How Many is too Many?
On the Relationship between Output and Impact in Research

Vincent Larivière[1] and Rodrigo Costas[2]

[1] *vincent.lariviere@umontreal.ca*
Université de Montréal, École de bibliothéconomie et des sciences de l'information, C.P. 6128, Succ. Centre-Ville, H3C 3J7 Montréal, Qc. (Canada) and Université du Québec à Montréal, Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Observatoire des Sciences et des Technologies (OST), C.P. 8888, Succ. Centre-Ville, H3C 3P8 Montreal, Qc. (Canada)

[2] *rcostas@cwts.leidenuniv.nl*
Leiden University, Center for Science and Technology Studies, Wassenaarseweg 62A, 2333 AL Leiden (The Netherlands)

Abstract
Over the last few decades, the massification of quantitative evaluations of science and their institutionalisation in several countries has led many researchers to aim at publishing as much as possible. This paper assesses the potential adverse effects of this behaviour by analysing the relationship between individual researchers' productivity and their proportion of highly cited papers. In other words, does the share of an author's top 1% most cited papers increase, decrease or remain stable, as her number of total papers increase? Using a large dataset of disambiguated researchers (N= 25,994,021) over the 1980-2012 period, this paper shows that the higher the number of papers a researcher publishes, the more likely they are amongst the most cited in their domain. This relationship was stronger for older cohorts of researchers, while decreasing returns to scale were observed in some domains for more recent cohorts. On the whole, these results suggest that at the macro-level, the culture of publishing as many papers as possible did not yield to adverse effects in terms of impact, especially for older researchers. For such researchers, who have had a long period of time to accumulate scientific capital, there can never be too many papers.

Conference Topic
Science Policy and Research Assessment

Introduction
In the second half of the 20th Century, but even more so over the last few decades, evaluations have become widespread in various spheres of society (Dalher-Larsen, 2011). Although scientific research has long been exempt from external evaluations thanks to Vannevar Bush and post WWII non-interventionist science policy, it has always been assessed internally through peer review. These means of evaluating research and researchers have, however, slowly changed since the 1980s, when researchers and administrators became aware of the roles that bibliometric analyses could play in such evaluations. Quantitative publication and citation analyses gained even more importance in the 2000s (Cameron, 2005), when tools for assessing individual researchers' output and impact became widespread. While in some cases, these methods have been developed to complement peer review in the allocation of research funding—such as the BOF-key in Flanders (Belgium) (Debackere & Glänzel, 2004), the Research Assessment Exercise/Framework in the UK—in other settings, these quantitative evaluations of research have become the main mean through which research is assessed and funded (Sörlin, 2007). Various publication-based and citation-based funding models can be found in Australia, Norway, Denmark, Sweden and Finland—and translates as the currency through which academic exchanges of tenure, promotion and salary raises are made (e.g. Fuyono & Cyranoski, 2006).
While there has always been subliminal bibliometrics performed through peer evaluation—as reviewers were skimming through reviewees' CVs through the process—the massification of

evaluations and their institutionalisation led many researchers and institutions to put large emphasis on the number of papers they published. This has led to adverse effects (Binswanger, 2015; Frey & Osterloh, 2006; Haustein and Larivière, 2014; Weingart, 2005). Indeed, like any social group, researchers are prone to change their behaviour once the rules of the games become explicit or what is expected from them; phenomenon that could be referred to as the Hawthorne effect (Gillespie, 1993), or to Goodhart (1975) or Campbell's laws (1979). As most evaluations and rankings are first based on numbers of published papers, this has created incentives for researchers to *author as many papers as possible.* In Australia (Butler, 2004), where publications counts were used without differentiating between publication venue or citations received, researchers have been found to increase their numbers of publications in journals with high acceptance rates and lower impact. Along these lines, the h-index, which together with the Impact Factor, is likely the most popular bibliometric indicator in the scientific community, is largely determined by numbers of papers published than on citations (Waltman & van Eck, 2012).

