

# Is science a pyramid scheme? The correlation between an author's position in the academic hierarchy and her scientific output per year

Charleeze P. Ponzi

**A grievance expressed by some PhD students and Postdocs is that science works like a pyramid scheme: Young scientists are encouraged to invest into building scientific careers although the chances at remaining in science are extremely slim. This issue is investigated quantitatively by connecting it with the way authorship on papers is distributed. I analyzed a large bibliographic dataset made available by Microsoft under the name Academic Graph to create a histogram with the number of articles an author produces per year. The histogram has the shape of a pyramid, and different layers in it correlate with positions in the academic hierarchy. The super-prolific authors at the top of the pyramid with more than 40 publications per year are usually heads of large institutes with many subgroups and large numbers of PhD students, while the bottom of the pyramid is populated by PhD students and Postdocs with less than 5 publications per year. The mechanism that allows 'manager scientists' to appropriate publications generated in their sphere of influence is related to other issues, such as the evaluation of scientific performance based on scientometric indicators and the lenient enforcement of authorship rules. A new index, the Ponzi factor, is proposed to quantify this phenomenon.**

## Introduction

Recent decades have witnessed an increasing reliance on bibliometric indicators to evaluate performance of researchers and universities. (1) A massive increase in access to higher education and doctoral studies (2) was accompanied by policies to foster accountability for tax payers' money, transparency in the distribution of grants and job opportunities in science, and efficiency by forcing researchers to compete for limited resources. (3)

In this context, hard numbers infuse a sense of impartiality, accountability and comparability. (4) Decision makers are confronted with a dilemma: Limited resources have to be allocated based on some criterion. How can one judge research which one does not fully understand as an outsider? Since the value of basic research can sometimes only be felt after many years, judging scientific achievement is a bet on the future. The decision makers do not have expertise in all fields of study. Therefore, ranking scientists, projects and universities by using well-defined metrics, is an attractive shortcut to legitimizing decisions and increasing competition. The detrimental effects of this trend have been pointed out (3, 5, 6).

An entire research field called scientometrics has developed around quantifying scientific activity. Private companies such as the Institute for Scientific Information (ISI) and more recently internet giants such as Google and Microsoft enabled this trend by compiling digital databases of publications. The scientific literature can be represented as a huge graph, in which nodes represent authors and publications and edges between them establish authorship and citations. Nowadays the

necessary information to build the academic graph is extracted automatically from journals. Powerful computers make it possible to analyze the entire scientific literature and compute bibliometric indicators. The most well-known indicators are the h-index, introduced by Hirsch (7), to compare authors and the impact factor introduced by Garfield (8) to rank journals.

However, scientometrics is blind to the social relations between authors. The claim of this contribution is that extreme differentials in scientific productivity can be explained largely by authors' social status in academia. Scientometric indicators such as number of publications, citations or more complex ones like the h-index, while purporting to measure 'excellence' and 'leadership', to some extent are just maps of the social stratification of academia. A high position in the hierarchy automatically begets the quantifiable signs of academic excellence. The social structure of academia resembles a feudal pyramid where loyalty and authorship on publications are exchanged for funding and access to resources (9). To substantiate these claims, the academic graph compiled by Microsoft is analyzed and complemented with information about author's academic standing obtained from their curricula vitae.

This article is structured as follows: After explaining how the data was extracted from the academic graph the main results are presented, which are then discussed in the wider context.

## Methods

Microsoft's Academic Graph (MAG) (10) is a relational database that contains bibliographic information about more than 230 million publications and some 233 million authors. It consists of tables for authors, papers, affiliations and fields of studies that can be joined via various bridging tables. The database may be queried using Microsoft's U-SQL query language. In this study the 2019-10-18 release of MAG was used.

I chose to analyze authors with affiliations to US or German institutions, because the US is the world's leading scientific superpower, while Germany is the largest country in the European Union. Unlike in the US, in Germany the academic landscape is less dominated by few major research universities. In advance one can therefore expect different findings for the two countries. The field of study was limited to science (vs. humanities) since the practices and organization of authors in the two domains differ too much to treat them jointly.

For each author the number of papers published in any journal in the year 2018 was computed and stored as a new table. Authors from the US or Germany were selected based on the domain name in the official webpage of their most recent affiliation. To obtain histograms for different disciplines,

To whom correspondence should be addressed. E-mail: c.ponzi@mail.ch

the selection of authors was further refined by matching the field of studies assigned to their papers to patterns such as '%physics%' or '%materials%'.

