Annals of Library and Information Studies
Vol. 67, March 2020, pp. 17-35

Revisiting De Solla Price: growth dynamics studies of various subjects over last one hundred years

Soumen Teli[a] and Bidyarthi Dutta[b]

[a]Librarian, Shyampur Siddheswari Mahavidyalaya, Ajodhya, Howrah 711 312, West Bengal &
Research Scholar, Dept. of Library and Information Science, Vidyasagar University, Midnapore-721 102, West Bengal, India,
E-mail: soumenn165@gmail.com

[b]Assistant Professor, Department of Library and Information Science, Vidyasagar University, Midnapore 721 102, West Bengal, India,
E-mail: bidyarthi.bhaswati@gmail.com

The stalwart historian of science, Derek J. De Solla Price delivered a series of lectures at Brookhaven National Laboratory in 1962, which was dedicated to discussing science and its interaction with society. The collection of these lectures was published in 1963 as a book entitled *Little Science, Big Science*. Here, Price empirically established that the initial exponential growth pattern of literature reaches a ceiling after a certain time span, which results a logistic growth pattern. This paper analyses Price's empirical theory on the basis of 198 articles that presented growth of literature of variant subjects published since 1913 to 2018. In all, 214 growth models were reported by the 198 articles that analysed growth of literature of more than 50 subjects. It is found that growth patterns reported by nearly 50% articles followed Price's empirical theory, i.e., exponential and logistic growth pattern while remaining 50% articles followed other growth patterns, viz., power model, linear model, etc. All growth models reported by the 198 articles were broadly categorised into five groups on the basis of statistical characteristics, viz., (exponential + logistic), growing without definite pattern (GWDP), linear, non-linear and decaying models. The null hypothesis formulated states that 214 growth models observed by different subjects described in 198 articles will follow either of the five patterns that will be guided by Bradford's Law of Bibliographic Scattering. The null hypothesis is accepted by Chi-square test. It is inferred that the distribution of different models of growth of literature is guided by Bradford's Law where the core or nucleus zone is occupied by the logistic and exponential model, i.e. Price's empirical model prevails in Bradford's nuclear (core) zone.

Keywords: Growth of literature; De Solla Price; Literature growth; Exponential growth; Logistic growth

Introduction

The term 'growth' indicates an augmentation in original size, which implies a change of state or size. The concept of growth holds extensive spectrum of connotations, for instance, cell growth, bacterial growth or organism growth in the context of biological sciences. The domain auxology covers all aspects of physical growth in the context of human physiology, growth of resource, goods, market and services in the context of economics, commerce and management etc. The patterns of growth curves are described in mathematics by different names, i.e. exponential, power, linear, logistic, hyperbolic etc. The growth of primary or secondary sources of information belonging to any subject domain over time is being studied since 1913. As the sources of information of any subject area is known as the literature, this kind of study is also popularly known as "study of growth of literature". The scope of this study is normally defined under bibliometrics, informetrics or scientometrics. This kind of study achieved special significance particularly after De Solla Price' masterpiece entitled *Little science, Big science* that he published in 1963[1].

The theoretical foundation of growth dynamics study of literature was laid down by De Solla Price in this book on the basis of simple logical analysis. It was shown argumentatively with aid of few observations that the growth pattern follows exponential graph initially with a ceiling that after a certain time span that is different for different subject domains. The resultant curve as a consequence acquires the logistic 'S'- shaped pattern. Recent studies on scientific growth focus mainly on two aspects— increase in scientific manpower and increase in the stock of scientific knowledge. The number of science periodicals including abstracting sources are the basic indicators of scientific growth.

Growth of scientific literature and knowledge studies are highly interdisciplinary in nature, and significant contributions are from library and information science field, as is evident from the bibliographic databases. This paper surveyed a sample of 198 relevant papers to explore 214 growth models of various subjects. The hypothesis is formulated on the basis of Price's theory and the same has been tested on the basis of practical data obtained from the collected sample.