Within this context, researchers have adopted many publication strategies. While some researchers focus on publishing few, high-quality papers—e.g. 'selective' (Costas & Bordons, 2007) or 'perfectionists' (Cole & Cole, 1973)—others publish as many papers as possible, without not all of them necessarily being of high quality—e.g. 'prolific scientists' (Cole & Cole, 1973) or 'big producers' (Costas & Bordons, 2008)). However, little is known on the publication strategy that yields the highest results in terms of impact. In order to better understand the relationship between productivity and impact, this paper compares, for a large dataset of disambiguated researchers (N= 25,994,021), their total number of papers with the proportion of these papers that made it to the top 1% most cited of their field. Thus, this paper aims at answering the following key question: Does an authors' share of top papers start to decrease with a certain number of papers published? Or is it stable, as production and impact are two distinct dimensions of scientific activity. In other words, *how many is too many?* What is the probability for an author to publish top cited papers relate to the number of papers published? A good analogy for this is archery: if an archer throws one arrow, what is the probability that it hits the center of the target? Does an increase in the number of arrows thrown leads to an increase in the proportion of arrows hitting the center of the target?

Two opposite hypotheses could be made. The first one would be that authors with just 'average' production—rather than low or high production— are the ones more likely to publish top cited papers, as these authors, perhaps, focus more on the 'quality' of their output than just on quantity (i.e. selective scholars). The second hypothesis would be that, it is the authors with very high number of papers who, on average, publish the highest proportion of top cited papers. This hypothesis would be on agreement with the theory of Merton's cumulative advantages (Merton, 1968), and supported by empirical work in the sociology of science (Cole & Cole, 1973). Similarly, in a Bourdieusian framework, the main goal of a researcher is to increase its rank in the scientific hierarchy and gain more scientific capital (Bourdieu, 2004). If publishing a high number of scientific papers and being abundantly cited are the ways through which researchers can reach this goal, then they will adapt their behaviour to reach these evaluation criteria.

This focus on publishing as many papers as possible—often referred to as 'salami slicing'—has been long discussed (e.g. Abraham, 2000; Jefferson, 1998). However, only a few authors have analysed the effect of 'salami slicing' on papers' citations. For instance, Bornmann and Daniel (2007) have shown, for a small sample of PhD research projects in biomedicine (N=96), that an increase in the number of papers associated with a project lead to an increase in the total citation counts of papers associated with the projects. However, they do not show whether the impact of each paper taken individually increases with the number of papers published. Similar to this study, Hanssen and Jørgensen (2015) analysed the effect of ‘experience’ on papers’ citations; experience being defined as the author’s previous number of publications. Drawing a sample of papers in transportation research (N=779) they show that experience is a statistically significant determinant of individual papers’ citations, although this increase becomes marginal once a certain threshold is met in terms of previous papers published.

Methods

This paper uses Thomson Reuters’ Web of Science (WoS) for the period 1980-2012. Only journal articles are included. Given that the units analysed in this paper are individual researchers, we used the disambiguation algorithm developed by Caron & van Eck (2014) to identify the papers of individual researchers. On the whole, the algorithm managed to attribute papers to 25,994,021 individuals, which were divided into seven cohorts based on the year of their first publication (Table 1).

Table 1. Number of disambiguated researchers per cohort

Year of first publication	Number of researchers
<=1985	3,574,667
1986-1990	2,733,002
1991-1995	3,282,421
1996-2000	3,810,652
2001-2005	4,310,886
2006-2011	6,930,289
>=2012	1,352,104

As we want to assess researchers’ contribution to research that has the highest impact, we isolated for each discipline the top 1% most cited papers published each year (normalized by WoS subject categories). Citations are counted until the end of 2013, and exclude self-citations. The broad disciplines used are those of the 2013 Leiden ranking which are based on the assignment of WoS Subject Categories to five main domains (CWTS, 2013). Figures in the paper presents classes of numbers of papers in which there are at least 100 researchers.

Results

Figure 1 presents, for the oldest cohort studied—researchers who have published their first paper before 1985—the relationship between the number of papers throughout their career and the proportion of those papers that made it to the top 1% most cited. For any specific number of papers, the expected value of top 1% papers is, as one might expect, 1%. Researchers for all five domains have one thing in common: authors with very few papers are, on average, much less likely to publish high shares of top 1% most cited papers. For *Biomedical and health sciences* and for *Social sciences and humanities* we observe a continuous increase in authors’ proportion of top papers as their overall number of papers increases. For *Life and earth sciences* the share of papers does increase with the number of papers, until about 10 papers where they starts to oscillate, although in general an increasing pattern is still observed, especially after 40 papers. Perhaps the most deviant pattern is found in *Mathematics and computer science* where for just for the very low levels of production there is an increase in the share of highly cited publications, but this share decreases between

4 and 20 papers. It then starts to increase again for higher numbers of papers, despite important fluctuations. Natural sciences and engineering follow a similar pattern, with a decrease in the share of top papers between 6 and 30 papers, followed, in this case, by a clear increase until very high levels of productivity.

