career age as a predictor of journal impact in the linear hierarchical Bayesian model. In this case, the distribution of the linear coefficient related to the effect of career age for each researcher has a mean drawn from another distribution with average value  $\mu_A$  (see the Methods Sec. IV D for details). Figure 4(b) shows that the distributions of  $\mu_A$  also vary among disciplines with most having negative or close to zero average values. These results indicate a reduction in the average journal impact over career years for most researchers from most disciplines. While it is hard to directly compare the effects of changing productivity with the effects of career progression, a 10-year career progression has more effect on journal prestige than increasing one unit of productivity ( $z$  score) of a typical researcher only for chemistry and physics (Fig. S13 [45]). Most importantly, Fig. 4(a) shows that the distributions of  $\mu_P$  with (colored-filled curves) and without (gray-filled curves) the career age effect change very little. Thus, the confounding effect of career age on the overall negative association between journal prestige and productivity is almost negligible—that is, an increase in productivity associates with a decrease in journal prestige regardless of career age.

The SJR data set (Figs. S14 and S15 [45]) extends this analysis for more disciplines and yields similar results for disciplines present in both data sets.

### III. DISCUSSION

We have investigated the association between yearly scientific productivity and average journal impact for more than 6000 top Brazilian researchers. Our results explore this association across disciplines and career stages and distinguish researchers with outlier and nonoutlier performances. Unlike previous works on the subject, our findings explicitly account for temporal inflation of the bibliometric indicators, scale-dependent effect on average journal prestige, and discipline-specific publication practices via robust standard score measures. These procedures have allowed us to build the journal prestige versus productivity plane—a coherent and straightforward aggregated representation of researchers' performances in productivity and journal impact. From this representation, we have categorized researchers into outliers and nonoutliers and further divided outliers into three categories: hyperprolific (outlier only in productivity), perfectionist (outlier only in journal impact), and hyperprolific

perfectionist (simultaneously outlier in productivity and journal impact).

Researchers with outlier performance comprise 30% of total scholars in our data set, and the most common behavior is performing as an outlier in only one career year (47.6% of cases). Among the outliers, the vast majority of researchers are exclusively hyperprolific or exclusively perfectionist. Despite that, 16 extremely hyperprolific researchers display career years solely in sector  $IP++$  when performing above a productivity threshold of  $P > 27.7$ . Only 14.4% of outlier researchers manage to be hyperprolific and perfectionist over their careers, and solely 6.7% simultaneously outperform in both categories (the hyperprolific perfectionists). This former group of 14.4% of outlier researchers (261 individuals) does not have a preferential outlier sector, displays productivity levels higher than exclusively hyperprolific and perfectionist scholars, and publishes in journals of higher prestige in comparison with exclusively hyperprolific and perfectionist researchers. Furthermore, we find that an increase in the number of hyperprolific career years reduces the probability of performing as a perfectionist for researchers who nonsimultaneously outperform in both categories for all disciplines except materials engineering in our data set. This negative association varies among disciplines, with mathematics presenting the most negative effect and physics the blandest effect. Together, these findings corroborate a negative association between productivity and journal prestige at outlier levels of both quantities. It is extremely hard for researchers to maintain overly high productivity while frequently publishing in very prestigious journals.

We have also explored short-term career patterns regarding productivity and journal impact. To do so, we have estimated the excess of transitions among sectors of the journal prestige versus productivity plane during consecutive career years of outlier and nonoutlier researchers. We have identified a persistent behavior in which researchers tend to stay within the same sector of the plane and thus display similar performance over consecutive years. Transitions among similar levels of productivity and journal prestige are about as frequent as by chance. Conversely, transitions among plane sectors with different productivity and journal impact levels occur much less often than by chance, indicating that researchers are averse to simultaneously changing their productivity and journal impact levels over consecutive career years.

We believe this aversion to simultaneously changing productivity and journal impact and the persistence in maintaining similar performances regarding these metrics indicate possible research and publication strategies in which researchers opt between productivity-focused or journal-impact-focused strategies [35]. To keep productivity levels, scholars may choose strategies based on expanding collaborations, avoiding prestigious journals, producing bite-sized articles, and selecting more traditional research themes. Conversely, impact-focused strategies may rely on searching for collaborators only when necessary and beneficial to solving research tasks, selecting high-impact journals as the top choice, publishing findings with maximization of understanding and impact in mind, and choosing novel research fields. Further research is needed for explicitly identifying these strategies. Still, our results suggest that publication strategies



may persist as a habit, and they possibly reflect individual characteristics and cultural conventions of research groups.