A word of caution: The Academic Graph database is neither curated nor cleaned, it has been created automatically as a mixture of publisher feeds and data mining (10). As such the conclusions drawn from its data should be taken with a grain of salt. Authors with common names might be accidentally combined into a single record, which overestimates the publications per year. On the other hand an author whose name contains non-ASCII characters that can be transliterated in different ways might be split into multiple records thus underestimating her scientific output. Papers published as preprints on arXiv, that were later published in a journal under a slightly revised title could lead to double counting. Contrasting data from different sources such as Web Of Science (WoS) and Google Scholar allows to estimate the uncertainties in the Academic Graph. WoS tends to report lower publications per year as compared to Microsoft Academic. The reason for this might be that WoS does not automatically merge authors with identical names. For some authors multiple author records are found in WoS that clearly belong to the same person based on the narrow field of research. On the other hand, publication lists generated from Academic Graph sometimes contain a few publications that belong to a completely different field (such as a medical field study in the publication list of a theoretical physicist) and are clearly misplaced. It is also difficult to draw sharp boundaries between research fields. In the Academic Graph *fields of study* records can be linked to papers but not directly to authors. Therefore, authors working in a specific field are selected based on the classification of their papers. If we restrict the selection to authors publishing in the field of chemistry, the resulting data set will contain some physicists, biologist and medical scientist as well, since these fields overlap. Since super-prolific authors work in many fields - interdisciplinary projects are highly valued by funding agencies - the same scholars might show up in histograms belonging to different disciplines. Classifying authors by countries is relatively straight-forward, since the affiliation table provides the URL of the official webpage of an institution. Except for US universities (which have .edu or .gov domains) a country code can be extracted from the domain name. However, some institutions outside the US also use .edu domains. The most reliable data can be obtained from research-group websites and scientific profiles that are maintained by the authors themselves such as Google Scholar accounts. Unfortunately no database collects academic titles (Prof., Dr.) of authors. The author byline usually does not mention academic titles, either. \* Since the goal of this study is to establish a connection between the academic hierarchy and the scientific output, this information had to be retrieved manually from the web for a sample of authors. Career stages were obtained from CVs published on the web.

## Results

**A. Histograms of Publications per Year.** Figure 1 shows histograms of authors' scientific productivities in the fields of physics, chemistry and medicine in Germany and the US in the year 2018. The number of publications authored per year

is shown on the vertical axis, the horizontal bars give the number of authors in logarithmic scale. The histograms have the shape of one side of a pyramid. Apart from the overall larger number of authors, the pyramids for Germany and the US look remarkably similar.

In the physics pyramid there is an additional big bump between 100 and 200 publications per year. It contains mostly authors working in high energy physics and astronomy. These fields are sometimes described as 'big science', because commissioning of expensive machines such as accelerators or telescopes and design of experiments takes decades and huge investments. Publications in high energy physics are very peculiar: The paper about the discovery of the Higgs boson lists several hundred authors in alphabetical order (11). Because the practice of assigning authorship in 'big science' is so much different from other fields of science we leave the physics pyramid aside and focus our further analysis on the chemistry and materials science pyramid for Germany, which conforms more to the ideal shape of a pyramid.

The vast majority of authors, 95 %, publish less than 10 papers per year. On the other hand, there is a minority of 0.3 % who publish more than 50 (!) papers per year. These are the super-prolific scholars that are at the focus of this study. These authors at the top of the pyramid are usually directors of research hospitals, heads of large research institutes such as those belonging to the Max Planck society or department chairs and professors at universities. Super-prolific scholars occupy the highest positions in the academic hierarchy both in terms of academic titles (Prof., Chair, Director) and in terms of the prestige of the institutions they lead. All of them have very high h-indexes. The groups or departments they head are typically very large, sometimes with multiple subgroups and a clear hierarchy, although sometimes it is not possible to deduce a clear structure from the group's presentation on the web. At universities the research groups are often named after their leaders. The web presentations provide much more information about the leader, her achievements and career path than about any other member of the group. A research group rarely has more than one such significant figure.

The lower part of the pyramid is populated mostly by Postdocs, PhD and Master students who are about to initiate their academic careers, i.e. they do not have published many papers so far. Although it is more difficult to obtain information about the authors at the bottom than at the top of the pyramid, it is likely that many of them have fixed-term contracts. The bottom of the pyramid is also home to the majority of professors and researchers without stellar publication rates. However, since the number of PhD students and Postdocs vastly exceeds the number of professors, sampling from the bottom of the pyramid mostly produces PhDs and Postdocs. The super-prolific scholars thus form a highly visible minority among professors.

It is important to note that the histogram shows the average number of publication *per year*, not the total number of publications accumulated over an author's lifetime. It is clear that scientist who are in business for a longer time will have long publication lists. The analysis so far demonstrates that senior, powerful authors have not only longer publication lists, but also a much higher publication rate per year than less powerful authors. All of this suggests that the publication rate is a measure of the author's position in the academic

\* This is probably motivated by the noble idea that science is a community of equals.

hierarchy.