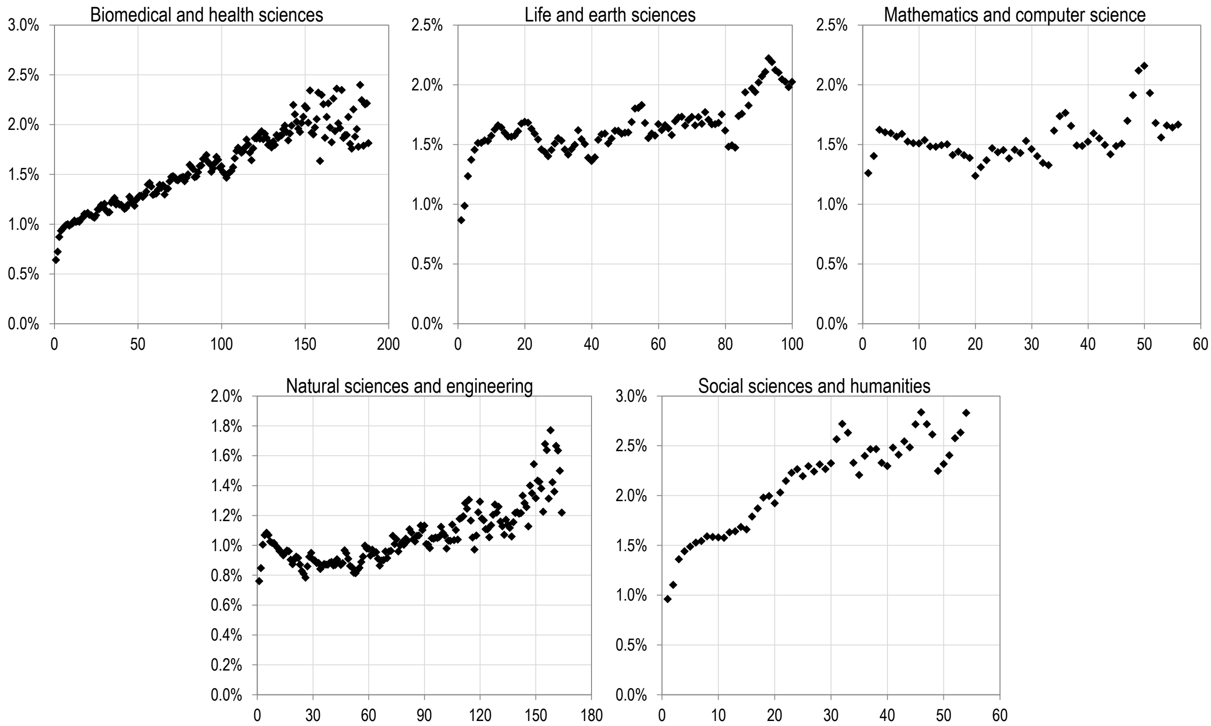

Figure 1. Proportion of top 1% most cited papers (*y* axis), as a function of the number of papers published (*x* axis), for the cohort of researchers who have published their first paper before 1985, by domain. Only classes of numbers of papers with 100 researchers or more are shown.

When researchers who have published their first paper between 2006 and 2011 are considered, different pattern are observed (Figure 2). For *Biomedical and health sciences* there is an increase in the share of highly cited publications up to around 15 publications, when some important fluctuations—or certain decreasing returns to scale—start to appear. A similar pattern is observed for the *Life and earth sciences* with the variability starting from levels of production of around 10 publications although, in this case, a decrease is clearly observed. For the other domains the pattern tends to be clearly increasing, although oscillations are also observed for the higher levels of production, which could also be seen as decreasing returns to scale. For the other three domains, there is clearly an increase in the share of top papers as the number of papers increases. However, we also observe for these three fields a decrease at very high levels of productivity.

An important characteristic of this cohort is that it got socialized to research recently—when the evaluation culture was more present—which might explain why they might be more prone to try to publish as much as possible. However, the drop in the share of top papers observed in each domain—although at different levels of productivity—suggests that these academically-younger scholars struggle to keep impact high once a certain threshold is met. This might be due to the fact that these scholars have not yet secured permanent or tenure positions and, thus, might feel that they cannot be as selective as older scholars who might choose their collaborators more easily.