We have investigated the aggregated effect of career age on the average journal prestige and productivity for all disciplines. We have identified that journal prestige is usually slightly larger in initial career stages with a subtle downward trend over career years for most disciplines. Productivity, in its turn, tends to increase over career years for all disciplines. We have also studied the effect of career age on the occupation of sectors of the journal prestige versus productivity plane for each discipline. Our findings indicate that disciplines have distinct occupation fractions of these sectors, reflecting the different publication practices of different fields. However, we have found low productivity sectors ( $I-P-$  or  $I+P-$ ) to be more populated during initial stages of researchers' careers from all disciplines. We have also identified an increasing occupation trend of high productivity sectors—including the hyperprolific sector ( $P++$ )—in later career stages for practically all disciplines. Conversely, researchers more often achieve perfectionist performances in early career stages. It is important to remark that both the trend of presenting larger journal prestige in initial career years and the higher probability of finding researchers occupying sector  $I++$  in yearly career stages may reflect a selection effect, as all researchers in our data set belong to the Brazilian scientific elite. Whether these trends would also hold for other academics is an interesting question that future research could address. The increase in productivity with career age was also verified by Sinatra *et al.* [47] and may reflect a series of achievements that tend to be usual in scientific career progression, such as more familiarity with research themes [24], larger availability of financial resources [24,50], and invitation to write review articles [24]. Similarly, the emergence of hyperprolific years in later career stages may coincide with achieving higher positions in research centers, which could overly enhance publication rates by the tradition of some research disciplines (such as in medical and life sciences) of including the head of scientific labs in all publications [51].

Our results have also shown that the relation between productivity and journal impact for nonoutlier researchers is similar to the one observed for those achieving outlier performance. For nonoutlier, we have used a Bayesian hierarchical model that accounts for researchers' heterogeneous behaviors and identifies the emergent pattern for each discipline. We have found an overall negative association for the majority of disciplines when considering only nonoutlier researchers—a result that is in line with the negative association observed at outlier levels of productivity. However, the intensity of the association varies among disciplines, with physics having the most negative association and mathematics having the blandest effect of productivity on journal prestige. We have verified that while career age is also negatively correlated with journal impact, the overall negative association between journal impact and productivity is not significantly affected by this confounding factor. These findings contradict the Nijstad *et al.* “dual pathway to creativity model” [28], which states that creativity—as perceived as novel and suitable ideas—can be achieved through flexibility (usage of a variety of ideas to generate new ones) and persistence (exploration of the same subject in depth) pathways. According to this theory,

researchers with high productivity should be either exploring and associating various themes, enabling the generation of creative ideas in the flexibility pathway, or intensively working and publishing on the same theme until creative ideas are generated in the persistence pathway. In this sense, since productivity does not positively correlate with journal prestige, the JIF and SJR may not be the most suitable indicators for evaluating creativity.

## IV. METHODS

### A. Data

The data used in our study were obtained from the Lattes platform [52] (*Plataforma Lattes*). This information system has been maintained by the Brazilian government since 1999 and hosts the official curricula vitae (CV) of academics in Brazil. The Lattes CV is widely used for individual and institutional evaluations, and researchers are required to keep their records up to date. We have initially selected all 14 487 Brazilian researchers (from 88 disciplines) holding the CNPq Research Productivity Fellowship (as of May 2017) and obtained their complete publication records (1 121 652 articles). We filtered out researchers whose CVs were not updated from 1 Jan 2016 and those having no information about discipline and Ph.D. conclusion date, reducing the number to 14 146 researchers. We further filled in missing information about publication year and journal by using the DOI reference with the CrossRef API.

