Growth dynamics of German university enrolments and of scientific disciplines in the 19th century: scaling behaviour under weak competitive pressure

Frank Havemann, Michael Heinz, and Roland Wagner-Döbler
Institute of Library Science, Humboldt-University Berlin, Dorotheenstr. 26, 10117 Berlin, Germany.
E-mail: FNHavemann@compuserve.de, michael.heinz@rz.hu-berlin.de, rfw-d@t-online.de

Abstract. According to authors like H.E. Stanley and others, growth dynamics of university research displays a quantitative behaviour similar to the growth dynamics of firms acting under competitive pressure. Features of such a behaviour are probability distributions of annual growth rates or the standard deviation of growth rates. We show that a similar statistical behaviour can be observed in the growth dynamics of German university enrolments or in the growth dynamics of physics and mathematics, both for the 19th century. Since competitive pressure was generally weak at that time, interpretations of statistical similarities as to pointing to a “firm-like behaviour” are questionable.

Introduction

Growth dynamics of economies is of obvious practical interest, but also in the centre of theoretical interest of many economists, beginning already in the 19th century. In the recent two decades or so, this interest concentrated on finding clues of chaotic behaviour, as it might be visible in fluctuations of economic time series. Growth dynamics of another kind of productive activity, namely science, can meanwhile be called a classical topic in scientometrics, of course connected with D. deSolla Price‘s investigations on exponential growth of that activity‘s output in the last two centuries. For scientometrics and informetrics, the recent work on growth dynamics is reviewed by Tabah (1999). With few exceptions (e.g. Goffman & Harmon, 1971; Tabah, 1999; Wagner-Döbler, 1998), scientometricians focused not so much on growth fluctuations, as economists are used to do, as on general trends of science and technology growth and on modelling these trends.

Applying statistical physics methods, in a series of articles some authors around H.E. Stanley (Boston), mostly coming from physics, recently studied growth dynamics and the annual fluctuations of growth rates of firms’ sales, of national economies, and eventually of research activities of universities (M.H.R. Stanley et al., 1996; Lee et al., 1998; Plerou et al., 1999). This work found also interest in scientometrics (Amaral et al., 2001). Among the scientometric indicators used were research and development expenditures of United States universities, patents issued to U.S. universities, and papers brought out by them, using databases, which covered the last or two decades, as a rule. The authors computed the standard deviations of the distributions of annual growth rates as a function of the size of the growing units and found that they fit to a power-law with an exponent in the range of 0.25. Dividing the universities or firms into three groups of small, medium, and large expenditures or sales,

Frank Havemann, Michael Heinz and Roland Wagner-Döbler, *Growth dynamics of German university enrollments and of scientific disciplines in the 19th century: scaling behaviour under weak competitive pressure*, In: Guohua Jiang, Ronald Rousseau, Yishan Wu (Eds), Proceedings of the 9th International Conference on Scientometrics and Informetrics – ISSI 2003, Dalian University of Technology, Dalian, China, 2003, 91-98.

respectively, they furthermore found that the logarithmised distributions of growth rates, if scaled with the growth rates‘ standard deviation of the corresponding group, „collapsed“ onto a single graph of a „tent“ form. These „tents“ are known as doubled-sided exponential or Laplace distributions. According to these authors, the distributions of university research growth rates obviously display a universal form and a scaling behaviour typical for competitive economic activities as they have been analysed e.g. by R. Gibrat, H. Simon, and by themselves. Similar results appeared for Canadian and English universities. Stanley et al. interpreted this as to possibly reflect increasing competition between funds-seeking universities. From such a view, universities exhibit a firm-like behaviour. Immediately after these new insights were published, Moed and Luwel (1999) pointed out that strong competition in business implies a shortsighted perspective of survival, whereas basic research is connected with a long-term vision of growth, according to Moed and Luwel. A vision more associated with the perspective of a forester who is thinking in terms of generations than with the perspective of a business, as one could say. Thus, Moed and Luwel brought forward the question whether strong competition in a „business of research“, as mirrored by scaling behaviour typical for firms, might undermine basic research in the long run. This is a question of utmost importance for science policy, of course.

