Wolfgang Glänzel and András Schubert

From Matthew to Hirsch: A Success-Breeds-Success Story

1 Preamble: A Submicro-Level Bisociation Study

It was 50 years ago that James S. Coleman's *Introduction to Mathematical Sociology* was published (Coleman, 1964). The book has received more than 500 citations since then in Thomson Reuters' *Web of Science Core Collection* (and more than 1500 citations according to Google Scholar). One of these citations is found in a 1982 review by a young information scientist from London: Blaise Cronin (Cronin, 1982). Another one, two years later, was given by two authors from Budapest: Wolfgang Glänzel and András Schubert (Schubert & Glänzel, 1984). Although both papers have been cited above average (as they belong to the Hirsch-core of their respective authors), they have never been co-cited. As we might learn from Koestler (1964), some of the most remarkable instances of human creativity can be attributed to what he called *bisociation*: "the perceiving of a situation or idea in two self-consistent but habitually incompatible frames of reference". In his views, "creation" is making connection between two well-known, but yet unconnected entities. In bibliometric terms, bisociation is encountered whenever two frequently used but so far unconnected keywords co-occur, or when two frequently cited items are co-cited. In such cases, there are increased chances for outstanding achievements. For keyword bisociation, the conjecture was empirically supported by an inorganic chemistry example (Schubert & Schubert, 1997; Schubert, 2013). No systematic studies on co-citation bisociation presently exist. The present article may be considered a self-experiment in order to determine whether an unprecedented co-citation of Cronin (1982) and Schubert and Glänzel (1984) will lead to success. The result is the sole responsibility of the authors.

2 Introduction

In the present paper, an overview is given of statistical models of bibliometric distributions on the basis of the principle called "Matthew's principle", "cumulative advantage", or "success-breeds-success", among others. A general introduction is followed by the description of the model, the properties of a particular distribution derived from the model—the Waring distribution—and a family of apparently

DOI 10.1515/9783110308464-011,

remote distributions all derivable from the model. In retrospect, the model can be considered a precursor of Barabási's celebrated "preferential attachment" network model resulting in power law-type distributions, and some relations with Hirsch's *h*-index are also revealed.

Skewness is one of the most conspicuous features of informetric distributions. Whether we refer to Pareto's 80/20 rule or any other specimen of unequal proportions, strong inequalities are readily illustrated on the distribution of publication productivity or citation impact. The skewness property is, however, only one of the common features of informetric distributions.

The mathematical models of these distributions are sometimes taken from the Gaussian distribution family, such as Poisson, negative binomial, etc., but more often—as in the case of the classical laws of bibliometrics: those of Lotka, Bradford, and Zipf—approximately follow an inverse power law (see Furner, this volume, for more on these and other power-law distributions). This is the second important feature of scientometric distributions: their heavy tail. The question about the usefulness of the two families of distributions in scientometrics has been discussed in detail, among others, by Brookes (1968) and Haitun (1982).

There is an apparent consensus that the generating mechanism of these distributions is some kind of positive feedback, what was called the Matthew principle by Merton (1968), "cumulative advantage" by Price and Gürsey (1976), reinforcement by Allison (1980), and "success-breeds-success" by Tague (1981).

From the mathematical viewpoint there are several ways of developing and describing such models. Proceeding from a simple point process, i.e., the random sum of independent exponentially distributed random variables, where all special properties of inequality, such as subject-specific peculiarities, the authors' social status, or academic age, etc., can be obtained via (1) *compound distributions* and stochastic processes, that is, by mixtures of random effects; (2) the *urn model* according to Pólya and Eggenberger (Eggenberger & Pólya, 1923), where success or failure might positively or negatively affect further trials; or (3) stochastic *birth-and-death processes* with particular transition rules controlling for the extent of cumulative advantage. All these models result in the same family of distributions and processes but highlight different aspects of the genesis of inequality properties like skewness and heavy tails.

