Scientometrics (2019) 120:793–805
https://doi.org/10.1007/s11192-019-03135-7

The strong nonlinear effect in academic dropout

Yanmeng Xing[1] · An Zeng[1] · Ying Fan[1] · Zengru Di[1]

Received: 24 January 2019 / Published online: 30 May 2019

Abstract
Survivability is one of the features for success in contemporary science ecosystem. In this paper, we analyze the publication records of physicists in American Physical Society journals, aiming to identify the career length of each researcher and accordingly investigate the dropout phenomenon in science by the example of physicists. We find that scientific career is a complex nonlinear evolution process and can be generally divided into four stages regarding the dropout rate. In the early career, the dropout rate from trainee phase to maturity is high and negatively correlated with the research performance of the scientists, in both productivity and impact. Moreover, a strong nonlinearity is observed when we study the detailed relationship between the dropout rate and research performance. Interestingly, in the more mature stage of the career, the dropout rate becomes stable and independent of the early performance of the scientists. In the late career stage, the dropout rate increases and is mainly determined by retirement and external factors. The findings in this paper may provide useful guidance for young scholars to allocate their research effort in the early career.

Keywords Academic career · Dropout rate · Nonlinearity

Introduction

The blockbuster expansion of large-scale digital data on the production of science—from productivity, collaboration to paper citations and individual academic trajectory—boosts unprecedented opportunities to systematically study science of science (SciSci) in last decade (Zeng et al. 2017; Fortunato et al. 2018). Large effort has been devoted to quantify the scientific influence of scientists, the evolution of scientists' creativity and discuss the key factors of career success. Yet much less attention has been paid to negative events in scientific research such as career dropout and death phenomenon (i.e. the researcher donot publish academic papers any more in any journals). Universities and research institutes are currently training

✉ An Zeng
anzeng@bnu.edu.cn

✉ Ying Fan
yfan@bnu.edu.cn

1 School of Systems Science, Beijing Normal University, Beijing 100875, People's Republic of China

more and more Ph.D.'s in most of science fields. However, many of them will eventually drop out academia due to different levels of competition or other various reasons (Editorial 2017). Most junior scholars who are attracted by academic calling and ideals and want a career in academic research may find that career success is not so easy and strongly associated with whether they outperform than others or not (Zeng et al. 2017; Levecque et al. 2017). In this context, the research of academic dropout behavior is of great significance. It will not only enable a new dimension for theoretically studying academic career, but also offer guidance for young scholars' career development in practice.

Research on career progression and uncertainty among academia are critical for scholars at a series of academic stages, especially for junior researchers. In recent years, there has been a disconnection between scientific progress and career progress in academia. A structural change of academic workforce space has taken place: there are much more Ph.D.'s trained than the available faculty positions (Afonso 2016; Merton 1968). It is also increasingly difficult for new scientists to establish independent research careers (Sinatra et al. 2016). As a result, many early-career academic researchers can no longer find a tenure-track research position but instead pursuing a temporary short-term and flexible contracts although without job security in perpetual position (Cyranoski et al. 2011). A research shows that the reseachers under short-term appraisal system face more competition and uncertainty. And they face bigger career danger of "sudden" early deaths because of productivity fluctuations not necessarily due to lack of individual brilliance and persistence (Stumpf 1995). In contrast, rewarding long-term academic success promotes scholars to produce high-impact works at a much higher rate than grants subject to short review cycles (Musch and Grondin 2001). Without doubt that the outflow of trainee academics who are likely to be our next generation of leaders in scientific discovery, is a serious loss for the progress of science.

In this paper, we focus on the dropout phenomenon in the case of the physicists' academic careers and their relationships to different research performance factors. Firstly, we compute the dropout rate in different career stages by using the publication data over a century from American Physical Society (APS) journals. We find that physicists' careers are generally divided into 4 periods. Also, we analyse dropout characteristics of different career stages. Next, we pay more attention to junior scholars' career uncertainty and consider the influence of quantity and quality of early publication on it. We reveal that the relationship between early publications and dropout rate is highly nonlinear. Finally, we explore the relationship between late dropout rate and early academic performances. Interestingly, we find evidence showing that the dropout rate in the later career stage is almost independent of the early research performance.

The paper is arranged as follows. The first section is the introduction of the background of our work. The second section presents the related works of scientific career factors and the impact of early related performance on long-term academic career. The third section is the concrete methods of data collection and analysis. The major results display of the nonlinear effect on academic dropout are in forth section. Next, we will present some discussions of above results. Then, the applicability and lessons learned from our analysis will be made in sixth section. The last Section is conclusions and future works we intend to study in depth.

Related works

Individual academic careers have been started and developed under a broad market not only for science production but also complex consumption. Depending on interdisciplinary data-driven effort from related fields such as natural, computational, and social science, many important scientific career factors have been captured based on different big data sets from different science fields. The balance between scientific work and personal life is not notably related to the reason for leaving the career pathway (Yoon et al. 2018). There exist gender differences on academic career-choice risk in publication rate and impact. And science faculty's subtle gender biases result in male students prevalence (Duch et al. 2012; Moss-Racusin et al. 2012). Systematic hierarchy and inequality in science faculty hiring networks have been revealed recently (Clauset 2015; Deville et al. 2015). It has already been found that cumulative advantage of scientists in their careers usually described as reputation is a critical driver for early sucecss (Petersen et al. 2014). Age at career onset could account for individual differences in career trajectories (Simonton 1997; Petersen et al. 2012). A good faculty affiliation and a stable position in professional faculties relieve career anxiety and indecision (Daniels et al. 2011; Weinstein et al. 2002). However, uncertainty in young adulthood occupational expectations results in higher drop-out rates and lower occupational attainment (Seifert 1994; Sikora 2018). Tolerance of career uncertainty plays a great role in career satisfaction and happenstance skills (Kim et al. 2016). Young scholars are thus advised to be aware of these scenarios and prepare for their careers before entering the first position of academia.