**B. Scientific Productivity and Career Stages.** We are confronted here with a chicken or egg causality problem: Either high scientific productivity (and a large h-index, which is considered a sign of scientific leadership) results in promotion to the top of the hierarchy, or it is the high position in the hierarchy that allows the author to put her name on all publications generated in the layers below? To answer this question, I analyzed the productivity of super-prolific authors in different stages of their career, i.e. I asked how many publications did they author per year as PhDs, Postdocs, professors or directors. Figure 2 shows the results for three typical European super-prolific authors 1,2 and 3. Interestingly, the publication rate of the same person also depends strongly on her position in the academic hierarchy over time.

This contradicts the assumptions made by Hirsch in his famous article introducing the h-index (7). Hirsch assumes a constant rate of publications per year,  $p$ . Of course, the rate of publication can vary slightly over the lifetime of a scholar for different reasons: Scientific writing has to be learned, and lack of experience and multiple revisions will make writing one's first paper a lengthy affair. After some practice a successful scholar should output publications at a steady rate. Close to retirement the productivity will drop again. The data contradicts this picture. Looking at the trajectories of super-prolific scholars in Fig. 2, we see that the rate of publication increases dramatically as they climb the hierarchy ladder. All super-prolific scholars in our sample start out with average rates in the humble range of around 0 to 5 publications per year. The explosion of productivity later in their careers can be linked approximately with important events chronicled in their curriculum vitae: leading one's own research group, tenure, promotion to full professor or becoming director of an important research institute. With each promotion comes the burden of additional responsibilities (teaching duties, sitting on committees, writing grant applications, reviewing, managing a research group), so that in theory there should be less time that can be dedicated to research. Nevertheless an increase of scientific output is observed. This paradox is resolved when we acknowledge that science works like a pyramid scheme: Each step up the ladder increases the number of subordinates that have to pay tribute in the form of publications. Access to facilities, funding and influential networks is exchanged for loyalty and publications. Resources flow from the top of the pyramid to the bottom, while publications flow in the opposite direction. Scientists at the top of the hierarchy become authors on publication that are produced in their field of influence.

The rate of publication in the early stages of a career is the one that is intrinsic to the author; this is what Hirsch probably means by the publication rate  $p$ . This rate does not vary greatly among all junior scientists and is below 10 publications per year. In later career stages the publication rate depends on the position in the academic hierarchy, i.e. the distribution of publication rates mirrors the social stratification of academia. The scientific output is thus tied to the position and less to the person occupying that position. Authorship on publications is a form of non-monetary compensation the comes with the position.

An interesting European peculiarity stands out in the career paths of authors 1-3: The CVs of all three European super-prolific authors document a one year stint at a prestigious US

university. Although the scientific output during this time is small compared to the maximum output at later times, these short stays seem to function as booster rockets that propel the careers of the super-prolific authors into high orbit.

By scientific metrics such as number of publications or h-index, authors 1,2 and 3 would be considered successful scientists. However, I argue that they represent a relatively new type of scientist, which I would like to call the 'manager scientist'. To differentiate them from 'traditional scientists' it is instructive to compare with the publication track record of Jorge Hirsch, the inventor of the h-index (bottom of Fig. 2).<sup>†</sup>

Hirsch seems to have a small (if any) research group, often writes single-author papers or papers with few co-authors. Periods of high productivity alternate with doldrums, leading to an average of 6 publications on average per year. His h-index is 65, which makes him a scientific leader by his own standards. J. Hirsch is situated at the lower part of the histogram of publications, although in some years he has published over 10 papers, which is no small feat. Nevertheless, he is still separated by a factor of more than 5 from the league of super-prolific scholars who publish more than 50 papers per year and occupy the highest positions in the academic hierarchy. Hirsch's publication rate is not constant over time, but it does not correlate with career level as in the case of the other authors. Clearly the publication patterns of authors 1 to 3 and J. Hirsch are completely different. This suggests that scientists can be divided in two classes, the 'manager scientists' and the 'traditional scientists'.<sup>‡</sup> The two classes differ in how they organize their work and how they build their scientific authority. The traditional scientists choose their own projects and are responsible for all steps of a study from the design to the execution and publication. Division of labour only occurs when different authors bring different expertise to the project, but not primarily to make the production of scientific results more efficient. Their scientific output depends directly on their own scientific activity. As opposed to this, the manager scientists derive their reputation from managing a research group. Sometimes division of labour is observed in managed research groups: one person only works in the lab, a second one analyzes the data, a third one writes the papers, a fourth one is responsible for writing grant applications.