A versatile model was proposed in the 1980's by Schubert and Glänzel (1984) based on a simple counting process using a deterministic birth model with immigration and emigration and a transition rate that is a linear function of the actual count. Depending on the supplementary conditions, the process may result in Poisson, negative binomial, geometric, or Waring distributions. Each of these distributions may be used to model bibliometric samples in certain conditions.

The model is based on the scheme of a simple Poisson process presented by Coleman (1964, p. 289). In his model, all transition rates are equal and independent of the previous counts. The more general model of a birth process can be obtained if the transition rates may change according to the previous counts and following particular rules (cf. Figure 1) (discussed in more detail later).

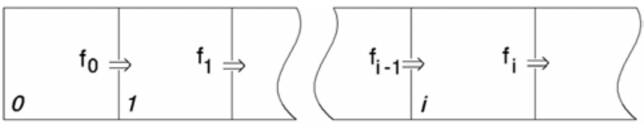

Fig. 1: Coleman's scheme of the Poisson process (with $f_i = a =$ const. for all $i \geq 0$).

In Coleman's brand loyalty example, "the states labelled 1, 2, 3, … are the states of having bought brand *A* one, two, three, etc., times in succession."

Turning to bibliometrics, we can think about authors publishing 1, 2, 3, … papers on a topic before getting into state 0 by changing topic (or ceasing to publish at all). In this simplest model there is no kind of "advantage", the chances of proceeding from the i-th state to the $(i + 1)$-th are equal independently of the present state (a).

3 The Schubert–Glänzel model

A somewhat modified and generalized version of Coleman's scheme was used in Schubert and Glänzel (1984). In order to build a (stochastic) birth process with the desired features to model informetric processes we first consider an infinite array of units or cells, indexed in succession by the non-negative integers, among which a certain substance is distributed. The content of the i-th cell is denoted by x_i; the (finite) content of all units or cells by x.

Obviously, $x = \sum_i x_i$. Then the fraction $y_i = x_i/x$ $(i \geq 0)$ expresses the share of elements contained by the i-th cell. The change of content is postulated to obey the following rules.

Substance may enter the system from the external environment through the 0-th unit at a rate s; (1)

substance may be transferred unidirectionally from the i-th unit to the $(i + 1)$-th one at a rate f_i $(i \in \mathbf{N}_0)$; and (2)

substance may leak out from the i-th unit
into the external environment at a rate g_i $(i \in \mathbf{N}_0)$. (3)

The next step towards a stochastic model is to interpret the above ratios y_i as the (classical) probability with which an element is contained by the i-th unit. The stochastic process is then formed by the change of the content of the cells (or units), i.e., by the change of purchases or papers published by the authors, who have entered the system. The discrete random variable $X(t)$ denotes the (random) number of purchases or papers published at time t, $P(X(t) = i) = y_i$ its probability. In order to use an example from bibliometrics, $P(X(t) = i)$ might, for instance, be the probability that an author has published exactly i papers in the period from the time of his/her entrance into the system, denoted by 0, until time t. Figure 2 visualizes the scheme of substance flow of this process.

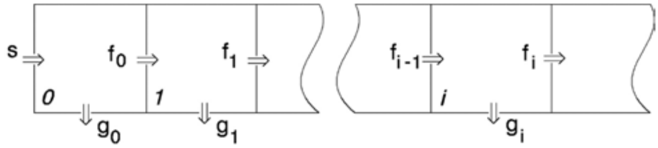

Fig. 2: Scheme of substance flow with immigration and emigration according to Schubert and Glänzel (1984).

According to Schubert and Glänzel, the following particular forms of the above rate terms are used:

$$s = \sigma \cdot x\,, \qquad (1^*)$$

$$f_i = (a + b \cdot i) \cdot x_i\,; \quad (i \geq 0)\,, \qquad (2^*)$$

$$g_i = \gamma \cdot x_i\,; \quad (i \geq 0)\,, \qquad (3^*)$$

where σ, a, b and γ are non-negative real values. The distribution of the substance over the units during time t can then be obtained as a solution of a relatively simple system of first order linear differential equations (cf. Schubert & Glänzel, 1984).