To define the journal prestige of these publications, we have obtained the journal impact factor (JIF) for all available scientific journals between 1997 and 2015 from Clarivate's journal citation reports. We thus combined these two data sets to assign the time-varying values of JIF to the articles published by the CNPq fellows. For each of these researchers, we calculated the number of articles published by year (productivity) and the average JIF of these publications (average journal prestige). Finally, we grouped these time series by discipline and selected the 14 disciplines having at least 50 researchers with published articles in each year between 1997 and 2015. This process led to our final data set comprising 6 028 researchers from 14 disciplines and 312 881 articles (Fig. S16 [45]). We have also considered the Scopus's SCImago journal rank (SJR) as a measure of journal prestige. To do so, we have obtained the SJR values for all available scientific journals between 1999 to 2015 from Scopus. By following the same approach used for the JIF, we obtain the productivity and the average SJR for 448 959 articles published by 8 465 researchers from 25 disciplines (Fig. S17 [45]).

While the JIF of a journal is simply defined as the number of citations received by articles from the two preceding years divided by the number of published articles in these two previous years [53,54], the SJR is a more complex network-based (an eigenfactor variant of the Page Rank algorithm) indicator [55]. Despite this difference, JIF and SJR are strongly correlated with Pearson correlation  $\approx 0.85$  (Fig. S1 [45]). The disciplines from both data sets are from science, technology, engineering, and mathematics (STEM disciplines), which in turn reflects the predominance of pro-

ductivity fellowship grants to researchers from these academic disciplines.

### B. Inflation and robust standard score units

The volume of scientific production has increased over time at global and individual levels [46,47]. This yields an inflation effect that prevents a direct comparison of raw productivity and journal impact values from different periods (Figs. S2 and S3 [45]). Disciplines also have distinct volumes of publication and citation dynamics [56,57], which in turn hampers the aggregation and comparison of raw productivity and journal impact values among different disciplines. Furthermore, average journal impact suffers from an additional size effect that decreases its variability with the rise of productivity. This size effect has been observed when comparing the impact factor of journals with different numbers of total publications [48,49] and, as argued by Antonoyiannakis [48,49], it represents a direct consequence of the central limit theorem.

To account for these issues, we have used  $z$ -score measures relative to year and discipline for productivity and  $z$ -score measures relative to year, discipline, and productivity level for journal prestige. Let  $p_j^k(y)$  and  $i_j^k(y)$  represent, respectively, the number of papers and the average journal prestige of the publications by researcher  $j$  from discipline  $k$  in year  $y$ . We calculate the  $z$  scores of productivity as

$$p_j^k(y) = \frac{p_j^k(y) - \mathbb{E}[p_j^k(y)]}{\mathbb{S}[p_j^k(y)]},$$

where  $\mathbb{E}[p_j^k(y)]$  and  $\mathbb{S}[p_j^k(y)]$  are (respectively) the average and the standard deviation of the productivity of researchers from discipline  $k$  in year  $y$ . Similarly, we calculate the  $z$ -score journal prestige as

$$i_j^k(y) = \frac{i_j^k(y) - \mathbb{E}[i_{\text{md}}^k(y, p_j^k(y))]}{\mathbb{S}[i_{\text{md}}^k(y, p_j^k(y))]},$$

where  $i_{\text{md}}^k(y, p)$  is the average journal impact of a random sample of  $p$  publications from discipline  $k$  in year  $y$ , and  $\mathbb{E}[i_{\text{md}}^k(y, p)]$  and  $\mathbb{S}[i_{\text{md}}^k(y, p)]$  represent, respectively, the average and the standard deviation of  $i_{\text{md}}^k(y, p)$  estimated over 1000 independent realizations. This definition is an adaptation of the  $\Phi$  index proposed by Antonoyiannakis [48,49] for ranking journals of different sizes, and it accounts for the fact that low-productivity researchers have high variability in their average values of journal prestige, while high-productivity researchers display significantly lower variability (Figs. S18 [45]).

Here we have further used Huber robust estimators for mean (location) and standard deviation (scale) in place of the usual estimators [58] due to the existence of outlier values for  $p_j^k(y)$  and  $i_j^k(y)$  (Fig. S19 [45])—that is,  $\mathbb{E}[\dots]$  and  $\mathbb{S}[\dots]$  represent respectively Huber's estimators of location and scale (as implemented in PYTHON package *statsmodels* [59]).

