Journal of Informetrics 11 (2017) 598–607

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Do mathematicians, economists and biomedical scientists trace large topics more strongly than physicists?

Menghui Li[a], Liying Yang[b], Huina Zhang[a], Zhesi Shen[c], Chensheng Wu[a], Jinshan Wu[c,*]

[a] Beijing Institute of Science and Technology Information, Beijing 100044, PR China
[b] National Science Library, Chinese Academy of Sciences, Beijing 100190, PR China
[c] School of Systems Science, Beijing Normal University, Beijing 100875, PR China

ARTICLE INFO

Article history:
Received 20 December 2016
Accepted 9 April 2017
Available online 15 May 2017

Keywords:
Research choice
Matthew effect
Large topics

ABSTRACT

In this work, we extend our previous work on largeness tracing among physicists to other fields, namely mathematics, economics and biomedical science. Overall, the results confirm our previous discovery, indicating that scientists in all these fields trace large topics. Surprisingly, however, it seems that researchers in mathematics tend to be more likely to trace large topics than those in the other fields. We also find that on average, papers in top journals are less largeness-driven. We compare researchers from the USA, Germany, Japan and China and find that Chinese researchers exhibit consistently larger exponents, indicating that in all these fields, Chinese researchers trace large topics more strongly than others. Further correlation analyses between the degree of largeness tracing and the numbers of authors, affiliations and references per paper reveal positive correlations – papers with more authors, affiliations or references are likely to be more largeness-driven, with several interesting and noteworthy exceptions: in economics, papers with more references are not necessary more largeness-driven, and the same is true for papers with more authors in biomedical science. We believe that these empirical discoveries may be valuable to science policy-makers.

1. Introduction

How researchers choose their research topics has received sustained attention (Busch, Lacy, & Sachs, 1983; Diamond, 1994; Foster, Rzhetsky, & Evans, 2015; Gieryn, 1978; Merton, 1938; Zuckerman, 1978). This will not only directly affects scientists' output and recognition, but also indirectly affect the science itself. This problem has been widely investigated from sociology of science (Latour, 1987; Merton, 1957, 1968, 1973), philosophy of science (Bikard, Murray, & Gans, 2015; Boyer-Kassem & Imbert, 2015; Kitcher, 1995; Kleinberg & Oren, 2011; Strevens, 2003, 2006; Zollman, 2009), and new economics of science (Dasgupta & David, 1994; Stephan, 1996). Quite often previous investigations on this issue are rather qualitative. In this work, we want to study this empirically based on data.

Researchers may determine their research topics primarily according to their research interests, their perceived potential in making progress on the topics or simply by the largeness of the topics, or some combination of all these factors (Foster et al., 2015). Here, largeness refers to how many publications are on the topic during a time interval denoted as $[t_0, t]$.

* Corresponding author.
E-mail address: jinshanw@bnu.edu.cn (J. Wu).

http://dx.doi.org/10.1016/j.joi.2017.04.004
1751-1577/

For example, based on knowledge network of chemical reactants, Rzhetsky et al. explore the probability of selecting a pair of molecules as a function of the importance (represented by degree) and the difficulty associated with combining them (measured by network distance). It is found that biomedical scientists prefer to explore the local neighborhood of central, important (i.e., higher-degree) molecules in biomedical chemistry (Rzhetsky, Foster, Foster, & Evans, 2015). In this work, we focus on the effect of largeness of topics on scientists' choices of topics. It has been found that the scientists tend to trace large topics in physics (Wei et al., 2013), and in environmental science (Grandjean, Eriksen, Ellegaard, & Wallin, 2011).

"The rich get richer" or Matthew effect is a common social phenomenon (Barabási & Albert, 1999; Price, 1976; Simon, 1955). Matthew effects in science have been investigated also in scientometrics, for instance, on scientists' credit (Merton, 1968) and on citations (Biglu, 2008; Bonitz, Bruckner, & Scharnhorst, 1997; Khosrowjerdi, Zeraatkar, & Hajipour, 2012). Only a few previous studies are on scientists' choice of topics in a certain discipline (Grandjean et al., 2011; Rzhetsky et al., 2015; Wei et al., 2013). Thus, in this paper we ask: Does tracing large topics differ across disciplines? What is the difference of degree of largeness tracing among different countries? To what degree is the intensity of largeness tracing relevant to properties such as the number of authors, references and affiliations of articles? We hope that discoveries from this study will provide valuable information for scientific policy makers especially on the issue of funding and evaluation.