The relationship of this model with Price's author categorization bears mentioning. Price and Gürsey (1976) distinguished the following four categories of authors: *newcomers, continuants, transients,* and *terminators*. Within the framework of this model, s represents the group of newcomers, g_i terminators, g_1 transients and substance remaining in the system, i.e., $\sum_i f_i$ represents the group of continuants.

In addition, we would further like to mention that there is a certain relationship with the epidemic model according to Goffman and Nevill (Goffman & Nevill, 1964; Goffman, 1965), although the birth model with immigration and emigration does not explicitly assume the occurrence of any epidemic situation, and is, on the other hand, somewhat more complex as it differentiates the status of "infection" further by counting events. Goffman and Nevill introduced the theory of *intellectual epidemics as a model of scientific communication* in 1964. According to their model, which in turn is based on the classical Reed-Frost model, the diffusion of ideas in a population of scientists could be compared to the spreading of an influenza virus in a population of people, causing an epidemic. The population can at any time be subdivided into three groups of infected (I), resistant or immune (R), and infection sensitive, i.e., susceptible (S) animals or persons. Goffman and Nevill considered a published article on a specific topic an infection. Using this model as an analogon, we have S appearing in the Schubert–Glänzel scheme as the group (s) entering the system I corresponds to and group R to those who left the system ($\sum_i g_i$). But most notably, in this context the two main cases of the epidemic model are also obtained, particularly, if $\sigma = \gamma$ (with the further two subcases $\sigma = 0$ and $\sigma \neq 0$), or $\sigma \neq \gamma$, respectively. In the following section we will analyze these cases more in detail.

3.1 Two Special Cases: "Closed Systems" and "Equilibrium Systems"

For the entire population, we can derive $x(t) = x(0) \cdot \exp\left((\sigma - \gamma) \cdot t\right)$, i.e., the system is asymptotically time-invariant (stationary) if $\sigma = \gamma$, otherwise, if $\sigma > \gamma$ or $\sigma < \gamma$, it exponentially grows or decays, respectively. The distribution of the "substance" (purchases, publication productivity, etc.) can be exactly determined in two special cases using the notation of the Schubert–Glänzel scheme.

(i) In a "closed system" we assume $\sigma = \gamma = 0$, after (finite) time t.

In this case, the distribution of the substance will take a negative binomial distribution (Pólya distribution, in the terminology of Coleman [1964, p. 301]).

$$y_i = \binom{k+i-1}{i}\left(\frac{k}{\mu+k}\right)^k\left(\frac{\mu}{\mu+k}\right)^i, \tag{4}$$

with a scale parameter $\mu = (a/b)(e^{bt} - 1)$, and a shape parameter $k = a/b$, where a and b are the two parameters of the transition rate, f_i. The first factor, $\binom{k+i-1}{i}$, is a binomial coefficient.

This model was successfully used by Allison (1980) in describing the publication productivity of a cohort of chemists in the first six years after the doctorate.

In the special case $b = 0$ (no cumulative advantage, as in Figure 1), we have $k \to \infty$, $\mu \to at$, and the distribution becomes Poisson.

$$y_i = \mu^i \mathrm{e}^{-\mu}/i!\,. \tag{5}$$

(ii) In an "equilibrium system" we assume $\sigma = \gamma > 0$, and $t \to \infty$.

In this case, $b = 0$ (no cumulative advantage) leads to a geometric distribution.

$$y_i = q(1-q)^i\,, \tag{6}$$

with $q = \sigma/(\sigma + a)$.

This model describes the asymptotic steady state ("equilibrium") of the productivity distribution of an author community with a constant supply of newcomers and a constant "dropout" of authors (for whatever reason: retirement, death, topic change, leaving academia, etc.) independently of the productivity level reached so far, provided that there is no cumulative advantage effect.