### C. Logistic regressions

To quantify the effect of performing as outlier in productivity on the chance of being an outlier in journal impact

(perfectionist), we use the following logistic model,

$$\Pi_{\text{perfectionist}} = \frac{e^{\alpha_0 + \alpha_1 Y_P}}{1 + e^{\alpha_0 + \alpha_1 Y_P}},$$

where  $\Pi_{\text{perfectionist}}$  is the probability of being a perfectionist researcher given the scholar has  $Y_P$  outlier career years in productivity,  $\alpha_0$  is the intercept, and  $\alpha_1$  is the logistic regression coefficient. Positive values of  $\alpha_1$  indicate that an increase in  $Y_P$  enhances the probability of performing as a perfectionist, while negative values of  $\alpha_1$  show that an increase in  $Y_P$  reduces the probability of being a perfectionist. We have adjusted this model (as implemented in the PYTHON package *statsmodels* [59]) to our data by considering all researchers who outperformed in journal impact or productivity at some point in their careers. We have also adjusted the same model when grouping researchers by discipline. Figure 1(c) shows  $\Pi_{\text{perfectionist}}$  as a function of  $Y_P$  for each discipline in the JIF data set and when considering all disciplines together (the inset depicts the values of  $\alpha_1$ ). Similarly, Fig. S4C [45] shows the corresponding results for the SJR data set.

We further use a similar logistic model to estimate the probability of being perfectionist as a function of the career length of researchers ( $L$ ). In this case, the model can be written as

$$\Pi_{\text{perfectionist}} = \frac{e^{\theta_0 + \theta_1 L}}{1 + e^{\theta_0 + \theta_1 L}},$$

where  $\theta_0$  (intercept) and  $\theta_1$  (regression coefficient) are the model parameters. We have adjusted this model considering all outlier researchers in JIF and SJR data sets. Figure S12 [45] shows  $\Pi_{\text{perfectionist}}$  as a function of  $L$  for both data sets. The adjusted parameters are  $\theta_0 = 1.849 \pm 0.132$  and  $\theta_1 = -0.051 \pm 0.006$  for the JIF data set, and  $\theta_0 = 1.921 \pm 0.108$  and  $\theta_1 = -0.054 \pm 0.005$  for the SJR data set.

### D. Bayesian hierarchical model

We use a Bayesian hierarchical model to estimate the effect of productivity on average journal impact for nonoutlier researchers. For a given discipline, we consider that the data are hierarchically structured such that each observation of  $I_j$  and  $P_j$  is nested within a researcher  $j$  (here we have dropped the index  $k$  for simplicity). We further assume a linear relation between these variables at the individual level, where  $c_j$  and  $\beta_j$  are (respectively) the intercept and the slope of the linear association for the  $j$ th researcher of a given discipline. We consider the parameters  $c_j$  and  $\beta_j$  as random variables distributed according to normal distributions whose parameters are also random variables. Mathematically, we can write this model as

$$I_j \sim \mathcal{N}(c_j + \beta_j P_j, \varepsilon), \quad (1)$$

where  $\mathcal{N}(\mu, \sigma)$  stands for a normal distribution with mean  $\mu$  and standard deviation  $\sigma$ ,  $\varepsilon$  accounts for the unobserved determinants of  $I_j$ , and

$$c_j \sim \mathcal{N}(\mu_c, \sigma_c), \\ \beta_j \sim \mathcal{N}(\mu_P, \sigma_P),$$

where  $\mu_c$  is the mean,  $\sigma_c$  is the standard deviation of a normal distribution associated with the intercept  $c_j$ , and  $\mu_P$  and  $\sigma_P$  are

the same for the distribution associated with  $\beta_j$ . The Bayesian inference process consists in determining the posterior probability distributions of the parameters at discipline level ( $\mu_c$ ,  $\sigma_c$ ,  $\mu_P$ , and  $\sigma_P$ ) and at researcher level ( $c_j$  and  $\beta_j$  for every  $j$  researcher of given discipline).

We perform this Bayesian regression for each area separately and use noninformative prior distributions [60] to not bias the posterior estimation; that is, we consider

$$\begin{aligned}\varepsilon &\sim \mathcal{U}(0, 10^2), \\ \mu_c &\sim \mathcal{N}(0, 10^5), \\ \mu_P &\sim \mathcal{N}(0, 10^5), \\ \sigma_c &\sim \text{Inv-}\Gamma(10^{-3}, 1), \\ \sigma_P &\sim \text{Inv-}\Gamma(10^{-3}, 1),\end{aligned}\quad (2)$$

where  $\mathcal{U}(x_{\min}, x_{\max})$  represents a uniform distribution between  $x_{\min}$  and  $x_{\max}$ , and  $\text{Inv-}\Gamma(a, b)$  stands for an inverse- $\gamma$  distribution with parameters  $a$  (shape) and  $b$  (scale). Figure S20 [45] shows a graphical representation of this model.

