A Model for Predicting Citation Enhancement / Maor Weinberger

Abstract

Introduction:

Assessing the current and predicting the future academic performance of scholars is important, since the performance determines the researcher's status in the world of academia (Pan & Fortunato, 2014; Waltman, 2016) and also influences decisions related to academic promotions and allocation of funds (Way et al., 2017). Academic *performance* evaluation is based primarily on measurement of the amount of publications, namely the scholar's *productivity*, and the number of citations, namely the *impact* (Sinatra et al., 2016). One of the main concerns of individual researchers and institutions is how to assess the future performance of scholars and identify their potential to become successful scientists. Scholarly performance is directly assessed by the scholar's publication and citation rates, and various combinations of these two parameters, such as the standard *h-index* measure - an author gets an h score when he has published h articles, each cited at least h times, while h is the maximal value (Hirsch, 2005). Therefore, recent studies aimed at predicting the future *h-index* or citation rate of a scholar used his/her past citation data as well as productivity and various social and demographic factors. However, citation and *h-index* scores are prone to biases and inconsistencies, and consequently do not always provide a fair and comprehensive evaluation of scholarly success.

Research objectives:

In this study, we present a new method for systematic investigation of the diversity in academic performance and factors influencing it among successful scholars. We identified the most successful scholars (i.e. most cited scholars) and their common traits, traits that distinguish them from their less successful peers, in order to predict the potential for future success of academic scholars. The differences between these two groups of scholars were examined using two sets of measures: the conventional scholarly performance evaluation metrics (e.g. the total number of citations and *h-index*) and measures that quantify the structure of their citation trajectories. We also investigated whether and to what extent these measures are influenced by gender, seniority, affiliation and field of study, and examined the potential effect of citation indices on the scholarly performance evaluation. Next, we modeled scholarly success in terms of the probability of a scholar belonging to a group of highly impactful scholars as determined by their citation trajectory structures, and applied the proposed models to predict the scholars' potential for future academic success.

To this end, we developed a new set of impact measures based on a scholar's citation trajectory structure (rather than on absolute citation, publication or *h-index* rates), that showed a stable trend and scale for those highly impactful scholars, independent of their field of study, seniority and citation index. The newly introduced measures were then incorporated as influence factors into the logistic regression models and probabilistic classifiers to identify the successful scholars in the heterogeneous corpus of the most and the least cited professors from two Israeli universities. The productivity and impact rates of 663 tenured professors, sampled from six faculties in two Israeli

universities, were collected from the two major citation indices: Web of Science (WoS) and Google Scholar (GS). A comparison was conducted among the highest impact, lowest impact and average impact scholars in the corpus for each citation index.

Results:
Significant differences were found between the scholars' performance rates in different impact-level groups in the two indices. The top performing group comprised 44 scholars who belonged to the highest impact sub-corpora according to both citation indices. Linear regression analysis found that women, despite being a minority in Israeli academia, outperformed men in terms of scientific impact. Interestingly, there were several differences among the two indices in terms of seniority and performance rates. Our findings provide evidence for the "rich get richer" phenomenon in GS compared to WoS. In WoS, mean performance rates stabilize after 15 years of seniority, while the GS performance rates of scholars constantly grow over time. Overall, 79-100% of the scholars in the highly-impactful group were classified as successful scholars according to the trajectory-based models. The models were primarily based on the proportions of incline and decline in the growth rates of the citation slope throughout the entire career of the scholar. The most influencing factors in determining the scholar group were: the total number of positive growth years and the difference between the number of positive growth years and the number of negative growth years throughout the scholars' career, particularly in the first 10-15 years of their career. The highly-impactful group had significantly higher average slope growth and longer incline periods compared to the decline periods than lowly-impactful scholar group. These differences were even more acute in the early to mid-stages of their careers. This is indicative of an overall more stable and positively directed career trajectories among the scholars in this group, compared to the ones in the lowly-impactful group, who had more skewed career trajectories.
However, we wish to note that the citation rates were not normalized by field of study or seniority (as suggested in the literature), due to unavailability of average citation rates for each field. Thus, there may still be some bias in the four sub-groups towards certain faculties and seniority levels (e.g. more Social Science and young scholars in the LI100 group than in the HI100 group).

Research Implications:
From the practical point of view, the study may yield useful insights and serve as an aid to academic institutions when making promotion decisions, as well as a self-assessment tool for researchers who strive to increase their academic influence and become leaders in their field. The study also contributes to the evaluation of scholarly success and performance diversity in the academic community. The research findings were published in three peer-reviewed venues: one journal publication in Scientometrics (Weinberger & Zhitomirsky-Geffet, 2021) and two conference proceedings: iConference 2020 (Weinberger et al., 2020a) and ASIS&T 2020 (Weinberger et al., 2020b). The fourth publication was recently submitted and is currently under review in Scientometrics.

MMS Number: 9926723029205776