In the cumulative advantage case, $b > 0$, the equilibrium distribution has a less well-known form, namely that of the Waring distribution (Schubert & Glänzel, 1984). In particular, we obtain the following limiting distribution.

$$\begin{aligned} y_i(\infty) &= \frac{\sigma(a+b)\ldots(a+b(i-1))}{(a+\sigma)(a+b+\sigma)\ldots(a+bi+\sigma)} \\ &= \frac{\alpha(N+1)\ldots(N+i-1)}{(N+\alpha)(N+\alpha+1)\ldots(N+\alpha+i)}\,, \end{aligned} \tag{7}$$

with parameters $\alpha = \sigma/b$, $N = a/b$.

3.2 Properties of the Waring Distribution

Although the closed-form definition of the Waring distribution looks a bit awkward, it obeys a rather simple recursive formula.

$$\begin{aligned} y_0 &= \alpha/(N+\alpha)\,, \\ &\vdots \\ y_i &= y_{i-1}(N+i-1)/(N+\alpha+i)\,. \end{aligned} \tag{8}$$

Also, the Waring distribution has some remarkable properties.

– It obeys Zipf's law:

$$y_i \approx i^{-(\alpha+1)}\,, \quad \text{as } i \text{ tends to infinity,} \tag{9}$$

i.e., the tail of the distribution follows an inverse power law. Due to the relation (8), the tail exponent can be estimated even from any two frequency values.

– The mean value of the Waring distribution has the very simple form:

$$\langle y \rangle = N/(\alpha - 1)\,. \tag{10}$$

– The Waring distribution has a “self-similarity” property.

$$y_{-j} \equiv y_{i-j}(\alpha, N) = y_i(\alpha, N + j) \quad \text{for any } i \geq j\,, \tag{11}$$

i.e., a Waring distribution truncated from left at j and shifted back with j units, is again a Waring distribution with unchanged parameter α (as can be expected, since the asymptotic Zipf behavior must not change) and with a parameter N increased by j units. The geometric distribution, which is a limiting case of the Waring distribution, if N, $\alpha \to \infty$ and $N/\alpha = (1/q - 1) > 0$, has the “lack-of-memory property,” that is, a geometric distribution truncated from left at j and shifted back with j units, is the identical geometric distribution with unchanged parameter.

– From equations (10) and (11) it follows, that

$$\langle y_{-j} \rangle = (N + j)/(\alpha - 1)\,, \tag{12}$$

i.e., the mean value of the left-truncated and left-shifted Waring distribution is a linear function of the point of truncation. This property is a characterization: the linear relation holds if and only if the distribution is Waring. This characterization is a special case of a more general characterization theorem (Glänzel et al., 1984).

3.3 Applications of a Characterization Theorem

The linear relation (12) can be used as the basis of an effective statistical test and extrapolation tool. Plotting the series of truncated mean values $\langle y_{-0} \rangle$, $\langle y_{-1} \rangle$, $\langle y_{-2} \rangle, \ldots$ against the point of truncation, 0, 1, 2, ... a straight line should be obtained indicating the subsistence of a Waring distribution. A statistical test has been elaborated and presented on a linguistic example by Telcs et al. (1985).

If one happens to have a left-truncated set of frequency data, the same straight line may help to extrapolate to the missing region. Most typically, publication frequency data start with 1, i.e., do not account for the “silent majority”: those researchers who happen not to publish within the framework studied. The number of these researchers can be estimated using the Waring model. A successful attempt has been reported by Schubert and Telcs (1987) on the example of estimating the size of "publication-worthy" researcher community ("publication potential") of U.S. states. Furthermore, the method has been adapted to evaluate research institutions in Sweden, and induced heated science policy debates that have not yet been settled (Koski, 2013).

Another interesting consequence of the linear relation (12) is connected with the so-called "Characteristic Scores and Scales" (CSS) method (Glänzel & Schubert, 1988). This is a method for marking thresholds to divide a sample into classes according to the value of a random variable ξ. While, e.g., in case of quantiles, the classes are defined to contain equal number of elements, CSS classes adjust themselves according to the nature of the distribution. CSS thresholds originated from iteratively truncating samples at their mean value and recalculating the mean of the truncated sample until the procedure is stopped or no new scores are obtained. This procedure is briefly described in the following.