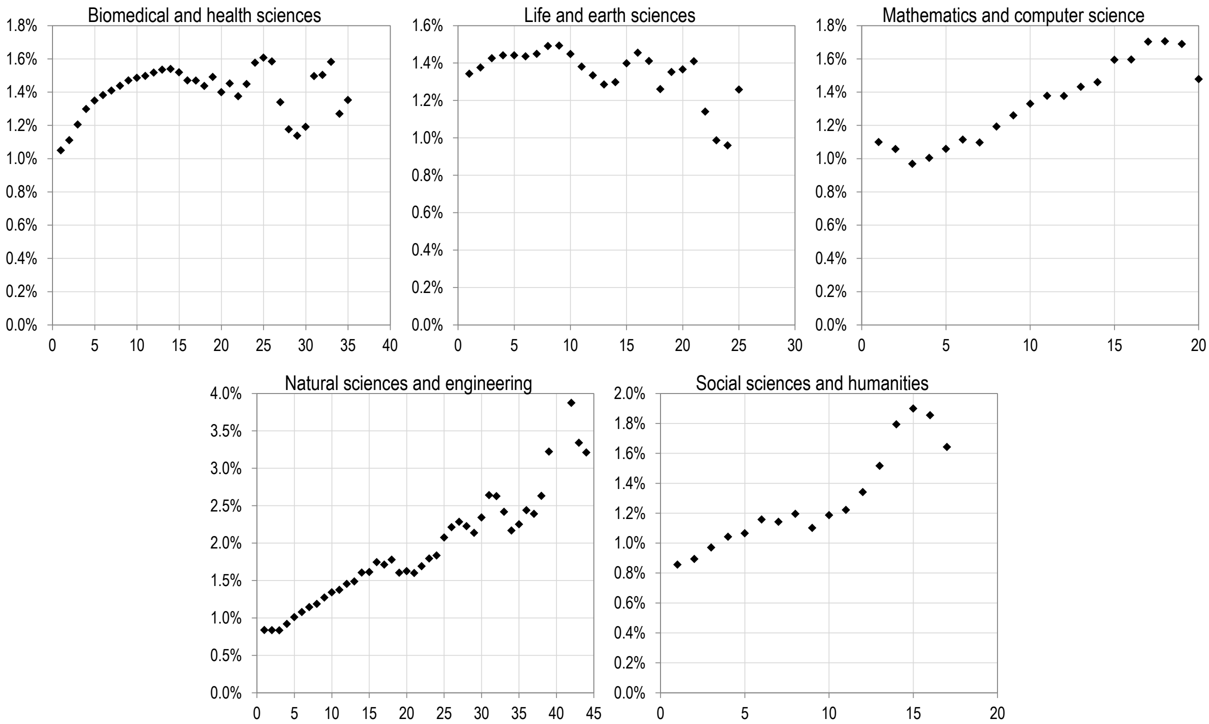

Figure 2. Proportion of top 1% most cited papers, as a function of the number of papers published, for the cohort of researchers who have published their first paper between 2006 and 2011, by domain. Only classes of numbers of papers with 100 researchers or more are shown.

Discussion and Conclusion

Previous research has shown that, in many contexts, the focus on indicators in research evaluation has had adverse effects, especially in terms of papers published (e.g. Binswanger, 2015). This paper aimed to provide an original analysis of one of these adverse effects, which is to aim to *publish as much as possible*. Our results have shown that, especially for older researchers, the higher the number of papers published throughout their careers, the higher the share of these papers ends up being amongst the top cited papers of their fields. This effect was higher for *Biomedical and health sciences* and for *Social sciences and humanities*, but in all fields the most active group of researcher was also having a higher share of top cited papers. A general exception to this trend was found in academically-younger researchers working in the field of *Life and earth sciences*, where higher scientific output was associated with lower impact than low-to-mid scientific output. Decreasing returns to scale were also more common for more junior researchers than senior ones.