Early academic performance, vocational interests and career preferences are key factors for understanding major persistence behavior and academic commitment which lead to academic career-decision readiness (Allen and Robbins 2008; Liao and Ji 2015). The pressures on academic career progress prioritise researchers' early science productivity and impact (Mills and Paulson 2014). The career development tasks of the exploratory stage have an impact on ultimate career stages (Seifert 1994). For biologists, publishing in early career has a significant advantage for future career survival rate (Afonso 2016). Meanwhile, early approach to prestigious central institutions offers more chances and greater advantages access to prestigious venues and reduces dropout rate in art (Fraiberger et al. 2018). A study about biochemists on position in the academic stratification system reveals that preemployment productivity has an insignificant effect on the prestige of the scientist's first academic position (Scott et al. 1979). Moreover, quality of academic outputs overweights quantity in eliciting recognition through appointment to important academic departments (Cole and Cole 1967). A most recent reserach based on three scientific disciplines shows that neither productivity nor the citation impact of early publications can dramatically influence terminal survival status (Milojević et al. 2018).

Despite above various progress, we still find a clear gap in general knowledge of scientific career pattern for researchers in various career states (not just the extremely prolific scientists). Furthermore, quantitative understanding of factors related to transition from trainee phase to maturity is underemphasized. The original junior researchers have received insufficient attention. However, they are the most important building blocks of science. In addition, based on large-scale data sets, studies of the drupout trends in academia over long periods of time is still unclear. This paper focuses on the entire career pattern of researchers as long as he published the first article with big data set over a century from American Physical Society (APS) journals. And we concentrate on the analysis of the significant factors for surviving in early academic workspace to stable phase.

Methods

Data collection

Our dataset is cleaned and provided by Barabasi et al. (www.magnus.net) and combines information on scientists' research agendas as their careers progress. In detail, the database used in this paper is from the American Physical Society (APS) journals in the period from 1893 to 2009. APS publications serve the international physics community with peer-reviewed research journals, news and commentary about the latest research published in the Physical Review journals, information about physics and its place in the world. The Physical Review journal collection of 14 leading peer-reviewed research journals includes Physical Review Letters, Physical Review X, Physical Review, and Reviews of Modern Physics. In more recent years, APS is active in the international physics community. APS divisions and topical groups cover all areas of physics research. Sections are organized by geographical region. International issues cut across essentially all aspects of the APS and physics community in this era of globalization. In the data set used in this article, 37% of the researchers come from the US. The rest of them come from 125 different countries all over the world. For each author, we can obtain his/her name and all the papers he/she has published. For each paper, the information about DOI, authors, publication date and the DOI of citing papers are all available in the database. It offers information on 458,584 papers with 4,620,025 citations from 325,491 authors, allowing us to systematically analyze scholars' academic career trajectory.

Quantifying researchers' scientific performance and dropout rate

We first briefly introduce the definitions and notations in this paper. We define the year when an author published his/her first paper as the beginning of his/her scientific career (Deville et al. 2014). For each author i, his/her career length of scientific research is denoted as cl, defined as the year of the last publication subtracts the year of the first publication. Researchers who drop out from scientific research, in our definition, are those researchers without any article published after a certain year in our data set. They are denoted as dropouts. To quantify the drop out phenomenon, we define the dropout rate as $\mathrm{DR} = N_q/N$, where N_q is the number of the dropouts and N is the number of the total scientific researchers. In addition, to study the factors that affect the dropout rate, we will focus on the quantity and quality of their publications, respectively denoted as *Pubs* and C_{10}, the number of citations 10 years after publication (Sinatra et al. 2016). The *Pubs* is the number of an author's publications in APS. In detail, we calculate C_{10} by two means of C_{10}^{mean} (i.e. the average C_{10} of authors' publications during a certain period) and C_{10}^{max} (i.e. the max C_{10} of authors' publications during a certain period).

To assign citations of a paper to its authors, there are two main statistical methods in scientometrics: full counting and fractional counting. While ignoring the impact of the number of authors, full counting is the most popular in practical calculation where each author is allocated the equal total credit for the article. In fractional counting, the total credit for a co-authored article is the same as a single author article, but all the collaborative authors share equal fractions of the total credit for this paper. For simplicity, we assume that each author makes the same contribution to artical and show the results calculated in the full counting system in this analysis.

The problem addressed in this paper is the analysis of researchers' dropout phnomenon and the critical factors which underlie activity in academia, which is straightforwardly illustrated in Fig. 1. We first compute the starting year of individual researcher's career by identifying her first published paper, and then track her later career trajectory to collect the information of her career length. We find that the fraction of dropouts varies significantly in different career stages, and the dropout rate is connected to scientists' research performance. Our questions are what the scientific career dropout patterns are (i.e. how does the dropout rate vary with career length) and how to maintain scientific career security (i.e. how to avoid scientific career uncertainties), especiallly for junior researchers.