Method and Data

In the present paper, in contrast to the work we sketched above, we aim to examine a growth dynamics of scientific institutions under quite different economic and socio-political circumstances, in a quite different historical period of time. The first half of the 19th century was a comparatively static period of time with respect to institutions of higher education and scientific activity in Germany, apart from the general fact that Germany, as all countries of that century, were the opposite of information or knowledge societies in our modern sense (if it has a sense). With regard to competition, we can at least hypothesize that the competition among universities for students was low—and may be low even nowadays in Germany, because universities are financed by state and not by student fees—, and that the same be valid for competition between subfields of science. (That by quantitative terms of migration between subfields, even in the 20th century competition between subfields is low, at least in mathematics, was shown by Wagner-Döbler, 1998a.) With this in mind, we analysed the growth dynamics of student enrolments of all German universities for 1830-1900, of the 34 most important mathematical fields, and of the 53 most important physics fields, both for 1800-1900. Data for student enrolments were compiled from a well-known university history (Prahl, 1978). They refer to all 21 universities located in the area of the „Deutsches Reich“ in its founding year 1871. In order to avoid seasonal effects of winter and summer semesters (with more students in winter than in summer, as a rule), two growth rates for each year were computed, referring to the preceding winter or summer semester, respectively.

Data for mathematics and physics stem from the printed subject indexes of the „Catalogue of Scientific Papers 1800-1900“ of the Royal Society of London that were transformed into a machine-readable bibliometric source by Wagner-Döbler and Berg (1996; 1999). 19th-century mathematics and physics together comprise more than 100,000 papers, and these papers were classified for the Index by a group of experts in retrospect in the first years of the 20th century. Some mathematical or physics subfields not directly related to mathematical or physics research were excluded (e.g., history).

As indicators of the growth dynamics of mathematics and physics we use annual counts of publications of a field, computed as moving 3-year-averages. For comparative purposes and to test the robustness of our computations, in addition numbers of active contributors to the fields were used, which can conceived of as an indicator of scientific manpower. Following suggestions of Goffman and Harmon (1971), an active contributor to a field is defined as a scientist from the year of his first publication until the year of his last contribution to a field during the whole time of his or her activity in the field—independent of the amount of activity. The indicator shows lower volatility than publication output. One has to bear in mind here that many of the mathematical or physics subfields experienced very low publication numbers, especially in the first decades of the 19th century. Because the individual last publications of authors entering science at the end of the 19th century are not covered by the „Catalogue“ increasing with time, we excluded the last 15 years from the computations. (For a graphical demonstration of that effect, see Wagner-Döbler & Berg, 1999.) Because the same is valid, at least in principle, for the first years of the 19th century for the individual first publications of authors, we excluded also the first five years from the computations. In addition, we excluded annual values equal or lower than two active contributors or publications because of their erratic statistical behaviour.

Findings

Figure 1 shows the frequency distribution of all enrolments of all semesters for the German universities, for better convenience as percentages of all enrolments. The parable— in a double-logarithmic scale—corresponds to a lognormal distribution, which is in agreement with Gibrat‘s model of firm growth as a multiplicative stochastic process and Gibrat‘s „Law of Proportional Growth“.

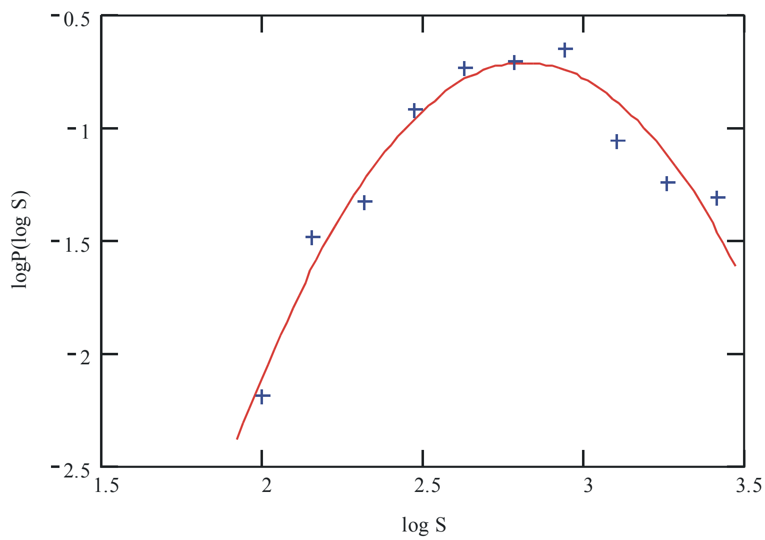

FIG. 1. Frequency distribution of enrolments of all semesters in all German universities 1830-1900 as percentages of all enrolments. Sources: Raw data from Prahl (1978); own computations. Graph fitted according to least squares of logarithmised values.