After putting $b_0 = 0$, the sample mean is chosen as the first threshold, denoted now as b_1.

$$b_1 = E(\xi)\,. \tag{13}$$

Further thresholds are defined recursively.

$$b_k = E(\xi \mid \xi \geq b_{k-1})\,. \tag{14}$$

That is, the second threshold, b_2, is equal to the mean value of all sample elements equal to or greater than the overall sample mean, and so on. The classes are then defined by the pairs of the corresponding threshold values, particularly on the basis of the half-closed intervals $[b_{k-1}, b_k)$ with $k \geq 0$. Depending on the sample size and the nature of the distribution, three to five classes are usually sufficient for a practical classification task.

Taking into account that in the above-average region Zipf's law (Eq. (9)) begins to come into force, the linear relation (12) leads to the following approximation (Glänzel, 2013):

$$b_k \approx b_1(a^k - 1)/(a - 1)\,, \tag{15}$$

where $a = \alpha/(\alpha - 1)$, α being the tail exponent, which is identical with the corresponding parameter of the Waring distribution (cf. Eq. (9)). Approximate thresholds can, therefore, be calculated from estimations of the mean value and the tail exponent even if the full distribution of the variable is not available.

3.4 Another Special Case

The $a = 0$ case (here a denotes the coefficient in the transfer rate of the above model according to Schubert and Glänzel) seems to lead to a trivial solution.

Eq. (2*) reduces to a formula often associated with "Gibrat's law"[1] (Gibrat, 1931):

$$f_i = bix_i\,, \tag{16}$$

consequently, $f_0 = 0$, thus no substance reaches beyond cell 0: $y_0 = 1$, $y_i = 0$ for all $i > 0$. If, however, the limiting distribution is sought for $i > 0$, $a \to 0$, a non-trivial solution is found both in the "closed-system" and the "equilibrium-system" case.

In a closed system, we have

$$y_i = q(1-q)^{i-1}\,. \tag{17}$$

This is a geometric distribution analogous to equation (6), but in this case $q = \mathrm{e}^{bt}$. In this case the geometric distribution emerges as a special case of the zero-truncated negative binomial distribution. This case is treated in (Coleman, 1964, p. 307).

In an equilibrium system the same procedure leads to a Yule distribution (see, e.g., Price, 1976).

$$y_i = \alpha\,\mathrm{B}(\alpha+1, i)\,, \tag{18}$$

where $\mathrm{B}(\cdot,\cdot)$ denotes the beta function. The Yule distribution is a prototype of distributions obeying Zipf's law (Eq. (9)); remarkably, Simon (1955) derived it from Gibrat's law. In this case, extrapolation to zero is meaningless; there is an infinite pool of zero-elements behind the apparent distribution.

3.5 A Summary of Distributions Emerging from the Schubert–Glänzel Model

Table 1 summarizes the distributions emerging from the model outlined in Figure 2 and equations (1)–(3) under various conditions.

The most striking feature is that while the 'closed-system' solutions are Gaussian in nature (i.e., have an exponential tail), the 'equilibrium' solutions are Zipfian (have a power-law tail). The special appeal of the model is that both of these classes, sometimes considered antagonistic, can be derived from it. The geometric distribution, which is a degenerate case of both classes has its generating scheme in both columns.

1 en.wikipedia.org/wiki/Gibrat's_law

Tab. 1: Distributions emerging from the Schubert–Glänzel model.