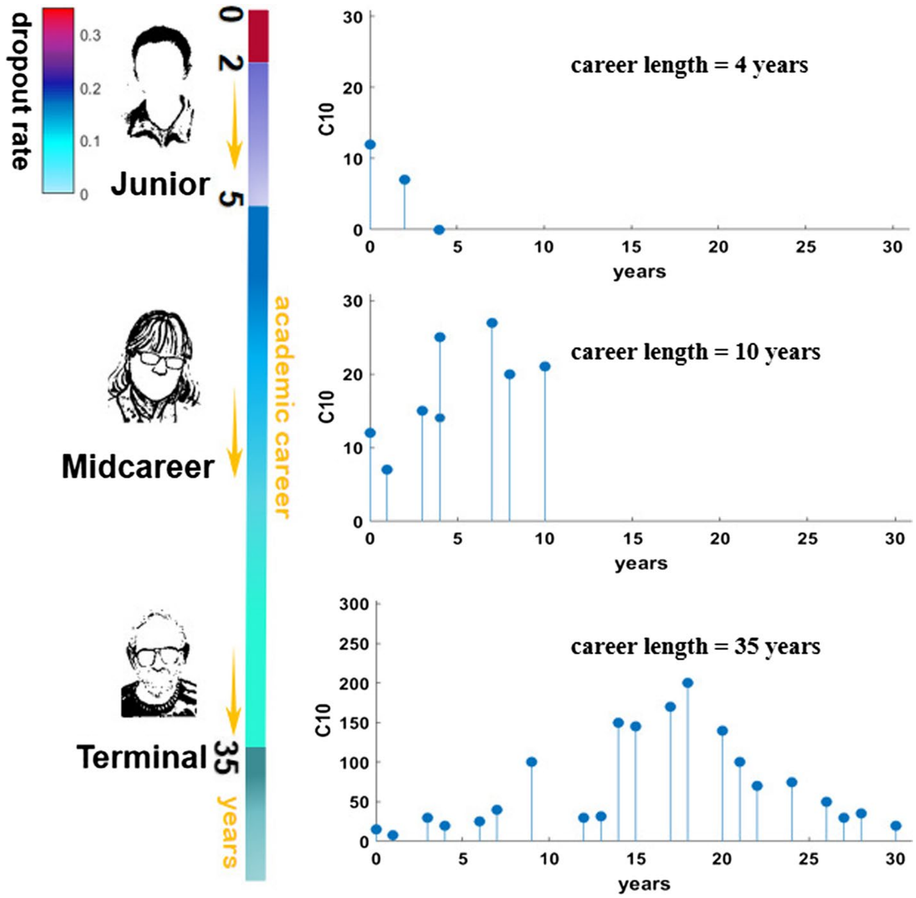

Fig. 1 (Color online) The illustration of the dropout phenonmen in academic careers. The figure shows three sample individual schloars with different academic performances and career length. Respectivly, they stand for junior droputs, midcareer dropouts and finally, the tenure scientists who almost achieved full careers. The colorbar indicates the variables of dropout rate. This paper aims to explore the genenral pattern of academic dropout phenonmen and the relationship between scientific performances (i.e. quantity and impact (C_{10}) of publications) and career length

Results

Using articles published by American Physical Society (APS) journals covering over 117 years (1893–2009), we collect publication records of individual authors who published more than one paper in the dataset. After a careful and extensive author-name disambiguation process, the data totally includes 458,584 papers with 4,620,025 citations from 325,491 authors. Firstly, in order to display a true and comprehensive scenario of academic career length, we only choose all researchers with a full career who published their first paper during 1893–1973 years in this section. We consider this group of scientists because, we are sure that they have already ended their academic career. In this case, the detected dropout is a real one instead of a temporal leave. We cannot see the true dropout rule of the career ending if we use the data after 1973. This is because we lack further publication recordings to calculate their academic career length. The result can display complete academic lifecycle and individual career evolution. Figure 2a shows that the cl for a researcher follows a exponential cutoff fat-tailed distribution where 40% of the researchers has one recorded paper, a few high-profile scholars have a superbly lengthy career. We observe universal statistical regularity corresponding to $k \approx 1.53$ in the career longevity cumulative distribution. The disparity in career lengths indicates that more than half of