We have also considered a generalized version of the model defined in Eq. (1), where career age  $A_j$  is also assumed to be linearly related with average journal prestige. The value of  $A_j$  refers to the career age of researcher  $j$  at a given year  $y$  with productivity  $P_j$  and average journal impact  $I_j$ . Thus, we include the career age  $A_j$  as an independent variable in the hierarchical model of Eq. (1), yielding

$$I_j \sim \mathcal{N}(c_j + \beta_j P_j + \gamma_j A_j, \varepsilon), \quad (3)$$

where  $\gamma_j$  is the slope of the linear association between career age and journal prestige. This linear coefficient is assumed to be distributed according to a normal distribution

$$\gamma_j \sim \mathcal{N}(\mu_A, \sigma_A),$$

where  $\mu_A$  is the mean and  $\sigma_A$  is the standard deviation. We adjust the model of Eq. (3) with the same noninformative prior distributions defined in Eq. (2), and use

$$\begin{aligned}\mu_A &\sim \mathcal{N}(0, 10^5), \\ \sigma_A &\sim \text{Inv-}\Gamma(10^{-3}, 1)\end{aligned}\quad (4)$$

as the noninformative prior distributions for the additional parameters related to the effects of career age. Figure S21 [45] shows a graphical representation of this generalized model that accounts for possible confounding effects of career age on the association between average journal prestige and productivity.

We implement these two models [Eqs. (1) and (3)] using the PyMC3 framework [61] via gradient-based Hamiltonian Monte Carlo no-U-Turn-sampler method for sampling the posterior distributions. We run eight parallel chains with 10 000 iterations (of which 5000 are burn-in samples) to allow good mixing of the Monte Carlo chains. We estimate the Gelman-Rubin convergence statistic (R-hat) for all regression analyses and the results were all close to one, an indication of convergence of the sampling approach.

#### ACKNOWLEDGMENTS

A.S.S. and H.V.R. acknowledge the support of the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq Grants No. 407690/2018-2 and No. 303121/2018-1). M.P. acknowledges the support of the Slovenian Research Agency (Grants No. J1-2457 and No. P1-0403).

- 
- [1] A. Zeng, Z. Shen, J. Zhou, J. Wu, Y. Fan, Y. Wang, and H. E. Stanley, The science of science: From the perspective of complex systems, *Phys. Rep.* **714**, 1 (2017).
- [2] S. Fortunato, C. T. Bergstrom, K. Börner, J. A. Evans, D. Helbing, S. Milojević, A. M. Petersen, F. Radicchi, R. Sinatra, B. Uzzì, A. Vespignani, L. Waltman, D. Wang, and A.-L. Barabási, Science of science, *Science* **359**, eaa0185 (2018).
- [3] P. Azoulay, J. S. G. Zivin, and G. Manso, Incentives and creativity: Evidence from the academic life sciences, *RAND J. Econ.* **42**, 527 (2011).
- [4] L. Bromham, R. Dinnage, and X. Hua, Interdisciplinary research has consistently lower funding success, *Nature (London)* **534**, 684 (2016).
- [5] S. Meirmans, R. K. Butlin, A. Charmantier, J. Engelstädter, A. T. Groot, K. C. King, H. Kokko, J. M. Reid, and M. Neiman, Science policies: How should science funding be allocated? An evolutionary biologists' perspective, *J. Evol. Biol.* **32**, 754 (2019).
- [6] J. Wilsdon, *The Metric Tide: Independent Review of the Role of Metrics in Research Assessment and Management* (Sage, London, 2016).
- [7] S. Wessely, Peer review of grant applications: What do we know?, *Lancet* **352**, 301 (1998).
- [8] R. Smith, Peer review: A flawed process at the heart of science and journals, *J. R. Soc. Med.* **99**, 178 (2006).
- [9] S. Ballelli, R. L. Goldstone, and D. Helbing, Peer review and competition in the art exhibition game, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 8414 (2016).
- [10] L. Bornmann and R. Mutz, Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references, *J. Assoc. Inf. Sci. Technol.* **66**, 2215 (2015).
- [11] J. P. Ioannidis, K. W. Boyack, and R. Klavans, Estimates of the continuously publishing core in the scientific workforce, *PLoS ONE* **9**, e101698 (2014).
- [12] B. D. Cameron, Trends in the usage of ISI bibliometric data: Uses, abuses, and implications, *Portal: Libraries Acad.* **5**, 105 (2005).
- [13] V. A. Traag and L. Waltman, Systematic analysis of agreement between metrics and peer review in the UK REF, *Palgrave Commun.* **5**, 29 (2019).
- [14] M. Gagolewski, Scientific impact assessment cannot be fair, *J. Informetrics* **7**, 792 (2013).