In Figure 2 the distribution $p(g)$ of the (decadicly logarithmised) annual growth rates $g = \log[S(t+1) / S(t)]$ of enrolments S is displayed with p on a logarithmic scale. The tent form of the distribution points to a Laplace distribution of the form

$$p(g) = \frac{1}{\sigma\sqrt{2}} \exp\left(\frac{-|g-\mu|\sqrt{2}}{\sigma}\right)$$

with standard deviation σ and average μ, in accordance with results of Stanley et al. We performed, in addition, Kolmogorov-Smirnov tests (KS tests) which yield the test statistics $D\sqrt{n} = 1.95$ for the Laplace and $D\sqrt{n} = 3.43$ for the Gaussian distribution (with the number of growth rates $n = 2814$). So both null hypotheses should be rejected with an error probability less than 1% but the test statistics is much more powerful for the Gaussian distribution. In the cited papers of Stanley et al. no result of any comparable statistical test is communicated. For KS test of Laplace distributions see Puig & Stephens, 2000.

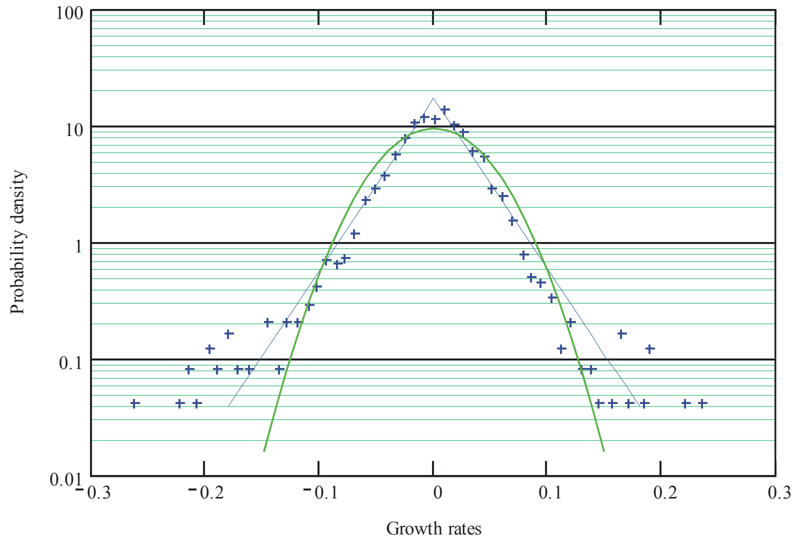

FIG. 2. Probability density of growth rates of student enrolments, German universities 1830-1900. The "tent" is the graph of the Laplace distribution, the parable that of the Gaussian distribution, both with standard deviation σ = 0.042 (obtained from data; average μ = 0.005 subtracted from data). Sources: See Fig. 1.

As we see in Figure 3, and in agreement with results of Stanley et al., the standard deviation of the growth rates is decreasing with size as a power law $\sigma(S) \sim S^{-\beta}$ with β = 0.29±0.04. There are time differences, however. In the period of university stagnation in Germany until 1870 we get β = 0.18±0.09. After 1870 a marked expansion of universities started, and we get a steeper curvature β = 0.34±0.04. The fitting is much better than for values before 1870 (Figure 4). Therefore, for universities we consider only the period after 1870 in what follows. In this period the distribution of growth rates does not conform a Laplace distribution, however, rather a Gaussian one. The KS test values with n = 1134 are $D\sqrt{n}$ = 1.62 for the Gaussian and $D\sqrt{n}$ = 1.85 for the Laplace distribution.

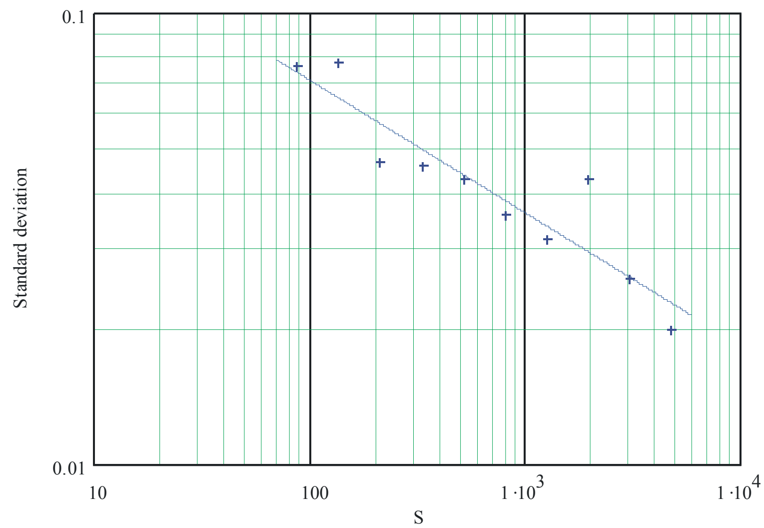

FIG. 3. Standard deviation of enrolment growth rates, plotted against enrolment size *S*. German universities 1830-1900. Sources: See Fig. 1. The ten classes have equidistant averages on the logarithmic *S* scale; therefore the points represent classes with different weights (s. Fig. 1); the line, however, was obtained by linear regression without weights.