	Closed system	**Equilibrium system**
General case	Negative binomial (Pólya); Eq. (4)	Waring; Eq. (7)
$b = 0$ (no cumulative advantage)	Poisson; Eq. (5)	Geometric; Eq. (6)
$a \to 0$, $i > 0$	Geometric; Eq. (14)	Yule; Eq. (15)
$i \to \infty$	Exponential	Inverse power (Zipf); Eq. (9)

4 An Alternative Explanation: Heterogeneity

It is important to note that the same distribution patterns can be explained without the cumulative advantage hypothesis. We may assume that any given researcher has a constant probability to publish the next paper independently of the number of papers already published (i.e., $b = 0$), yet a set of researchers may have a productivity distribution given in Table 1. This is possible, if the population is heterogeneous, namely, if a (constant for each single researcher) has a gamma distribution. A Poisson distribution with a gamma distributed parameter, μ, provides a negative binomial distribution, while a geometric distribution with gamma distributed parameter, q, results in a Waring distribution.

As Coleman (1964) notes:

> It is impossible to choose between a contagious[2] interpretation and a heterogeneity interpretation merely on the basis of the empirical distribution itself, no matter how well it fits a theoretical distribution. What are required in addition are over-time data, which can show the development of contagion if it exists.
>
> (p. 301)

In two rare cases of such longitudinal studies, Huber and Wagner-Döbler (2001a, b) found that in two fields: mathematical logic and physics, heterogeneity rather than cumulative advantage is responsible for the observed empirical productivity distributions.

2 In Coleman's terminology, "contagion" is analogous with cumulative advantage.

5 Network Aspects

Given the fact that scientific publications are increasingly the result of collaborative authorship (see, e.g., Glänzel & Schubert, 2004)—or hyperauthorship (Cronin, 2001), it seems obvious to seek cumulative advantage patterns in co-authorship distributions as well. At the time of elaborating the cumulative advantage publication productivity models, attempts to extend the idea to co-author networks were largely hindered by the lack of a proper mathematical theory. The Erdős-Rényi model of random graphs led to a Poisson distribution, and the cumulative advantage element could not be included into the model.

An alternative approach was found by Barabási and his group at the turn of the millennium. The scale-free network concept (Barabási & Albert, 1999) was inspired by well-tested physical models, and was successfully applied for such popular examples as the Internet and the network of movie actors. They coined the term "preferential attachment" for the cumulative advantage phenomenon. The applicability of the model to co-author networks has been demonstrated (Newman, 2001; Barabási et al., 2002) and became milestone papers of the topic.

6 The Hirsch-Connection

Shortly after Hirsch (2005) let the h-index genie out of the bottle, Glänzel (2006) attempted to find the position of the index in the framework of statistical theory. He found that it is closely connected with Gumbel's theory of characteristic extreme values. For distributions having inverse power tail (obeying Zipf's law), he could relate the h-index to traditional statistical parameters of the sample with the following formula:

$$h \approx x^{\alpha/(\alpha+1)} n^{1/(\alpha+1)} , \tag{19}$$

where h is the h-index, x is the sample mean, n is the sample size, and α is the tail exponent.

The formula found widespread empirical support, among others, on the example of h-indices of journals (Schubert & Glänzel, 2007) and countries (Csajbók et al., 2007).

7 Network-based *h*-indices

Far beyond the scope of bibliometrics, in general, and research evaluation, in particular, the h-index raised interest in network/graph research. Eppstein and Spiro (2009) interpreted it as a graph invariant used in constructing dynamic graph algorithms particularly efficient for scale-free networks. Korn et al. (2009) used it under the name of lobby-index as a centrality measure characterizing the network's capability for efficient communication. Its use as a measure of coherence in a community of researchers (authors) was demonstrated by Schubert et al. (2009).

An interesting Hirsch-type index characterizing social networks was defined by Schubert (2012a) and amended by Rousseau (2012). Actually, the index is a special case of an idea of Zhao et al. (2011). The partnership ability index, φ, can be defined for arbitrary actions and actors as follows: An actor is said to have a partnership ability index φ, if with φ of his/her n partners had at least φ joint actions each, and with the other $(n - \varphi)$ partners if they had no more than φ joint actions each. In the original paper, the index was exemplified on a co-authorship sample, but it proved to be applicable also in a network of jazz musicians (Schubert, 2012b).