Review of literature

Many studies in information science (and other subjects) have investigated the growth of science[2,3,4]. The systematic study of growth and obsolescence of literature of any subject is termed as “the study of literature dynamics”. Tabah[5] stated, “the information science approach is to follow the published literature and infer from the growth of the literature the movement of ideas and associations between scientists”. Besides *Little science, Big science*, Price undertook many significant works on literature dynamics in the years 1961, 1951 and 1965[6]. He analyzed the references listed in the 1961 edition of the Science Citation Index (SCI, Thomson Reuters) and the research papers published in the *Philosophical Transactions of the Royal Society of London*. His results show that science is growing exponentially in a certain period by a certain percentage and doubles every 10 to 15 years. The exponential growth in science established by Price has become today a generally accepted theory which has also been confirmed by other studies[6,7]

The hypergraph model was proposed to represent generalised network of literature of science, where the papers were considered as hypergraph nodes[8]. Kwiek[9] studied systematic inequality in knowledge production as argued by Lotka and Price across highly productive academics in 11 European countries. The academic attitudes, behaviours and perceptions as predictors of becoming top performers across European systems were also incorporated. Urban[10] analysed social, political and cultural impact on growth of science through regression analysis. Price’s theory of differences among the sciences included three important points in regard to knowledge in science. It was analysed through citation context studies that described the process of knowledge building[11,12]. These papers described Price's tool to describe and compare differences among the sciences in their processes of knowledge growth along with the continuous change of sciences under the influence of new instruments or new sponsors. The rate of growth of science and the increase of obsolescence with age of scientific papers was observed by Gilbert et al[13].

Gilbert[14] reviewed a number of indicators of the growth of science to assess their strengths and weaknesses. The study focused on the problems involved in measuring two aspects of scientific growth—growth in manpower and growth in knowledge. Dedijer[15] found that since 1945, a rapid growth is seen in the disciplines of sociology, philosophy, psychology and history of science. An index for its socioeconomic development was introduced on the basis of quantitative data on a nation's research effort. Brookes[16] carried out simultaneous growth, utility and obsolescence study of scientific periodical literature (1970). The study of Tague, Beheshti and Lorna[17] showed that the innovative features of an article are reflected through citation counts as predicted by Price and other bibliometricians. In 1963, Price said[1], “There is a possibility that the exponential law is breaking down”. Exponential growth cannot go on forever. The study concluded that growth studies of literature need to become more exact in the description of their models and more rigorous in the application of statistical tests to determine how well these models fit reality.

Fernandez-Cano[18] conducted a study to analyze Price's model of scientific growth. The study showed an integrative review using retrieved empirical studies that exposes the complexity and diversity of models of scientific growth and the absence of consistent patterns. Szydlowski & Krawiec[19] discussed the concepts of knowledge and its accumulation used in economic growth theory. They applied differential equations to model the evolution of science including additional aspects such as the death of results, the time required to learn or to apply results to new discoveries. Heinzkill[20] analyzed 9556 footnotes in 15 different journals. The study showed that about 70 percent of all material cited is over ten years old. Meadows[21] verified that the overgrowth had previously been acknowledged in the 19th century, provoking exasperated reactions due to the declining readability of scientific literature.

Many works are based on Price's classic, *Little Science, Big Science*, usually abbreviated as LSBS. Lievrouw[22] discussed the possibility of comeback of little science modes of communication contrasting big science conventions dominating research policy, scientific institutions, and the publishing industry.

The growing use of more participatory, interactive "Web 2.0" technologies and social media in science today (e.g. wikis, blogs, tagging and bookmarking, conferencing, etc.) may signal such possibilities. Furner[23,24]carried out genesis study of LSBS in the context of the of science in the UK and the USA in the late 1950s. He showed that Price's ideas were formulated during a pivotal period in the development of socio-historical studies of science.