Both types of scientists are evaluated using the same h-index. Actually the manager scientists have extremely high h-indexes (larger than 100). Therefore I propose a new p-index ( $p$  for Ponzi) that complements the h-index to evaluate a scientist and distinguish between a traditional scientist and a manager scientist.

**C. The Ponzi index.** The Ponzi index of an author  $A$  at time  $t$ ,  $\text{Ponzi}(A, t)$ , is simply the number of peer-reviewed articles authored by  $A$  in the previous year, i.e. the rate of publication. The Ponzi index can be defined for publications and journals, as well. The Ponzi index of a publication  $P$  published at time  $t(P)$  by the collective of authors  $C(P) = \{A_1, \dots, A_n\}$  is the largest of the Ponzi indexes of any author at the same time,

$$\text{Ponzi}(P) = \max_{A \in C(P)} \text{Ponzi}(A, t(P)). \quad [1]$$

<sup>†</sup> There is no particular reason for singling out J. Hirsch as an example of a 'traditional scientist', there are many others, except that his name is mentioned anyway in connection with the h-index.

<sup>‡</sup> The term 'traditional' is used in opposition to 'manager' to highlight the changes in the academic system which can be compared to changes in production during the industrial revolution: Power was shifted from the traditional craftsmen to the manager of a factory.

If a paper has a large Ponzi index this means that its author byline probably contains at least one super-prolific scholar. The Ponzi index of a journal  $J$  is the Ponzi index averaged over all articles published in the same year,

$$\text{Ponzi}(J) = \text{avg}_P \text{Ponzi}(P). \quad [2]$$

If the Hirsch index quantifies an individual's scientific leadership, the Ponzi index quantifies the individual's managerial abilities and in extreme cases the willingness to break authorship rules.

Figure 3 shows a scatter plot of journal Ponzi factors versus journal impact factors for scientific journals. The Ponzi factors are computed for the year 2018 and the impact factors for the year 2019 as  $\text{IP}(2019) = \frac{\text{Citations}(2017) + \text{Citations}(2018)}{\text{Publications}(2017) + \text{Publications}(2018)}$ . There is no clear correlation between the two indicators. Most journals cluster around low impact and low to medium Ponzi indexes. Not surprisingly, several high energy and astrophysical journals are found at very high values of Ponzi. Material science journals also score high values for the Ponzi indicator. Nature journals lie approximately on the diagonal, they have both high impact and high Ponzi factors. Since publications in Nature are highly valued and serve as a ticket into the upper echelons of the academic hierarchy, manager scientists will go to great lengths to secure authorship on a Nature paper.

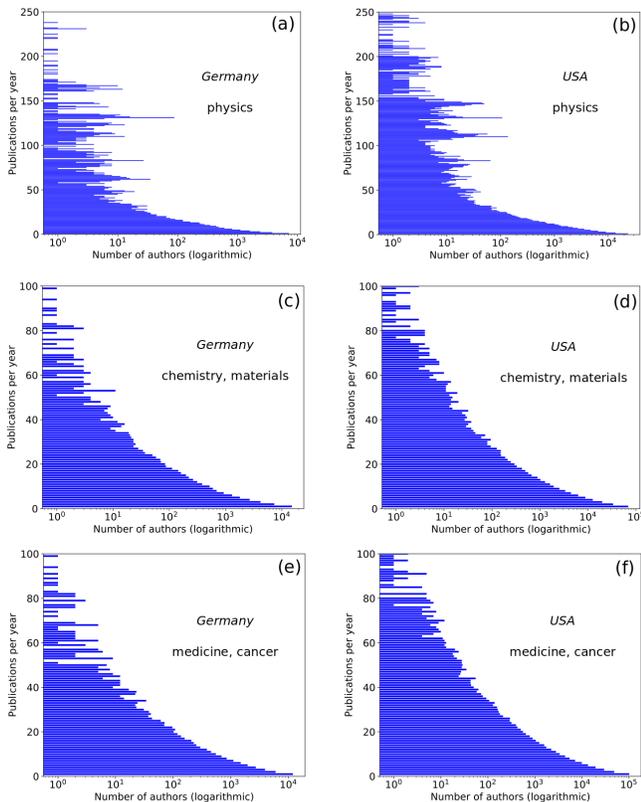


Fig. 1. Histograms of number of publications per year for different fields of study in Germany and in the US (in 2018).