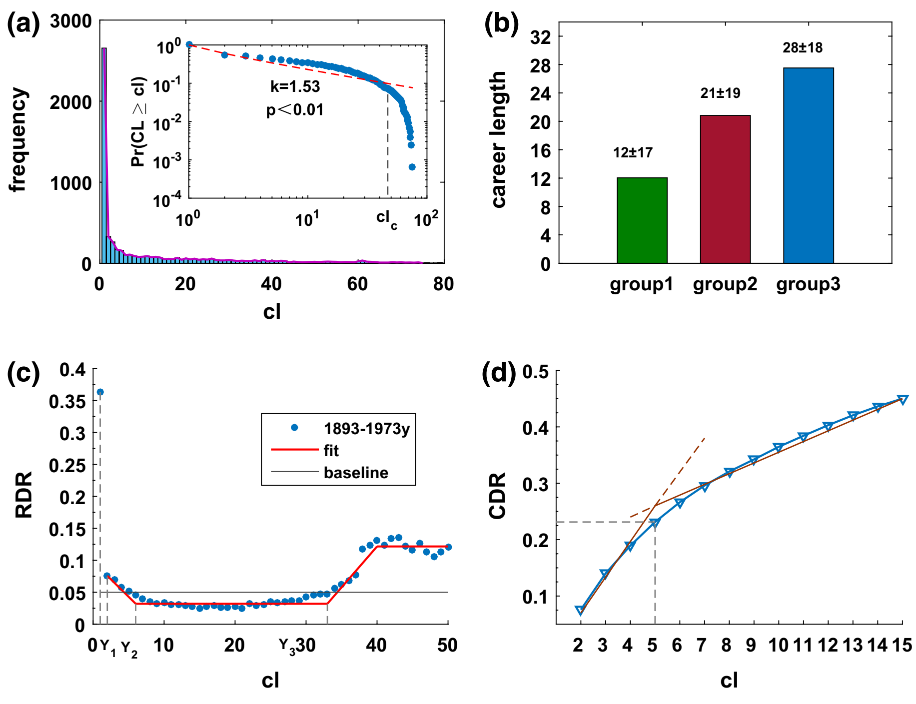

Fig. 2 (Color online) The characters of academic career length and dropout rate of researchers published their first paper in 1893–1973. **a** the career length cumulative distribution followed an exponential cutoff fat-tailed distribution. **b** The average academic career length of APS researchers divided into 3 groups: (i) individuals whose cl ≥ 1, (ii) individuals whose cl ≥ 2, (iii) individuals who have longer career longevity more than 5 years. **c** the trend of relative dropout rate (RDR). **d** The trend of cumulative absolute dropout rate (CDR)

individuals end their career with a relatively short time interval and even just making their debut. For example, we find that roughly 52% of researchers only have a career length less than 6 years. While the relatively high fraction of short careers, we can see some examples of extremely long careers as well, corresponding to nearly the full academic lifecycle of the individual. Specifically, nearly 14 percent of researchers has an academic career longer than 35 years. In short, the exponential cutoff in cl that follows after the crossover value cl_c arising from the finite human lifetime. Because finite-size effects dominate the asymptotic behavior (Petersen et al. 2011). We then investigate the average academic career length of researchers in the APS data. We divide the scholars "born" (i.e. published the first paper) in 1893–1973 years who at least published one paper into three types. Then we compute the mean value and standard deviaton of cl respectively for different groups in Fig. 2b: (i) individuals whose $\mathrm{cl} \geq 1$, resulting the career length of the selected scientists as $\langle \mathrm{cl} \rangle = 12 \pm 17$. (ii) individuals whose $\mathrm{cl} \geq 2$, resulting the career length of the selected scientists as $\langle \mathrm{cl} \rangle = 21 \pm 19$. (iii) individuals who have career longevity more than 5 years with $\langle \mathrm{cl} \rangle = 28 \pm 18$. One can see that the group 3's average cl is about 28 years and much longer than that of the group 1's. We assume that it stands for the mean career longevity of researchers who had gain tenured faculty before 1973s.

Then we focus on the dropout rate of academic career in the same data set. To explore and better quantify the relationship between dropout rate and career length, we do not directly calculate the absolute value but consider the relative value of the dropout rate. The relative dropout rate *RDR* is defined as the ratio between the number of researchers who terminate his career at a certain career length cl and the total number of researchers whose career length $\mathrm{CL} \geq \mathrm{cl}$. The results are shown in Fig. 2c. One can see a clear and intriguing pattern about dropout rate. First of all, the rate of the first year is remarkable more than 0.35. In the next few years from Y_1 to Y_2 , the fraction of *RDR* has been experiencing a dramatic downward trend since the Y_1, leading to a departure in the share of junior scholars. However, starting from a certain year, the *RDR* maintains stationary state less than 0.05. After the stable period for some time, *RDR* begins to deviate from a baseline performance at a certain point Y_3 in the career, which then sustains a little higher value. In order to obtain Y_2 exact value, we recalculate the cumulative absolute dropout rate *CDR* instead of relative dropout rate in Fig. 2d. As the result shows, two linear equation slopes come cross a crossover point at $\mathrm{cl} = 5$, which represents two distinct status on the left and right side of the point. In summary, the evolution of academic career length based on researchers' dropout rate could be divided into four periods including starting year, trainee phase, maturity period and retiring stage, which is consistent with general cognition about science progress. Young scholars used to be trained for 5 years nearly during their Ph.D. period in order to get stellar career. The trend indicates that it is the most competitive period in the first five years of academic career. These results raise an important question: what factors are related to the temporal regularities observed across diverse career histories?

To answer the question, we first investigate the references aimming at researching the mechanisms of academic career dynamics. Investigation results demonstrate that there are mainly four factors playing roles on dropout and survival rate, including scientific creativity, personal characteristics (i.e. age and gender), cumulative advantage and collaborators (Zeng et al. 2017). To exemplify this complexity, we are mainly interested in the scientific creativity quantified by *Pubs* and C_{10}. Note that we use the sample of researchers who published their first paper in 1928–1999 years because we explore the dropout rate in first 5 yeras at the beginning of career path in this part. It is calculated by $\mathrm{DR} = N_q/N$, where N_q is the number of the dropouts in first 5 years and N is the number of the total scientific researchers in the data set. Firstly, Fig. 3a shows the probability density distribution