Andersen and Hammarfelt[25] studied the production of dissertations in eight research fields in the natural sciences, the social sciences and the humanities on the basis of Price's theory which used PhD dissertations as one of several indicators of scientific growth. Glänzel and Schoepflin[26] said, "Since the beginning of the eighties, bibliometrics has evolved into a distinct scientific discipline with a specific research profile, several subfields and the corresponding scientific communication structures (publication of the international journal *Scientometrics* in 1979 as the first periodical specialised on bibliometric topics). The funding of big projects seems to have become the regular way of financing research in scientometrics. Thus, from "Little Scientometrics" the field has become "Big Scientometrics"." Price's idea of transitional phase of science research from 'little science' to 'big science' is reflected in Glanzel's paper in the context of scientometrics/ bibliometrics.

Growth dynamics study: objectives and limitations

The number of articles published in science periodicals including abstracting periodicals are simple indicators of scientific growth. Price[1] argued that scientific literature over the years show exponential growth pattern and calculated the growth rate as 5% over the eighteenth and nineteenth centuries. He observed that once in fifteen years science literature doubled[27,28,29]. Neelameghan[30] analysed the documents on the history of medicine in India over the period 1954-61. The notable point was that during the period, Indian contribution was 65% and foreign contribution was 30%. He also studied the coverage of Indian medical literature in *Index Medicus* and *Excerpta Medica* and it was found that they covered only 38% and 13.5% of the Indian literature respectively. There are number of articles published on this topic, particularly on the growth of literature in different subjects. These articles chiefly focus the following four issues, i.e., numerical growth of literature and its temporal variation, obsolescence studies, coverage by Science Citation Index (SCI) and other indexing & abstracting databases, and analysis of growth pattern to theorise different growth models based on mathematical functions. This paper focuses on the last point, i.e. theorising growth models.

The empirical theory enunciated by De Solla Price is tested here on the basis of 198 articles that lead to growth of literature of variant subjects published since 1913 to 2018. These articles reported 214 growth models of more than 50 subjects over the years. The research problem is to testify to what extent Price's empirical theory is followed by the concerned subjects. The next objective is to carry out the growth dynamics study of literature on growth of literature of various subjects and to find out the specific subjects considered till date to carry out growth dynamics study.

Methodology and sample collection

In all, 198 articles on growth dynamics published since 1913 to 2018 are collected to find out growth models of the concerned subjects discussed therein. The complete bibliographic details of these 198 articles are given in Annexure I, which comprise the sample for this study. The thorough inspection of these 198 articles instantly categorises the growth models observed therein as follows, i.e. exponential (35%), irregular (31%), logistic (15%), linear (9%), power (5%), decaying (2.3%), epidemic (2%), Gompertz (1%) and logarithmic (0.5%) (Fig. 1).

Nearly one-third (31%) of the articles followed no definite mathematical function that indicates high empirical nature of the subject domain. Of these, epidemic model may be classed under exponential model as it indicates the sharp exponential growth. As the initial part of the logistic growth is exponential, an exponential graph may be considered

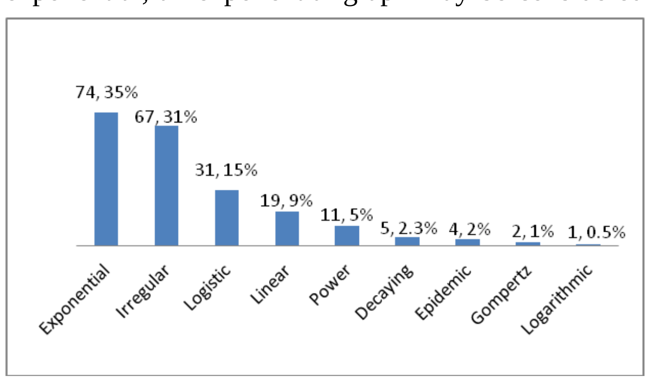

Fig. 1 — Growth models reported in 198 articles

as a component of a logistic graph. The continuing exponential growth results absurd conclusion that is practically impossible. The logistic curve actually limits the exponential growth curve by placing a ceiling of saturation at the tail of it.