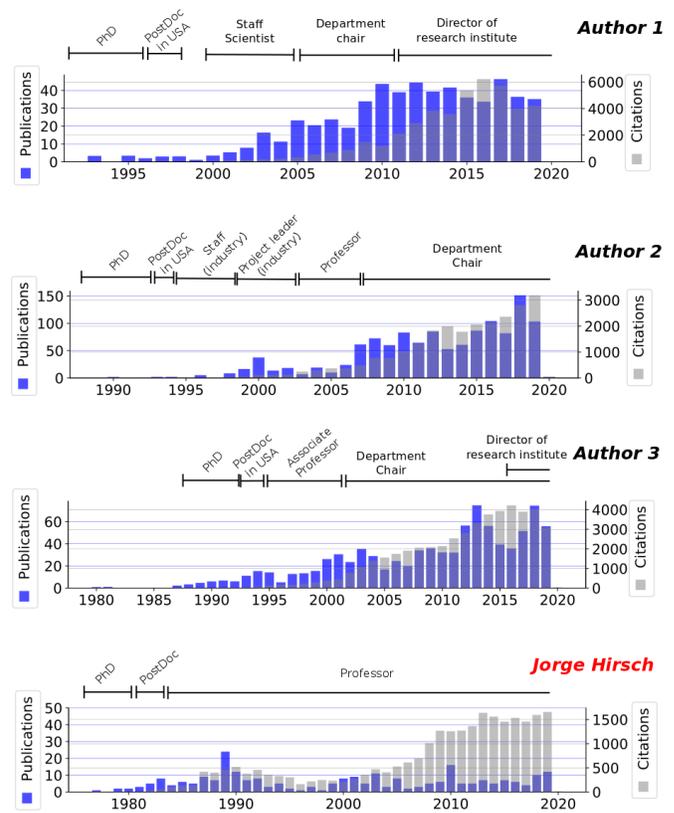
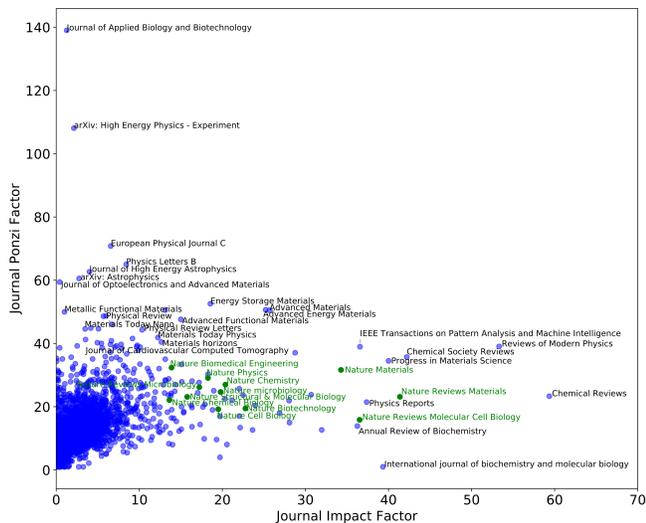


Fig. 2. Publications per year (blue), citations (gray) and career stages for three super-prolific authors from Europe and Jorge Hirsch. In order to prevent that the publication 'fingerprint' can be used to identify an author, each data point was multiplied by a random number from the range [0.85, 1.15]. The career stages were taken from the curriculum vitae, which the authors themselves or their institutions publish on their webpages. For anonymity years were randomly moved by  $\pm 1$ . For authors who hold multiple positions at the same time (professor at X, visiting professor at Y and Z) only the positions of highest rank are shown.



**Fig. 3.** Scatter plot of journal impact factors vs. Ponzi factors. Only journals whose names contain any of the following patterns are plotted: physic, chem, bio, comput, machine, material, quantum. Where several points overlapped, some journal names were deleted. Names are shown for journals with Ponzi  $\geq 40$  or impact  $\geq 35$ . Journals belonging to the Nature group are shown in green.

## Discussion

**A. The Meaning of Authorship.** This study touches on some interesting questions: What is authorship? What does it mean to be listed as an author of a paper? Since authorship requires a significant intellectual contribution to a publications, is there an upper limit on the number of publications per year a scholar can reasonably accomplish without violating basic ethical standards? There definitely is, and although it is hard to settle for one specific number, it should be clear that the limit is probably closer to 10 than to 50.

Publications are the currency of the academic world, but unlike money publications cannot be easily transferred, because authorship is tied to having made a significant intellectual contribution to that particular publications. One would expect that the requirements of authorship preclude the emergence of ‘publication millionaires’. Interestingly, however, it seems that the maximum number of publications per year is not limited by the self-constraint of the super-prolific authors or by some regulations, but simply by the maximum size of research organizations.