Although the indicator is meaningful for networks of any structure, for scale-free networks Glänzel's relation (19) is supposed to hold. Indeed, both mentioned studies supported the validity of this approximation, similarly to a study on a large co-author sample reported by Cabanac (2013).

The reader should be reminded that, similarly to the empirical distribution in the case of publication productivity models, the network properties (e.g., the inverse power tail of the degree distribution) is not a sufficient basis for inferring to the generating mechanism. Cumulative advantage (preferential attachment) is a reasonable possibility, but other mechanisms (e.g., heterogeneity) cannot be excluded.

8 Conclusion

Success is an extremely multifaceted concept. In scientific and scholarly research, success is the creation and communication of new knowledge. In informetrics, publication output and citation impact are considered the main measurable aspects of success. In both aspects, success appears to be 'contagious': success breeds success, advantages cumulate. Theories of informetrics should give an account of this feedback mechanism—that is what our model attempted, and maybe not quite unsuccessfully achieved.

A particular recipe for success is attributed to Mark Tyson: "Confidence breeds success and success breeds confidence …Confidence applied properly surpasses genius."[3] As a net result, in this 'model', as well, success seems to breed success and, although the mechanisms can be disputed, Matthew's truth still prevails. And as it could be seen from the overview compiled in this paper, success-breeds-success is not only a successful way to reach success in publication and citation terms, but also a fruitful and successful topic of bibliometric modeling.

Cited References

Allison, P. D. (1980). Estimation and testing for a Markov model of reinforcement. *Sociological Methods & Research*, *8*, 434–453.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science* *286*(5439), 509–512.

Barabási, A. L., Jeong, H. Néda, Z., Ravasz, E., Schubert, A. & Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A*, *311*(3–4), 590–614.

Brookes, B. C. (1968). The derivation and application of the Bradford-Zipf distribution. *Journal of Documentation*, *24*(4), 247–265.

Cabanac, G. (2013). Experimenting with the partnership ability φ-index on a million computer scientists. *Scientometrics*, *96*(1), 1–9.

Coleman, J. S. (1964). Introduction to Mathematical Sociology. New York: Free Press.

Cronin, B. (1982). Invisible-colleges and information-transfer – A review and commentary with particular reference to the social sciences. *Journal of Documentation*, *38*(3), 212–236.

Cronin, B. (2001). Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices? *Journal of the American Society for Information Science and Technology*, *52*(7), 558–569.

Csajbók, E., Berhidi, A., Vasas, L., & Schubert, A. (2007). Hirsch-index for countries based on essential science indicators data. *Scientometrics*, *73*(1), 91–117.

Eggenberger, F., & Pólya, G. (1923). Über die Statistik verketer Vorgänge, *Zeitschrift für Angewandte Mathematik und Mechanik*, *1*, 279–289.

Eppstein, D., & Spiro, E. S. (2009). The h-Index of a Graph and Its Application to Dynamic Subgraph Statistics. Algorithms and Data Structures. *Lecture Notes in Computer Science*, *5664*, 278–289

Gibrat, R. (1931). *Les Inégalités Économiques*. Paris: Librairie du Recueil Sirey.

Glänzel, W. (2006). On the h-index - A mathematical approach to a new measure of publication activity and citation impact. *Scientometrics*, *67*(2), 315–321.

Glänzel, W. (2013). High-end performance or outlier? Evaluating the tail of scientometric distributions. *Scientometrics*, *97*(1), 13–23.

Glänzel, W., & Schubert, A. (1988). Characteristic scores and scales in assessing citation impact. *Journal of Information Science*, *14*(2), 123–127.