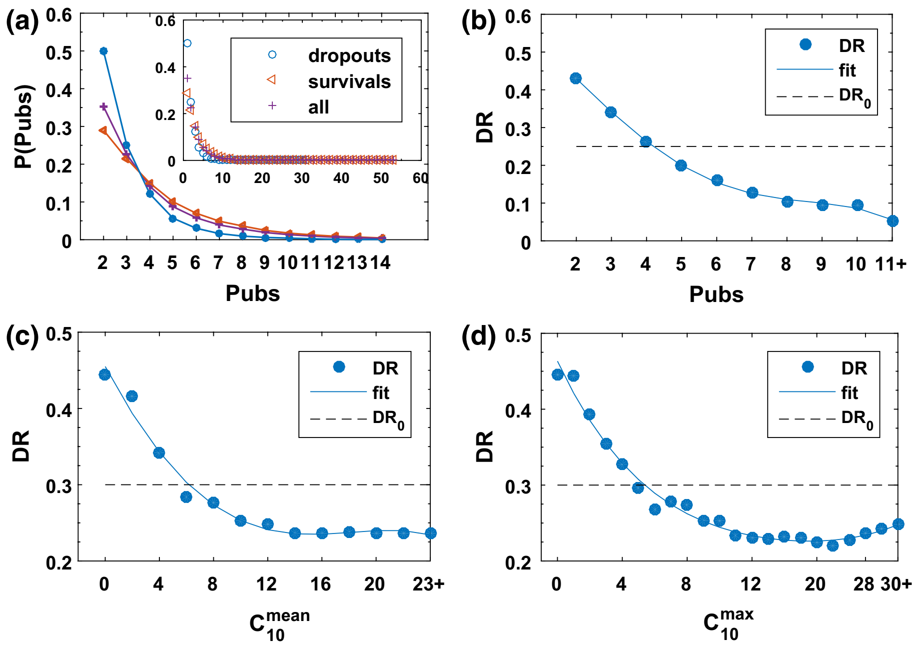

Fig. 3 (Color online) **a** the fraction of researchers' publications at first 5 years and **b** the relationship between DR (dropout rate) and early Pubs (publications), which shows nonlinear effect and **c** the relationship between DR (dropout rate) and early C_{10}^{mean} (mean impact), which shows nonlinear effect. and (d) the relationship between DR (dropout rate) and early C_{10}^{max} (max impact), which also shows nonlinear effect

of *Pubs* during researchers' trained period. In this part, we estimate the career length as the interval of publications and do not consider careers with only one publication. If we presume the crossover point as the critical publication number, the distribution reflects that almost three-quarters dropout researchers have publications less than critical number and 50% survivals' publications are more than critical number. Here we seek to identify quantitative patterns in the scientific career trajectory towards a better understanding of how scientific production responds to dropout rate. The relationship between dropout rate and publication is shown in Fig. 3b. The dotted gray line represents the average dropout rate. In practice, we leave behind the young scholars without any publications from 2 to 5 years in their early career but lately return to do science research (i.e. we don't find their second publication until 5th year later). This reduces the size of dataset by approximately 25%. Obviously, one can see that the dropout rate *DR* decreases with *Pubs* and the rate of descent gradually slows down. The crossover point is a little larger than 4 but smaller than 5 which means that producing about 4 publications could satisfy the mean research production performence of surviving scholars. On the other hand, in science, the ability to attract future opportunities is strongly related to citations. As we know, citations are used to measure the academic performance of scientists. We then take into account factors of C_{10}^{mean} and C_{10}^{max}, respectively standing for the average and top level of publications in the second stage of research career. In Fig. 3c, one can see that DR indeed decreases with C_{10}^{mean}, and the relationship between DR and C_{10}^{mean} is highly nonlinear. Specifically, *DR* decreases significantly with C_{10}^{mean} when C_{10}^{mean} is relatively small, yet becomes almost independent of C_{10}^{mean}

when C_{10}^{mean} is large. The tipping point is around $C_{10}^{\mathrm{mean}} \approx 12$ where the relationship between *DR* and C_{10}^{mean} changes. As long as $C_{10}^{\mathrm{mean}} > 12$, additional effort in improving C_{10}^{mean} will not result in a substantial lower dropout rate. Although some scholars have extremely large C_{10}^{mean}, they still sustain the same dropout rate as the weaker survivals. In Fig. 3d, we can see the similar results in the relationship between DR and C_{10}^{max}. Over this horizon, the scientific performance of early career can significantly alter the career trajectory. In order to avoid being weed out from the survivals, as for the young scholars, it is beneficial to try their best to improve productions to 4 publications and scientific impact to C_{10}^{max} and C_{10}^{mean} equaling to 12. However, if such academic impact and productions are not available, a question is naturally posed: what kind of scientific research performance can help young scholars continue to engage in scientific research work after trained period.

As Fig. 3a displays that some young scholars who have high research outputs drop out academia. However half of survivals only have publications amount less than the critical number. We speculate that integrated factors account for the uncommon phenomenon where academic publication amount and impact together are key features. Specifically, using the same set of young researchers, a companion study has analyzed the ordered publications sequence of each scientist with a focus on the statistical regularities underlying publication impact C_{10}. To explore and better quantify the relationship,we present the average curve in addition to the scatter plot. The publications of a scientist is sorted according to its publication date in first 5 years. The average impact $C_{10}(t)$ of the xth publication is averaged over the $C_{10}(t)$ of xth publication of the scientists in the corresponding group (i.e. dropouts or survivals). Interestingly, we find two visibly distinct trends of dropouts and survivals. The average impact of survivals stays almost unchanged regardless of sequence length. This observation is consistent with a recent study (Sinatra et al. 2016) which demonstrated that a scholar's most cited paper can be any of his or her papers and is independent of the age or career stage when it is published. Survivability in the workplace requires that the individual maintains his/her performance level in order to avoid being replaced by others. By contrast, junior dropouts perform worse and worse in academic career in which the gap of scientific performance is getting wider untill career termination as purple line shown in Fig. 4. Hence, in short trained systems, the dropout rate of academic careers is determined not only largely by scientific creativity but also by persistence. Many careers have difficulty making forward progress due to the relative disadvantage associated with lacking early career continuous innovation. These findings may also be helpful for young scholars when making efforts to get tenure position.