- [15] G. Siudem, B. Żogała-Siudem, A. Cena, and M. Gagolewski, Three dimensions of scientific impact, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 13896 (2020).
- [16] San Francisco Declaration on Research Assessment, 2012, <https://sfidora.org/read/>.
- [17] D. Hicks, P. Wouters, L. Waltman, S. D. Rijcke, and I. Rafols, Bibliometrics: The leiden manifesto res. metrics, *Nature (London)* **520**, 429 (2015).
- [18] Nuffield Council on Bioethics, The findings of a series of engagement activities exploring the culture of scientific research in the UK, 2014, <https://www.nuffieldbioethics.org/assets/pdfs/The-culture-of-scientific-research-report.pdf>.
- [19] K. Powell, Young, talented, and fed-up: Scientists tell their stories, *Nature (London)* **538**, 446 (2016).
- [20] D. Moher, F. Naudet, I. A. Cristea, F. Miedema, J. P. Ioannidis, and S. N. Goodman, Assessing scientists for hiring, promotion, and tenure, *PLoS Biol.* **16**, e2004089 (2018).
- [21] L. A. Schimanski and J. P. Alperin, The evaluation of scholarship in academic promotion and tenure processes: Past, present, and future, *F1000Res.* **7**, 1 (2018).
- [22] W. Dennis, Productivity among American psychologists, *Am. Psychologist* **9**, 191 (1954).
- [23] K. G. White and M. J. White, On the relation between productivity and impact, *Austr. Psychologist* **13**, 369 (1978).
- [24] S. Lawani, Some bibliometric correlates of quality in scientific research, *Scientometrics* **9**, 13 (1986).
- [25] D. K. Simonton, *Scientific Genius: A Psychology of Science* (Cambridge University Press, Cambridge, UK, 1988).
- [26] G. J. Feist, Quantity, quality, and depth of research as influences on scientific eminence: Is quantity most important?, *Creativity Res. J.* **10**, 325 (1997).
- [27] N. Haslam and S. M. Laham, Quality, quantity, and impact in academic publication, *Eur. J. Social Psych.* **40**, 216 (2010).
- [28] B. A. Nijstad, C. K. D. Dreu, E. F. Rietzschel, and M. Baas, The dual pathway to creativity model: Creative ideation as a function of flexibility and persistence, *Eur. Rev. Social Psych.* **21**, 34 (2010).
- [29] C. Bosquet and P.-P. Combes, Are academics who publish more also more cited? Individual determinants of publication and citation records, *Scientometrics* **97**, 831 (2013).
- [30] G. Abramo, T. Cicero, and C. A. D'Angelo, Are the authors of highly cited articles also the most productive ones?, *J. Informetrics* **8**, 89 (2014).
- [31] U. Sandström and P. van den Besselaar, Quantity and/or quality? The importance of publishing many papers, *PLoS ONE* **11**, e0166149 (2016).
- [32] V. Larivière and R. Costas, How many is too many? On the relationship between research productivity and impact, *PLoS ONE* **11**, e0162709 (2016).
- [33] V. Garousi and J. M. Fernandes, Quantity versus impact of software engineering papers: A quantitative study, *Scientometrics* **112**, 963 (2017).
- [34] M. J. Michalska-Smith and S. Allesina, And, not or: Quality, quantity in scientific publishing, *PLoS ONE* **12**, e0178074 (2017).
- [35] S. Kolesnikov, E. Fukumoto, and B. Bozeman, Researchers' risk-smoothing publication strategies: Is productivity the enemy of impact?, *Scientometrics* **116**, 1995 (2018).
- [36] L. Bornmann and A. Tekles, Productivity does not equal usefulness, *Scientometrics* **118**, 705 (2019).
- [37] B. Forthmann, M. Leveling, Y. Dong, and D. Dumas, Investigating the quantity-quality relationship in scientific creativity: An empirical examination of expected residual variance and the tilted funnel hypothesis, *Scientometrics* **124**, 2497 (2020).