3 http://mmaquotes.blogspot.hu/2014/04/confidence-breeds-success-and-success.html

Glänzel, W., & Schubert, A. (2004). Analyzing scientific networks through co-authorship, In: H. F. Moed et al. (eds.), *Handbook of Quantitative Science and Technology Research*, pp. 257–276.
Glänzel, W., Telcs, A., & Schubert, A. (1984). Characterization by truncated moments and its application to Pearson-type distributions. *Probability Theory and Related Fields* (at the time of publication: Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete), *66*(2), 173–183. (Correction: *Probability Theory and Related Fields*. 74 (1987) 317.)
Goffman, W. (1965). An epidemic model in an open population, *Nature*, *205*, 831.
Goffman W., & Nevill, V. A. (1964). Generalization of epidemic theory. *Nature*, *204*, 225.
Haitun, S. D. (1982). Stationary scientometric distributions. I-II-III. *Scientometrics*, *4*(1), 5–25; *4*(2), 89–104; *4*(3), 181–194.
Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*. *102*(46), 16569–16572.
Huber, J. C., & Wagner-Döbler, R. (2001a). Scientific production: A statistical analysis of authors in mathematical logic. *Scientometrics*, *50*(2), 323–337.
Huber, J. C., & Wagner-Döbler, R. (2001b). Scientific production: A statistical analysis of authors in physics, 1800–1900. *Scientometrics*, *50*(3), 437–453.
Koestler, A. (1964). *The Act of Creation*, Dell: New York.
Korn, A., Schubert, A., & Telcs, A. (2009). Lobby index in networks. *Physica A-Statistical Mechanics and its Applications*, *388*(11), 2221–2226.
Koski, T. (2013). *On the Waring distribution, the Glänzel-Schubert model and their applications — a historiett*. Retrieved from http://www.math.kth.se/~tjtkoski/goranwaring.pdf
Merton, R. K. (1968). The Matthew Effect in Science. *Science*, *159*(3810), 56–63.
Newman, M. E. J. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America*. *98*(2), 404–409.
Price, D. De Solla (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, *27*(5), 292–306.
Price, D. D., & Gürsey, S. (1976). Studies in scientometrics. Part 1. Transience and continuance in scientific authorship. *Internation Forum on Information and Documentation*, *1*, 17–24.
Rousseau, R. (2012). Comments on "A Hirsch-type index of co-author partnership ability". *Scientometrics*, *91*(1), 309–310.
Schubert, A. (2012a). A Hirsch-type index of co-author partnership ability. *Scientometrics*, *91*(1), 303–308.
Schubert, A. (2012b). Jazz discometrics: a network approach. *Journal of Informetrics*, *6*(4), 480–484.
Schubert, A. (2013). A follow-up study of title word bisociations in Inorganica Chimica Acta. *ISSI Newsletter*, *9*(3), 54–55.
Schubert, A., & Glänzel, W. (1984). A dynamic look at a class of skew distributions – A model with scientometric applications. *Scientometrics*, *6*(3), 149–167.
Schubert, A., & Glänzel, W. (2007). A systematic analysis of Hirsch-type indices for journals. *Journal of Informetrics*, *1*(3), 179–184.
Schubert, A., Korn, A., & Telcs, A. (2009). Hirsch-type indices for characterizing networks. *Scientometrics*, *78*(2), 375–382.
Schubert, A. P., & Schubert, G. A. (1997). Inorganica Chimica Acta: its publications, references and citations. An update for 1995–1996, *Inorganica Chimica Acta*, *266*(2), 125–133.

Schubert, A., & Telcs, A. (1989). Estimation of the publication potential in 50 U.S. states and in the District of Columbia based on the frequency distribution of scientific productivity. *Journal of the American Society for Information Science*, *40*(40), 291–297.

Simon, H. A. (1955). On a class of skew distribution functions, *Biometrika*, *42*(3–4), 425–440.

Tague, J. (1981). The success-breeds-success phenomenon and bibliometric processes. *Journal of the American Society for Information Science*, *32*(4), 280–286.

Telcs, A., Glänzel, W., Schubert, A. (1985). Characterization and statistical test using truncated expectations for a class of skew distributions. *Mathematical Social Sciences*, *10*(2), 169–178.

Zhao, S. X., Rousseau, R. & Ye, Y. (2011). h-degree as a basic measure in weighted network. *Journal of Informetrics*, *5*(4), 668–677.