The growth models following either of exponential, or logistic patterns are categorised under (Exponential + Logistic) model. The power, Gompertz and logarithmic models are categorised under non-linear model. The irregular growth model indicates the growth pattern following no definite mathematical function or erratic growth most likely resulting in unpredictable inference and are classed under Growing Without Definite pattern (GWDP). Besides, five articles reported negative growth or decaying of literature. The growth models of 198 articles are finally analysed under five categories, i.e. Decaying, Non-Linear, Linear, Growing Without Definite Pattern (GWDP) and (Exponential + Logistic). Some articles reported more than one growth model resulting in the 198 articles belonging to 214 models (Table 1).

Testing of hypothesis formulated

Null hypothesis is that the 214 growth models observed by different subjects described in 198 articles will follow either of the five patterns that will be guided by Bradford's Law of Bibliographic Scattering[31], which estimates the exponentially diminishing returns of searching for references in science journals. It is also stated as if journals in a subject domain are sorted by number of articles into three or more groups, each with about one-third of all articles, then the number of journals in each group will be proportional to $1:n:n^2:n^3$...... [32] Thus, according to null hypothesis, the ratio of 214 growth models described in 198 articles will follow Decaying, Non-Linear, Linear, Growing Without Definite Pattern (GWDP) and (Exponential + Logistic) patterns will be in the ratio, 1:2:4:8:16 (Taking n=2). It is the minimum possible ratio as per Bradford's Law as the minimum possible integral value of 'n' is 2. The total frequency in this case is 1+2+4+8+16 = 31, and the expected frequencies are: (16/31)*214=110, (8/31)*214=55, (4/31)*214=28, (2/31)*214=14 and (1/31)*214=7. As the (exponential + logistic) patterns are logically established by Price's theory, it is taken as most likely model whereas the decaying pattern is taken as most unlikely model as it is just opposite to growth function.

As $\chi^2 = \sum\{(f_o-f_e)^2/ f_e\}$, where f_o = Observed frequency and f_e = Expected frequency, the value of χ^2= 6.079 (Table 1). As all reported growth patterns are grouped into five growth models, therefore the number of classes is five, and the degrees of freedom is (5-1) = 4. Since the observed value of χ^2 (viz. 6.079) is less than the tabulated value 13.28 at 1% for four degrees of freedom, therefore the null hypothesis cannot be rejected at 1% level of significance. The conclusion is that the data are in agreement with the hypothesis that the ratio of 214 growth models described in 198 articles will follow Decaying, Non-Linear, Linear, Growing Without Definite Pattern (GWDP) and (Exponential + Logistic) patterns in the ratio at per Bradford's law.

Analysis

The first article that reported growth of literature on yeast was published in 1913 in German. In all, 198 articles were published since 1918 to 2018, the number of publications (frequency) in different years are presented in Table 2. The cumulative frequencies are also presented. The regression analysis of all observed cumulative frequencies data yielded the polynomial graph, i.e.

$y = a*x^4 + b*x^3 + c*x^2 + d*x + e$, where a, b, c, d and e are constants. The values of these constants are: $a = -4.325660102\cdot10^{-6}$; $b = 1.154595826\cdot10^{-3}$; $c = -7.295568847\cdot10^{-2}$; $d = 1.868687038$ and $e = -6.154763858$.

Figure 2 represents the frequency-time graph based on the data in Table 2. The continuous line represents the expected graph and the dots represent the observed values. The Residual Sum of Squares (RSS) = 500.1483258 and the Coefficient of Determination: R^2 = 0.9979708133. As the observed values are in close proximity of the expected values, it may be asserted that the growth of literature on growth dynamics studies follows polynomial pattern.