There are some unwritten rules concerning the order of authors on a paper: Often it is understood that the first author did most of the work, the last author is the leader of the research group, while the space in the middle is reserved loosely for authors that contributed to a minor extent or simply for returning favours. These unwritten rules stand in stark contrast to the official guidelines for good scientific practice established by the German Research Foundation (DFG) in Ref. (12):

Authors of scientific publications are always jointly responsible for their content. Only someone who has made a significant contribution to a scientific publication is deemed to be its author. A so-called “honorary authorship” is inadmissible.

The guidelines clearly state that managerial functions, al-

though being very important for the overall success of a project, should not be compensated by authorship (12):

Therefore, the following contributions on their own are not sufficient to justify authorship:

- merely organizational responsibility for obtaining the funds for the research, . . .
- only reading the manuscript without substantial contributions to its content, . . .
- directing an institution or working unit in which the publication originates, . . .

Neither the position of institute director and supervisor nor former supervisor justify designation as co-author.

Guidelines published by the UK Research Integrity Office (13) and the International Committee of Medical Journal Editors (14) contain similar requirements and agree on stipulating a significant intellectual contribution for authorship.

Whether an author has made a significant intellectual contribution to a publications is impossible to judge by an outsider. However, statistically an author who publishes in excess of 50 papers per year, is very likely to violate the guidelines in regard to several of those publications. In spite of this, super-prolific scholars receive generous funding from the German Research Foundation as documented in the funding database (15).

The DFG guidelines also oblige scientific journals to commit themselves to upholding the principles of good scientific practice (12), although in practice they lack the incentives and resources to actually enforce those rules. The activities of super-prolific authors are tolerated by journals and funding agencies since they increase the scientific output, which in turn is used to demand and justify large investments into research. However, increases in funding do not improve the situation, since they only make the bottom of the pyramid broader.

**B. The Scientific Evaluation and Reward System.** My findings also have implications for the way scientific projects are evaluated and which kind of performance is rewarded. I claim that scientific metrics are very good at measuring a manager’s performance, but that it is much more difficult to convert the performance of a traditional researcher into a number. When using scientometrics as the basis on which positions in the scientific hierarchy are distributed, one tacitly equates managerial success with scientific leadership.

In particular the h-index has been criticized by professional bibliometricians because it does not necessarily measure “importance” or “quality” (16). In fact, for a sample of 250 Danish professors working in health science, it was shown that there is a very good correlation between h-index and the number of publications (17). This means that the most reliable path to boosting one’s h-index is to increase the rate of publication. Since a high h-index is rewarded by funding agencies, the ‘manager scientist’ takes decisions that maximize the h-index of her institution. These include (1) hiring as many PhD students and Postdocs as financially possible, (2) imposing a minimum number of publications for graduation, and (3) participating in many projects, since sticking a tiny contribution onto a larger project guarantees authorship at a low cost. All these strategies lead to large dysfunctional research groups.

Since they are oriented towards maximizing the h-index of their leadership they do not provide any career opportunities for younger scientists. Cynically this is justified by saying that a PhD or Postdoc is a form of apprenticeship (18). Since the fate of a research group depends on its leader's ability to attract funding, maximizing her h-index can easily be framed as the common good. In this way, violation of authorship rules masquerades as a customary practice that benefits all involved. To be fair, some leaders of research groups are probably forced into this position by circumstances.

Hirsch himself likens the h-index to a stock at the stock market and advocates taking decisions about academic appointment and funding based on this metric (19). The advantage of metrics such as the valuation of a company's stock clearly consists in abstracting away the underlying details. Aggregates of various metrics can be formed for evaluating universities or entire research fields. However when using this analogy, one should not forget, that the stock market only works because of strict regulation and enforcement of laws. To avoid the worst abuses, the stock market is highly regulated with auditing requirements and anti-fraud units that can investigate offenders and bring them to court. As compared to the financial industry, in the knowledge producing industry regulation is almost absent (20). Academia is expected to police itself, often via an ombudsperson (12), who unlike a proper judge is neither independent nor qualified.

If the h-index is used to allocate resources, the high turnover of PhDs and Postdocs at the bottom of the pyramid gives an unfair advantage to senior scientists over their younger competitors. Since junior coauthors are generally removed from academia after at most a few years of Postdoc their h-indexes become irrelevant, while their contribution to the h-index of their superiors persists. <sup>§</sup>

**C. Implications for Careers in Science.** The types of work contracts also play an important role in shaping the pyramid. The few permanent positions in science are those related to management of research groups, while the majority of junior scientists have temporary Postdoc contracts. <sup>¶</sup>

The pyramid structure is characterized by an extremely high turnover at the PhD and Postdoc level. The market value of a Postdoc decreases rapidly over time. If not able to secure a permanent position, she will be rapidly replaced by one of the junior PhDs / Postdocs. This has the obscene consequence that in science (unlike most other occupations) experience counts for little and most publications are written by inexperienced newcomers. This is often justified by the claim that young people bring fresh ideas. The main reason, however, is that an academic system where every Postdoc is awarded a permanent position would be unsustainable as the size of the professorate would grow exponentially (21). Postdocs which leave academia are therefore like disappointed investors in a pyramid scheme (22) that give up on their investment.