- [38] V. Larivière and C. R. Sugimoto, *The Journal Impact Factor: A Brief History, Critique, and Discussion of Adverse Effects* (Springer, Berlin, 2019), pp. 3–24.
- [39] E. C. McKiernan, L. A. Schimanski, C. M. Nieves, L. Matthias, M. T. Niles, and J. P. Alperin, Meta-research: Use of the journal impact factor in academic review, promotion, and tenure evaluations, *eLife* **8**, e47338 (2019).
- [40] L. Bornmann and L. Leydesdorff, Skewness of citation impact data and covariates of citation distributions: A large-scale empirical analysis based on Web of Science data, *J. Informetrics* **11**, 164 (2017).
- [41] V. A. Traag, Inferring the causal effect of journals on citations, *Quant. Sci. Studies* **2**, 496 (2021).
- [42] L. Kim, J. H. Portenoy, J. D. West, and K. W. Stovel, Scientific journals still matter in the era of academic search engines and preprint archives, *J. Assoc. Inf. Sci. Technol.* **71**, 1218 (2020).
- [43] J. C. Correa, H. Laverde-Rojas, J. Tejada, and F. Marmolejo-Ramos, The Sci-Hub effect on papers' citations, *Scientometrics* (2021).
- [44] L. Waltman and V. A. Traag, Use of the journal impact factor for assessing individual articles need not be statistically wrong, *F1000Res.* **9**, 366 (2020).
- [45] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevResearch.3.033158> for Figs. S1–S21 and Tables S1 and S2.
- [46] D. J. de Solla Price, *Little Science, Big Science* (Columbia University Press, New York, 1963).
- [47] R. Sinatra, D. Wang, P. Deville, C. Song, and A.-L. Barabási, Quantifying the evolution of individual scientific impact, *Science* **354**, aaf5239 (2016).
- [48] M. Antonoyiannakis, Impact factors and the central limit theorem: Why citation averages are scale dependent, *J. Informetrics* **12**, 1072 (2018).
- [49] M. Antonoyiannakis, Impact factor volatility due to a single paper: A comprehensive analysis, *Quant. Sci. Studies* **1**, 639 (2020).
- [50] D. Hicks, Performance-based university research funding systems, *Res. Policy* **41**, 251 (2012).
- [51] M. Price, Some scientists publish more than 70 papers a year. Here's how—and why—they do it, *Science* (2018).
- [52] <http://lattes.cnpq.br/>.
- [53] E. Garfield, Citation analysis as a tool in journal evaluation: Journals can be ranked by frequency and impact of citations for science policy studies, *Science* **178**, 471 (1972).
- [54] E. Garfield, The history and meaning of the journal impact factor, *JAMA* **295**, 90 (2006).
- [55] V. P. Guerrero-Bote and F. Moya-Anegón, A further step forward in measuring journals' scientific prestige: The SJR2 indicator, *J. Informetrics* **6**, 674 (2012).
- [56] M. Bordons, M. Fernández, and I. Gómez, Advantages and limitations in the use of impact factor measures for the assessment of research performance, *Scientometrics* **53**, 195 (2002).

- [57] J. G. Foster, A. Rzhetsky, and J. A. Evans, Tradition and innovation in scientists' research strategies, *Am. Sociol. Rev.* **80**, 875 (2015).
- [58] P. J. Huber, *Robust Statistics* (Wiley, New York, 2004).
- [59] StatsModels, 2020, <https://www.statsmodels.org>.
- [60] A. Gelman, Prior distributions for variance parameters in hierarchical models, *Bayesian Anal.* **1**, 515 (2006).
- [61] J. Salvatier, T. V. Wiecki, and C. Fonnesbeck, Probabilistic programming in PYTHON using PyMC3, *Peer J. Comp. Sci.* **2**, e55 (2016).