The analyses mentioned above show that there is an obvious effect of scientific creativity quantified by *Pubs* and C_{10} on young scholars' careers progress. Finally, we study whether early scientific performances noted as $\mathrm{Early}\langle \mathrm{Pubs}\rangle$ and $\mathrm{Early}\langle C_{10}\rangle$ will have a further impact on the dropout rate of survival scholars at mature stage. Nearly every individual faces with the potential risk of ending the career suddenly, possibly as the result of poor performance, bad health, family responsibility, economic downturn, or even a change in the development strategy of institution or employer. However, the relationship between early performance and later dropout is unknown yet. To this end, we select two groups of survivals who successfully get through the training period into a stable period. Respectively, their career length are $15 \le \mathrm{cl} < 20$ and $25 \le \mathrm{cl} < 30$, including in total 16,037 scholars. In order to analyze the fluctuant properties of dropout rate in our sample, we define the normalized value $\mathrm{DR}/\mathrm{DR}_0$. The quantity DR_0 is the average dropout rate which is calculated by the value of group 1 dividing the total value of the two groups. Fig. 5a shows the characteristic dropout rate trajectory with $\mathrm{Early}\langle \mathrm{Pubs}\rangle$. It is obtained by calculating normalized *DR* belonging to each subset. Rather than the regularity at trained period, it is sustained

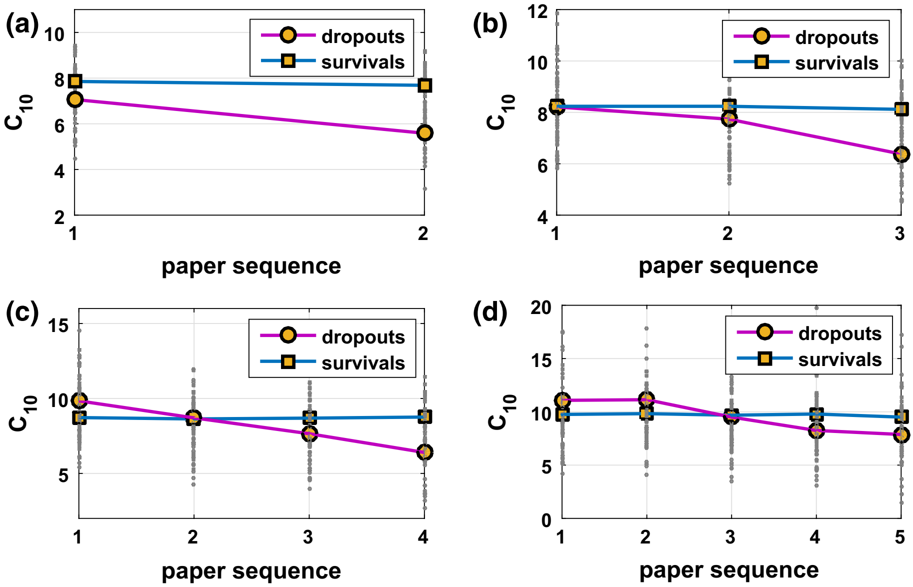

Fig. 4 (Color online) The evolution of the impact C_{10} of researchers' xth papers sorted according to publication time. The results of dropouts and survivals are compared. Each panel shows the results of scientists with different number of publications in their first 5 years, and a consistent trend can be observed

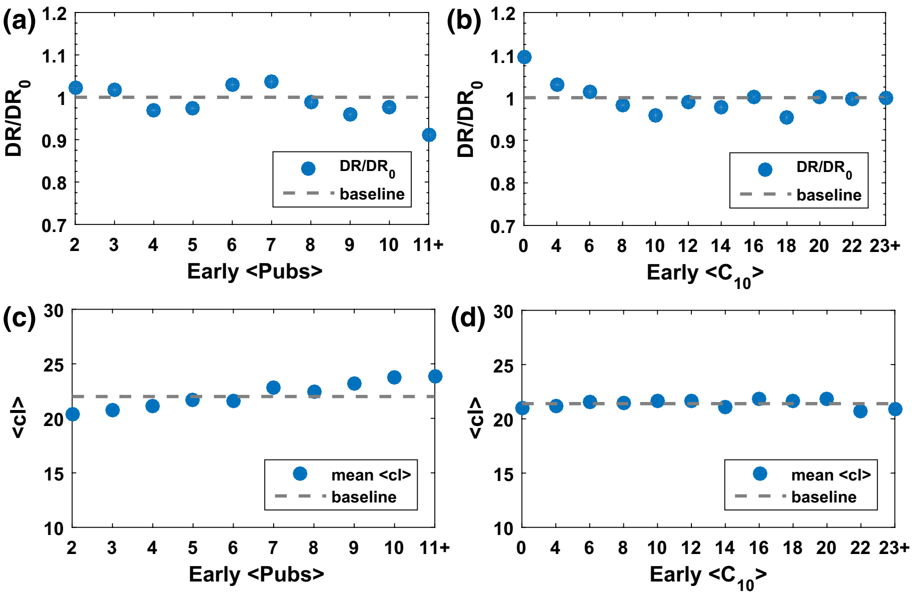

Fig. 5 (Color online) **a**, **b** The effects of the quantity and mean quality of early publications on normilized dropout rate in mature stage, respectively. The selected objects are researchers at different career status: leaving academia after 15–20 years and 25–30 years after the first publication. **c**, **d** The trends of survivals' mean career longevity based on $Early\langle Pubs\rangle$ and $Early\langle C_{10}\rangle$