The overproduction of PhD students (21, 23) and the plight of Postdocs is directly connected to the violation of authorship rules. If group leaders and directors would not be allowed to

put their names on dozens of publications, the incentive for creating huge multi-leveled research groups would disappear. Researchers would only hire as many PhD students as they can actually supervise and Postdocs or subgroup leaders would receive due credit for independent work.

The Matthew effect in science (24) is compounded by the fact that in research groups of super-prolific scholars most papers are written by PhDs or junior Postdocs with little supervision, who might not have a complete overview of the literature in their field. When asked to cite relevant papers, they will lavish citations on the most visible recent papers of prominent scholars, while relevant older papers with little citations are overlooked. De Solla Price succinctly states that "citations are generated by a pull mechanism from previous citation rather than from a push mechanism of the papers that do the citing" (25). This is an autocatalytic process that inflates the h-indexes of those highly visible, super-prolific scholars at the top of the pyramid further and breeds incompetence and subservience to metrics of scientific productivity at the bottom.

A young scientist (PhD student or Postdoc) might ask: What do I have to do to succeed in science? The immediate answer would be to publish as much as possible to make a name for herself. However, as I have shown, the number of publications does not depend so much on the personal effort but on the position in the hierarchy. In order to boost one's publication output one has to move up the ladder quickly. In an academic world driven by inflated metrics, the level of scientific output needed to reach visibility greatly exceeds what can be accomplished by single person. Therefore the appropriation of work performed by subordinates is an essential factor for success. To be successful, a scientist has to secure a position where she is in charge of managing other scientists.

**D. Implications for PhD studies.** Another, maybe overlooked aspect of the dire career prospects in science is that it changes the attitude of students towards a PhD. Very specific knowledge acquired during research cannot be converted into money outside of academia. A defensive strategy employed by PhD students is to shift their efforts from digging deeply into a problem (and maybe making a tiny scientific contribution to a difficult problem) to preferentially working on topics that provide them with skills which are in high demand in industry (such as software development and recently machine learning). To save on training costs companies like to hire PhDs which have proven track record in these technical fields. Therefore research projects are designed and advertised to potential PhD students as pretexts for increasing their employability (along the lines 'A PhD in theoretical physics is a safe bet, because later one can work in finance, data science, software development etc.'). This has the effect that funds allocated to research are actually diverted to schooling the workforce of the knowledge-based economy. It might be a consequence that is actually intended by decision makers (18).

Recent changes in the examination regulations at some German universities also cater to super-prolific authors and their thirst for publications. For heads of research groups keen on increasing scientific output it makes sense to hire candidates with little interest in a scientific career, since a minimum of 1-3 publications in peer-reviewed journals are often a prerequisite for having a PhD thesis accepted. (26) The surge of cumulative theses also manifests the priority of

<sup>§</sup>The h-index has certainly had a very big impact, but not necessarily the intended one. It is fair to say that the h-index has done to scientific integrity what the H-bomb has done to the Bikini island.

<sup>¶</sup>A director of an institute or university chair is like an impresario managing a traveling Postdoc circus. Postdocs are hired for one season to do conjuring tricks in front of the funding agencies in the hope that their contracts are extended for another season. However, the precariousness of the "circus life" does not allow for any serious scientific endeavors that require long term continuity.

professors to increase scientific output: A collection of articles is accepted as a substitute for writing a doctoral thesis as a single self-contained document. The PhD thesis, which demonstrates background knowledge and can be attributed only to a single author, is considered a waste of time. By requiring that work presented as a thesis is published in peer-reviewed journals prior to graduating, the duty of quality control is shifted from the thesis defense committee to the peer-review process. Thesis defenses thus degenerate into mere formalities.

## Conclusion

Now on to the final verdict: Is science a pyramid scheme? A large number of PhDs and junior scientists is necessary to sustain the high scientific output of super-prolific authors. Credit for publications is transferred to the top of the pyramid by violation of authorship rules. Junior scientists are encouraged to pursue scientific careers although the structure of the pyramid is sustainable only if the vast majority of them later lose their investment of time and effort. These characteristics turn the current academic system into a Ponzi scheme. The irony is that it is the most educated section of society that falls for this scam.