Table 1 — Observed and expected frequencies of growth models

Growth model	(Exponential + Logistic)	Growing Without Definite Pattern (GWDP)	Linear	Non-Linear	Decaying	Total
Frequency (Observed, f_o)	109	67	19	14	5	214
Frequency (Expected, f_e)	110	55	28	14	7	214

Table 2 — Publication timeline of articles that have reported growth of literature

Year	Frequency (No. of articles published)	Cumulative Frequency (Observed)	Cumulative Frequency (Expected)	Year	Frequency (No. of articles published)	Cumulative Frequency (Observed)	Cumulative Frequency (Expected)
1913	1	1	-4.4	1985	6	71	67.8
1917	1	2	1.5	1986	2	73	70.8
1923	1	3	7.0	1987	1	74	73.8
1927	1	4	9.1	1988	3	77	77.0
1929	1	5	9.8	1989	2	79	80.2
1930	1	6	10.1	1990	3	82	83.5
1931	2	8	10.4	1991	3	85	86.9
1934	1	9	10.9	1992	7	92	90.4
1935	4	13	11.1	1993	4	96	93.9
1937	1	14	11.3	1994	2	98	97.6
1938	1	15	11.4	1995	1	99	101.2
1939	1	16	11.5	1996	2	101	105.0
1947	1	17	12.9	1997	3	104	108.8
1949	1	18	13.5	1998	5	109	112.7
1952	1	19	14.7	1999	7	116	116.7
1957	1	20	17.7	2000	6	122	120.7
1960	1	21	20.2	2001	5	127	124.8
1963	1	22	23.3	2002	2	129	129.0
1966	2	24	27.0	2003	5	134	133.2
1969	1	25	31.5	2004	3	137	137.5
1970	5	30	33.1	2005	5	142	141.8
1971	4	34	34.9	2006	6	148	146.1
1972	3	37	36.7	2007	4	152	150.5
1973	1	38	38.5	2008	5	157	155.0
1974	2	40	40.5	2009	4	161	159.5
1975	1	41	42.6	2010	8	169	164.0
1976	1	42	44.7	2011	3	172	168.6
1977	4	46	46.9	2012	4	176	173.2
1978	3	49	49.2	2013	3	179	177.8
1979	4	53	51.6	2014	2	181	182.5
1980	3	56	54.1	2015	5	186	187.1
1981	3	59	56.7	2016	2	188	191.8
1982	2	61	59.3	2017	7	195	196.5
1984	4	65	64.9	2018	3	198	201.2

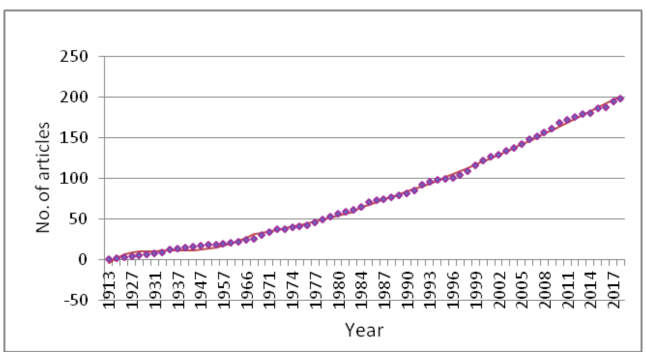

Fig. 2 — Cumulative number of articles published from 1913 to 2018

The subject domains of the 198 articles are listed in Table 3. The Dewey Decimal Class numbers up to second summary of the concerned subjects are also given with respective frequencies and percentages. The variation of subject domains are presented in Fig. 3 while the same in accordance with broad disciplines are presented in Fig. 4. It has been found that largest number of growth dynamics studies were performed in pure sciences (16.2%), followed by medical science (13.1%), life science (11.1%), chemistry (9.1%) and physics (8.6%). Other notable subject areas are engineering science, library and

Table 3 — Subject domains of 198 articles

Subject domains	DDC Class No. (2nd summary)	Frequency & Percent
Computer Science	004	1 (0.5%)
Library & Information Science	020	11 (5.6%)
Psychology	150	1 (0.5%)
Philosophical Logic	160	1 (0.5%)
Social Science	300	6 (3.0%)
Sociology & Social Anthropology	301	3 (1.5%)
Economics	330	2 (1.0%)
Social Service & Social Problems	360	1 (0.5%)
Education	370	1 (0.5%)
Pure Science	500	32 (16.2%)
Mathematics	510	6 (3.0%)
Astronomy	520	2 (1.0%)
Physics	530	17 (8.6%)
Chemistry	540	18 (9.1%)
Earth Sciences (Geology)	550	9 (4.5%)
Life Science (Biology)	570	22 (11.1%)
Botany	580	3 (1.5%)
Zoology	590	1 (0.5%)
Technology (Applied Sciences)	600	3 (1.5%)
Medical Science	610	26 (13.1%)
Engineering Science	620	15 (7.6%)
Agricultural Science	630	1 (0.5%)
Home Science	640	2 (1.0%)
Business & Management	650	2 (1.0%)
Chemical Technology	660	8 (4.0%)
Literature	800	2 (1.0%)
English Literature	820	2 (1.0%)