to level around 1 with the increasing of Early$\langle$Pubs$\rangle$. A similar pattern with Early$\langle C_{10}\rangle$ emerges in Fig. 5b. Next, we test how career longevity depends on early publications and citations, among the set of researchers including 37,304 scientists who accumulate 5 years of backgroud experience in Fig. 5c, d. The mean career length of survivals from different subsets both fluctuate slightly near the baseline (i.e. mean career length of whole survivals). Therefore, for survivals, the results suggest that early performances have little impact on ultimate career length. Consequently, in such systems, the long-term appraisal time scale averages out fluctuations. So late careers are significantly less vulnerable to academic exploratory periods and hence more sustainable because they are not determined primarily by early fluctuant performances.

Discussion of results

The findings revealed in this article are easy to understand as it is consistent with intuitive scenario. Specifically, we take advantage of the entire scientific labor force using comprehensive longitudinal data not just tenure-track scientists. The first result indicates that group academic career can be clearly divided into 4 periods with significant shares of the research workforce dropping out of research-active academic positions at each stage. According to our definition, they are starting year, trainee phase, maturity period and retiring stage, respectively. At early exploratory stage (i.e. the first 5 years including starting year and trainee phase), young scientists generally stay in postdoctoral short-term academic position or other forms of contingent academic work with drastic indeterminacy. Our results show that 75% of them can move into a safe academic career in which the relative dropout rate is less than 5% after five years of volatility. These analyses reveal the general science workforce pattern in career trajectory.

In current competitive environment, we show some significant evidence and factors about promoting the successful transition from temporary workforce to tenure position in academic career. We demonstrate that effect of the number and impact of early publications are both negatively nonlinear on the dropout rate in pathway to maintain a long active career in science. In details, publishing not less than 4 papers and reaching the level 12 of C_{10}^{mean} could meet the criteria for admission to academia according to our data set. Moreover, it should be noted that the continuous knowledge and innovation investments are essential in this stage. In this sense, our paper provides guidance for junior scholars in allocating their research effort in the early career. Furthermore, we find this effect is only obvious in progress of entering a profession rather than the dropout phenomenon during mature period. In other words, the career length for the mature survivals does not depend on their early research performance.

Applicability and lessons learned

Objective quantification of the research of scientific careers dynamics is a long-standing challenge in science of science (SciSci). Much effort and many findings have been made in this filed. However most studies so far have focused on relatively small samples and qualitative analysis. Unlike other occupations, academia is more mysterious and complex because only very few elites engage in. For junior researchers who are attracted by academic calling but lack much of academic experience to develop the first academic career , they may show concern about the uncertainty and prospects of their academic careers. In this paper, from a new perspective of dropout rate in scientific career, we systematically analyze general academic dropout pattern and identify the critical factors related to avoid to be a dropout and obtain permanent academic position in early academic career.

The findings in this paper are made based on publication records in physics filed. APS data includes collection of 14 leading peer-reviewed research journals which is a good sample for researching the dropout rules of physicists. We cannot guarantee strictly that all of them are applicable to researchers from different disciplinary backgrounds. Meanwhile, some researchers would like to shift their research interest to other fields rather than dropping out from the whole field of science. While we cannot address these with our current data, we remark that our statistical approach is general and possible to expand this study to other fileds. While these findings do not imply causality, it is still suggestive for contemporary scientists, especially for junior scholars. The first stage (i.e. the starting year) of an academic career are fraught with uncertainty where junior researchers face competition and have other occupational planning out of academia. After deciding to get involved in academic work, they will have a high probability (75%) of surviving in the academic work space as long as they make continuous knowledge and innovation investments. They need to improve the number and impact of their articles as much as possible for surviving to next phase at that duration. However, excellent performances in early acamemia do not affect long-term survival. Taken together, a possible implication of our results is that it may point out explicit directions of academic effort for young researchers suffering from complex science ecosystem.

Conclusions and future work

In this paper, we define a new conception relative dropout rate *RDR* in academia. And we find a simple rule that physicists' complex research lifecycle are nonlinear and divided into 4 segments by dropout and survival analysis. Early academic performances *Pubs* and C_{10} both have a strong nonlinear effect on dropout rate in early career rather than mature phase. Before the tipping point , the dropout rate *DR* decreases quickly. After that, the *DR* begins to stay stable under the effect of Pubs and C_{10}. In an attempt to render a more objective review process for tenure and other lifetime achievement awards, we use latest quantitative measures C_{10} for scientific publication impact. However, we do not take into account the level and size of initial collaboration, age and gender factors in this study. The complexity of collaborations play a great role on scientific success in per career stage. There are some further researches called for the relationship between collaboration and career dynamics. We will focus on the effect of team size and academic level on dropout rate in future work. What's more, with methods of network science, we can explore whether topological properties of the collaboration network have direct relationship with dropout rate or nor. Futhermore, more mechanism models that focus on modeling scientific careers in the context of survivability need to be designed to help us better understand career dynamics. We would like to make a multiple regression and statistical model to reproduce the nonlinear evolution rule in academic dropout.