1. David Adam. The counting house. *Nature*, 415:726–729, 2002.
2. David Cyranoski, Natasha Gilbert, Heidi Ledford, Anjali Nayar, and Mohammed Yahia. Education: the phd factory. *Nature*, 472:276–279, 2011.
3. Peter Weingart. Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics*, 62(1):117–131, 2005.
4. C Nicolini, S Vakula, M Balla, and E Gandini. Can the assignment of university chairs be automated? *Scientometrics*, 32(2):93–107, 1995.
5. Neal S Young, John PA Ioannidis, and Omar Al-Ubaydli. Why current publication practices may distort science. *PLoS medicine*, 5(10):e201, 2008.
6. John Smyth, Smyth, and Christie. *Toxic University*. Springer, 2017.
7. Jorge E Hirsch. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46):16569–16572, 2005.
8. Eugene Garfield. Citation analysis as a tool in journal evaluation. *Science*, 178(4060):471–479, 1972.
9. Christopher Holligan. Feudalism and academia: Uk academics' accounts of research culture. *International Journal of Qualitative Studies in Education*, 24(1):55–75, 2011.
10. Arnab Sinha, Zhihong Shen, Yang Song, Hao Ma, Darrin Eide, Bo-june Paul Hsu, and Kuansan Wang. An overview of microsoft academic service (mas) and applications. In *Proceedings of the 24th international conference on world wide web*, pages 243–246. ACM, 2015.
11. Tatevik Aad Abajyan, B Abbott, J Abdallah, S Abdel Khalek, Ahmed Ali Abdelalim, O Abdinov, R Aben, B Abi, M Abolins, and many others. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1–29, 2012.
12. German Research Foundation. Proposals for safeguarding good scientific practice, 2013. URL [https://www.dfg.de/download/pdf/dfg\\_im\\_profil/reden\\_stellungnahmen/download/empfehlung\\_wiss\\_praxis\\_1310.pdf](https://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf).
13. UK Research Integrity Office. Code of practice for research. promoting good practice and preventing misconduct, 2009. URL <https://ukrio.org/wp-content/uploads/UKRIO-Code-of-Practice-for-Research.pdf>.
14. International Committee of Medical Journal Editors. Defining the role of authors and contributors, 2020. URL <http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html>.
15. German Research Foundation. German project information system, 2019. URL <https://gepris.dfg.de/>.
16. Cameron Barnes. The h-index debate: an introduction for librarians. *The Journal of Academic Librarianship*, 43(6):487–494, 2017.
17. Natascha Gaster and Michael Gaster. A critical assessment of the h-index. *Bioessays*, 34(10):830–832, 2012.
18. Max Planck Society. Scholarships are also a sign of quality. what is a doctoral thesis all about?, 2011. URL <https://www.mpg.de/5724370/scholarships>.
19. Jorge E Hirsch. Does the h index have predictive power? *Proceedings of the National Academy of Sciences*, 104(49):19193–19198, 2007.
20. Gerhard Fröhlich and Leonie Mück. Interview with prof. dr. gerhard fröhlich on the schön scandal: "self control mechanisms are a myth in science to avoid any serious external control". *JUNQ-Journal of Unsolved Questions*, 1(2):XII–XIII, 2011.
21. Richard C Larson, Navid Ghaffaradegan, and Yi Xue. Too many phd graduates or too few academic job openings: the basic reproductive number  $r_0$  in academia. *Systems research and behavioral science*, 31(6):745–750, 2014.
22. Beryl Lief Benderly. Does the us produce too many scientists. *Scientific American*, 2010.
23. Mark Taylor. Reform the phd system or close it down. *Nature News*, 472(7343):261–261, 2011.
24. Robert K Merton. The matthew effect in science: The reward and communication systems of science are considered. *Science*, 159(3810):56–63, 1968.
25. Derek de Solla Price. A general theory of bibliometric and other cumulative advantage processes. *Journal of the American society for Information science*, 27(5):292–306, 1976.
26. Friedrich Schiller University Jena. Implementing provisions for cumulative doctoral theses, 2018. URL [https://www.chemgeo.uni-jena.de/chemedia/Fakultät/Einrichtungen/Promotions\\_+und+Habilitationsstelle/Promotionsordnungen+und+Durchführungsbestimmungen+\\_+Doctoral+Examination+Regulations+and+implementing+provisions/Durchführungsbestimmungen+kumulative+Dissertationen\\_English\\_2019\\_01\\_16-p-2386.pdf](https://www.chemgeo.uni-jena.de/chemedia/Fakultät/Einrichtungen/Promotions_+und+Habilitationsstelle/Promotionsordnungen+und+Durchführungsbestimmungen+_+Doctoral+Examination+Regulations+and+implementing+provisions/Durchführungsbestimmungen+kumulative+Dissertationen_English_2019_01_16-p-2386.pdf).