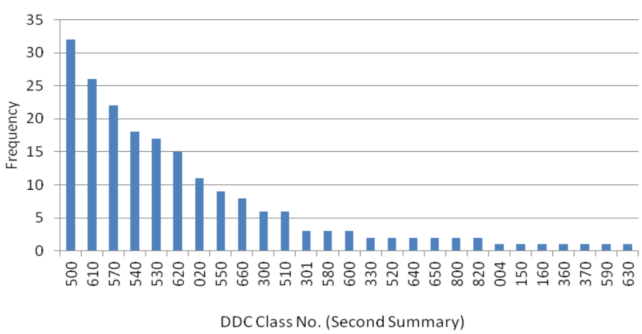

Fig. 3 — Variation of subject domains

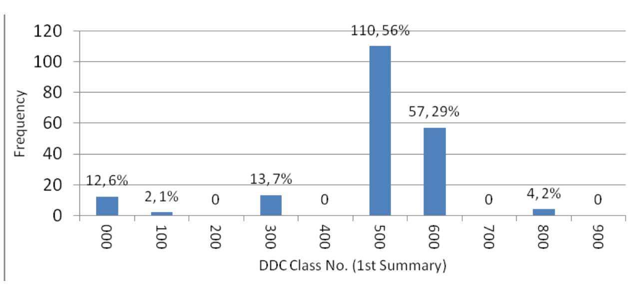

Fig. 4 — Variation of subject domains (Discipline-wise)

information science, earth science chemical technology etc. It is clear from Fig. 4, that the two broad disciplines, pure science and applied science together figure 85% of all growth dynamics studies.

Conclusion

From the Chi-square test, the null hypothesis is accepted, i.e., it is concluded that the distribution of different models of growth of literature over variant subjects is guided by Bradford's Law where the core or nucleus zone is occupied by either of logistic and exponential model. It may be pointed out that Price's empirical model prevails in Bradford's nuclear (core) zone in case of growth dynamics studies. This study shows an application of Bradford's law in Price's empirical theory. Also, the cumulative growth of literature on growth dynamics studies are found to follow fourth degree polynomial pattern as the best fit curve. It is found that largest number of such studies were performed in pure sciences (16.2%), followed by medical science (13.1%), life science (11.1%), chemistry (9.1%) and physics (8.6%) that figures nearly 60% of all studies.

The empirical theory of Price thus needs to be verified by other subject areas like management science, social science, creative and performing arts, language and literature etc. It is still necessary to verify applicability of Bradford's law of scattering in Price's theory for subject areas other than pure science and technology. This study emphasizes the necessity of growth dynamics study as an important tool for genesis and developmental analysis of a subject that may navigate properly in carrying out state-of-the-art-report or trend report of a subject.

References

1 Price DJS, *Little Science, Big Science*, (Columbia University Press, New York), 1963.

2 Evans JA, Future science, *Science*, 342 (6154) (2013) 44–45.

3 Henderson A, Diversity and the growth of serious / scholarly /scientific journals, *InScholarly publishing Books, Journals, Publishers and Libraries in the Twentieth Century*, edited by R E Abel, L W Newlin and K Strauch, (Wiley, New York), 2002, P. 133–162.

4 Henderson A, The growth of printed literature in the twentieth century, *InScholarly publishing Books, Journals, Publishers and Libraries in the Twentieth Century*, edited by R E Abel, L W Newlin and K Strauch, (Wiley, New York), 2005, P. 1–24.