Acknowledgements This work is supported by the Natural Science Foundation of Beijing (Grant No. L160008) and the National Natural Science Foundation of China (Grant Nos. 61603046).

References

Afonso, A. (2016). Varieties of academic labor markets in Europe. *PS, Political Science and Politics*, *49*(4), 816–821.

Allen, J., & Robbins, S. B. (2008). Prediction of college major persistence based on vocational interests, academic preparation, and first-year academic performance. *Research in Higher Education*, *49*(1), 62–79.

Clauset, A., Arbesman, S., & Larremore, D. B. (2015). Systematic inequality and hierarchy in faculty hiring networks. *Science Advances*, *1*(1), e1400005.

Cole, S., & Cole, J. R. (1967). Scientific output and recognition: A study in the operation of the reward system in science. *American Sociological Review*, *32*(3), 377–390.

Cyranoski, D., Gilbert, N., Ledford, H., Nayar, A., & Yahia, M. (2011). The PhD factory. *Nature*, *472*, 276–279.

Daniels, L. M., et al. (2011). Relieving career anxiety and indecision: the role of undergraduate students' perceived control and faculty affiliations. *Social Psychology of Education*, *14*(3), 409–426.

Deville, P., et al. (2015). Career on the move: Geography, stratification, and scientific impact. *Scientific Reports*, *4*(1), 4770.

Deville, P., Wang, D., Sinatra, R., Song, C., Blondel, V. D., & Barabai, A. L. (2014). Career on the move: Geography, stratification, and scientific impact. *Scientific Reports*, *4*, 4770.

Duch, J., et al. (2012). The possible role of resource requirements and academic career-choice risk on gender differences in publication rate and impact. *PLoS One*, *7*(12), e51332.

Editorial (2017). Spread your wings, Nature, 550, 429.

Fortunato, et al. (2018). Science of science. *Science 359*, eaao0185.

Fraiberger, S. P., et al. (2018). Quantifying reputation and success in art. *Science*, *362*(6416), 825–829.

Kim, B., et al. (2016). Tolerance of uncertainty: Links to happenstance, career decision self-efficacy, and career satisfaction. *The Career Development Quarterly*, *64*(2), 140–152.

Levecque, K., et al. (2017). Work organization and mental health problems in PhD students. *Research Policy*, *46*(04), 868–879.

Liao, C. N., & Ji, C. (2015). The origin of major choice, academic commitment, and career-decision readiness among taiwanese college students. *The Career Development Quarterly*, *63*(2), 156–170.

Merton, R. K. (1968). The Matthew effect in science. *Science*, *159*, 56–63.

Mills, D., & Paulson, J. (2014). Making social scientists, or not? Glimpses of the unmentionable in doctoral education. *Learning and Teaching*, *7*(3), 73–97.

Milojević, S., Radicchi, F., & Walsh, J. P. (2018). Changing demographics of scientific careers: The rise of the temporary workforce. *Proceedings of the National Academy of Sciences.*, *115*(50), 12616–12623.

Moss-Racusin, C. A., et al. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences.*, *109*(41), 16474–16479.

Musch, J., & Grondin, S. (2001). Unequal competition as an impediment to personal development: A review of the relative age effect in sport. *Developmental Review*, *21*(2), 147–167.

Petersen, A. M., et al. (2011). Quantitative and empirical demonstration of the Matthew effect in a study of career longevity. *Proceedings of the National Academy of Sciences*, *108*(1), 18–23.

Petersen, A. M., et al. (2014). Reputation and impact in academic careers. *Proceedings of the National Academy of Sciences*, *111*(43), 15316–15321.

Petersen, A. M., Riccaboni, M., Stanley, H. E., & Pammolli, F. (2012). Persistence and uncertainty in the academic career. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *109*(14), 5213–5218.

Scott, J., Long, & Paul, D. A. (1979). Entrance into the academic career. *American Sociological Review*, *44*(5), 816–830.

Seifert, K. H. (1994). Improving prediction of career adjustment with measures of career development. *The Career Development Quarterly.*, *42*(4), 353–366.

Sikora, J. (2018). Aimless or flexible? Does uncertainty in adolescent occupational expectations matter in young adulthood? *Australian Journal of Education*, *62*(2), 154–168.

Simonton, D. K. (1997). Creative productivity: A predictive and explanatory model of career trajectories and landmarks. *Psychological Review*, *104*(1), 66–89.

Sinatra, R., et al. (2016). Quantifying the evolution of individual scientific impact. *Science*, *354*(6312), aff5239–aff5239.

Stumpf, H. (1995). Scientific creativity: A short overview. *Educational Psychology Review*, *7*(3), 225–241.

Weinstein, F. M., Healy, C. C., & Ender, P. B. (2002). Career choice anxiety, coping, and perceived control. *The Career Development Quarterly*, *50*(4), 339–349.

Yoon, S., et al. (2018). Factors influencing career progress for early stage clinician-scientists in emerging Asian academic medical centres: a qualitative study in Singapore. *BMJ Open*, *8*(3), e020398.

Zeng, A., et al. (2017). The science of science: From the perspective of complex systems. *Physics Reports*, *714–715*, 1–73.