The question of hot topics, prominent topics, or research fronts itself has also been widely investigated in scientometrics (Boyack & Klavans, 2010; Chen, 2004, 2006; Cozzens et al., 2010; Klavans & Boyack, 2017; Small, Boyack, & Klavans, 2014; Upham & Small, 2010). Research fronts are generally defined as the areas attracting the most scientific interest in a period of time, especially before publication size of the considered field becomes really big and the field is clearly under exponential growth. Therefore, the studies of hot topics conventionally refer more to identifying emerging hot fields. We will refer these studies as studies on emerging hotness. One might use the number of received citations of papers on a topic (Boyack & Klavans, 2010; Chen, 2004, 2006; Klavans & Boyack, 2017; Small et al., 2014; Upham & Small, 2010), or simply use the number of publications (or scientists) on a topic (Cozzens et al., 2010) for this purpose. In this work, we do not need to identify emerging research fronts and our use of hotness or largeness is more like the current size (in terms of number of publications) of the fields. We call the hotness in this sense the realized hotness, or simply largeness.

Furthermore, for each of the four disciplines that we study, which are mathematics, physics, economics and biomedicine, there is an established hierarchical classification system of topics and all the publications under our investigation have been annotated with such a hierarchical codes from the system. Therefore, we do not even need to do classification or clustering of publications to identify topics. In this work, articles with same code at a certain level are considered to be on the same topic. Both the number of papers and the number of received citations of the papers on a topic can be used to quantify the largeness of the topic. However, these two quantities are not really independent: the number of citations more or less follows a power law relation with respect to the number of papers (Dong, Li, Liu, Wu, & Wu, 2017; Katz, 1999). In the following analysis, we use only the number of publications belonging to a topic as a measure of the largeness of the topic as for example in Grandjean et al. (2011), Rzhetsky et al. (2015). In principle, one can also consider how new publications are attracted by received citations of topics, but this will be the topic of another investigation. Thus, using established hierarchical classification systems of topics and classifications of papers in these systems, our definition of large topics in this paper is much simpler than that of research fronts: We only need to count how many papers belong to a topic at a given level. The more papers, the larger the topic and the bigger the realized hotness. Using the established hierarchical classification systems of topics is not ideal since new topics emerge constantly. One may use various clustering methods based on citation relation among publications to establish classification systems and identify topics (Boyack & Klavans, 2010; Klavans & Boyack, 2017; Subelj, van Eck, & Waltman, 2016; Waltman & van Eck, 2012). However, in this work, we will use the established coding systems based on controlled vocabularies for the above four disciplines.

Using the above measure of largeness, in this work, we investigate the degree of largeness tracing in mathematics, economics and biomedical science, and then perform a comparison among them and also between these three fields and physics. Intuitively, it might be expected that mathematicians would be more likely to choose their topics of investigation according to their scientific interests and the scientific value of the questions, partly due to the fact that mathematicians intend to work individually and partly since often mathematicians claim so. One might also guess that it is possible that biomedical scientists choose research topics mainly according to medical values of problems rather than the largeness of topics. Economists might choose their topics according to their urgency to the current economic situation rather than their largeness. Here, we empirically examine whether or not this is the case.

Aside from pure curiosity, such a comparison among various fields might have value in the study of science policy. For example, we wish to examine the correlation between the impact of papers and their degree of largeness tracing. As researchers, we would prefer to see that less largeness-driven papers have a greater impact. For science policy makers and administrators of universities, funding agencies and institutes, such a correlation, whether positive or negative, could potentially provide guidance regarding their duties, as one of their goals is to seek strategies to improve the scientific impact of the academic units under their administration. We would also like to examine whether larger teams tend to produce more value-driven papers or more largeness-driven papers. An answer to this question might have strategic policy value regarding how large research teams should be supported.

In our previous work, using data obtained from the American Physical Society (APS) concerning publications in APS journals, we investigated and confirmed the occurrence of largeness tracing in physics (Wei et al., 2013). In this work, we extended this analysis to the fields of mathematics, economics and biomedical science based on publications from the datasets of the American Mathematical Review (AMR), the Journal of Economic Literature (JEL) and PubMed, respectively.