These results conform to the Mertonian theory of cumulative advantages (Merton, 1968): the higher the number of papers an author contributes to, the more he or she gets known and, hence, is likely to attract citations. In Bourdieusian terms, the more an author publishes and accumulates citations in a domain, the more this capital will yield additional papers and citations. The relationship could also be in the other direction, as highly cited authors might have more opportunities to contribute to papers, given the scientific capital they have accumulated. Still, the results show that top cited authors do not only contribute on average to more papers, but also to more *highly cited* papers. On the whole, these results suggest that, at the macro-level, the culture of publishing *as many papers as possible* did not yield to adverse effects in terms of impact, especially for senior researchers. For such researchers, who have had a long period of time to accumulate scientific capital, there can never be *too many papers*.

References

Abraham, P. (2000). Duplicate and salami publications. *Journal of Postgraduate Medicine, 46*(2), 67.

Binswanger, M. (2015). How nonsense became excellence: Forcing professors to publish. In Welpe, I.M. et al. (eds). *Incentives and Performance* (pp. 19-32). Switzerland: Springer International Publishing.

Bornmann, L., & Daniel, H. D. (2007). Multiple publication on a single research study: does it pay? The influence of number of research articles on total citation counts in biomedicine. *Journal of the American Society for Information Science and Technology*, *58*(8), 1100-1107.

Bourdieu, P. (2004). *Science of Science and Reflexivity*. Cambridge, UK: Polity Press.

Butler, L. (2003). Modifying publication practices in response to funding formulas. *Research Evaluation*, 12(1), 39-46.

Cameron, B. D. (2005). Trends in the usage of ISI bibliometric data: Uses, abuses, and implications. *portal: Libraries and the Academy*, *5*(1), 105-125.

Campbell, D. T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, *2*(1), 67-90.

Caron, E., & van Eck, N. J. (2014). Large scale author name disambiguation using rule-based scoring and clustering. In *19th International Conference on Science and Technology Indicators. Context Counts: Pathways to Master Big Data and Little Data* (pp. 79-86). CWTS-Leiden University Leiden.

Cole, J. R., & Cole, S. (1973). *Social Stratification in Science*. Chicago: University of Chicago Press.

Costas, R., & Bordons, M. (2007). The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics*, *1*(3), 193–203.

Costas, R., & Bordons, M. (2008). Is g-index better than h-index? An exploratory study at the individual level. *Scientometrics*, *77*(2), 267–288.

CWTS. (2013). *Leiden Ranking 2013 Methodology*. http://www.leidenranking.com/Content/CWTS%20Leiden%20Ranking%202013.pdf

Dahler-Larsen, P. (2011). *The Evaluation Society*. Palo Alto, CA: Stanford University Press.

Debackere, K., & Glänzel, W. (2004). Using a bibliometric approach to support research policy making: The case of the Flemish BOF-key. *Scientometrics, 59*(2), 253-276.

Frey, B. S., & Osterloh, M. (2006). *Evaluations: Hidden Costs, Questionable Benefits, and Superior Alternatives*. Institute for Empirical Research in Economics, University of Zurich.

Fuyuno, I., & Cyranoski, D. (2006). Cash for papers: putting a premium on publication. *Nature*, *441*(7095), 792-792.

Gillespie, R. (1993). *Manufacturing Knowledge: A History of the Hawthorne Experiments*. Cambridge, New York: Cambridge University Press.

Goodhart, C.A.E. (1975). *Problems of Monetary Management: The U.K. Experience*. Papers in Monetary Economics (Reserve Bank of Australia).

Hanssen, T. E. S., & Jørgensen, F. (2015). The value of experience in research. *Journal of Informetrics, 9*(1), 16-24.

Haustein, S., & Larivière, V. (2015). The use of bibliometrics for assessing research: possibilities, limitations and adverse effects. In Lempe, I.M. et al. (eds.) *Incentives and Performance* (pp. 121-139). Switzerland: Springer International Publishing.

Jefferson, T. (1998). Redundant publication in biomedical sciences: Scientific misconduct or necessity? *Science and Engineering Ethics*, *4*(2), 135-140.

Merton, R. K. (1968). The Matthew effect in science. *Science, 159*(3810), 56-63.

Sörlin, S. (2007). Funding diversity: performance-based funding regimes as drivers of differentiation in higher education systems. *Higher Education Policy*, *20*(4), 413-440.

Waltman, L., & Van Eck, N. J. (2012). The inconsistency of the h-index. *Journal of the American Society for Information Science and Technology*, *63*(2), 406-415.

Weingart, P. (2005). Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics*, *62*(1), 117-131.