5 Tabah AN, Literature dynamics: studies on growth, diffusion, and epidemics, *Annual Review of Information Science and Technology*, 34 (1999) 249–286.

6 de Bellis N, *Bibliometrics and Citation Analysis: From the Science Citation Index to Cybermetrics*, (Scarecrow Press, Maryland), 2009.

7 Bornmann L and Mutz R, Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references, *Journal of the Association for Information Science and Technology*, 66 (11) (2015) 2215-2222.

8 Lung RI, Gaskó N and Suciu MA, A hypergraph model for representing scientific output, *Scientometrics*, 117 (3) (2018) 1361-1379.

9 Kwiek M, The European research elite: a cross-national study of highly productive academics in 11 countries, *Higher Education*, 71 (3) (2016) 379-397.

10 Urban D, Mobility and the growth of science, *Social Studies of Science*, 12 (3) (1982) 409-433.

11 Yagi E, Badash L and Beaver DB, Derek J. de S. Price (1922-83): historian of science and herald of scientometrics, *Interdisciplinary Science Reviews*, 21 (1) (1996) 64-84.

12 Cozzens SE, Using the archive: Derek Price's theory of differences among the sciences, *Scientometrics*, 7 (3-6) (1985) 431-441.

13 Gilbert GN and Steve W, Essay Review: The quantitative study of science: an examination of the literature, *Science Studies*, 4 (3) (1974) 279-294.

14 Gilbert GN, Measuring the growth of science: a review of indicators of scientific growth, *Scientometrics*, 1 (1) (1978) 9-34.

15 Dedijer S, Measuring the growth of science, *Science*, 138 (3542) (1962) 781-788.

16 Brookes BC, The growth, utility and obsolescence of scientific periodical literature, *Journal of Documentation*, 26 (4) (1970) 283-294.

17 Tague J, Beheshti J and Rees-Potter L, The law of exponential growth: evidence, implications and forecasts, *Library Trends*, (Summer) (1981) 125-149.

18 Fernandez-Cano A, Torralbo M and Vallejo M, Reconsidering Price's model of scientific growth: an overview, *Scientometrics*, 61 (3) (2004) 301-321.

19 Szydlowski M and Krawiec A, Scientific cycle model with delay, *Scientometrics*, 52 (1) (2001) 83-95.

20 Heinzkill R, Characteristics of references in selected scholarly English literary journals, *The Library Quarterly*, 50 (3) (1980) 352-365.

21 Meadows J, Early reactions to information growth, *Scientometrics*, 51 (3) (2001) 553-561.

22 Leah AL, Social media and the production of knowledge: a return to little science? *Social Epistemology*, 24 (3) (2010) 219-237.

23 Furner J, Little book, big book: Before and after little science, big science: a review article, Part I, *Journal of Librarianship and Information Science*, 35 (2) (2003a) 115-125.

24 Furner J, Little book, big book: Before and after little science, big science: a review article, part II, *Journal of Librarianship and Information Science*, 35 (3) (2003b) 189-201.

25 Andersen JP and Hammarfelt B, Price revisited: on the growth of dissertations in eight research fields, *Scientometrics*, 88 (2) (2011) 371-383.

26 Glänzel W and Schoepflin U, Little scientometrics, big scientometrics... and beyond?, *Scientometrics*, 30 (2-3) (1994) 375-384.

27 Price DJS, Quantitative measures of the development of science, *Archives Internationalesd'Histoire des Sciences*, 14 (1951) 85–93.

28 Price DJS, Science since Babylon, (Yale University Press, New Haven), 1961.

29 Price DJS, Networks of scientific papers, *Science*, 149 (3683) (1965) 510–515.

30 Neelameghan A, Documentation of the history of medicine in India, *Annals of Library Science and Documentation*, 10 (3-4) (1963) 116-142.

31 Vickery BC, Bradford's law of scattering, *Journal of Documentation*, 4 (3) (1948) 198-203.

32 Walton C and Morris A, A bibliometric study of taxonomic botany, *Journal of Documentation*, 69 (3) (2013) 435-451.