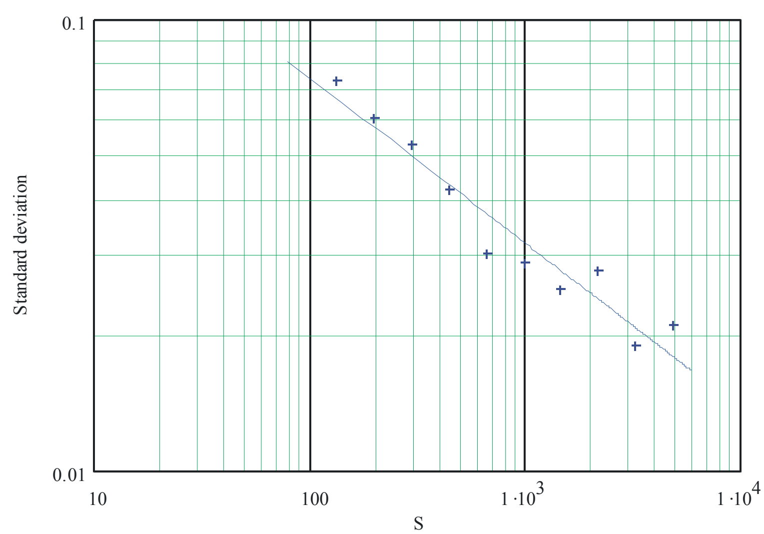

FIG. 4. Standard deviation of enrolment growth rates, plotted against enrolment size *S*. German universities 1871-1900. Sources: See Fig. 1. The ten points represent classes of the same weight: they all include 113 growth rates. Their *S* values are the geometric mean of their minimum and their maximum value of *S*. The line was obtained by linear regression.

Now, in agreement with Stanley et al., we choose bigger classes of growth rates than in

Figure 3 and calculate the corresponding conditional probability distributions $p(g \mid S)$ of the growth rates. If we partition the rates into three classes of equal size, two of the distributions $p(g \mid S)$ conform the Gaussian distribution (KS test values are smaller than the critical value 0.89 for 5% error probability); the third one (with lowest values of S) has a better test value for the Laplace (1.16) than for the Gaussian distribution (1.59). These findings are in contrast to those of Stanley et al. who in all cases communicated that also the conditional distributions tend to show the tent shape when plotted on logarithmic scale.

We now turn to the two scientific disciplines. The size distribution of the 53 physics subfields in the 19th century, indicated through journal articles, is approximately lognormal. The probability density of growth rates conforms more to a Laplace than to a Gaussian distribution, as Figure 5(a) shows (KS test values 2.67 and 4.14, respectively). The same is true for the 34 mathematical subfields, as Figure 5(b) shows (KS test values 2.07 and 3.39, respectively). Figure 6 shows that the standard deviations plotted against size of the physics and mathematical subfields conform to power laws with $\beta = 0.34 \pm 0.03$ and $\beta = 0.46 \pm 0.02$, respectively.

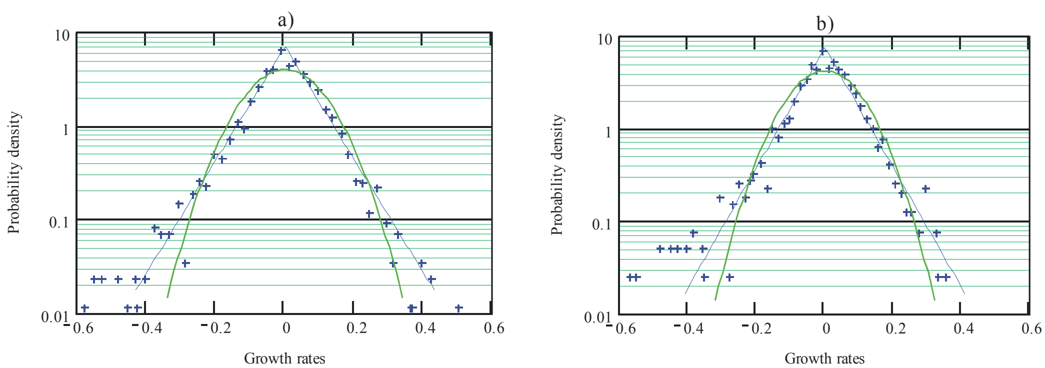

FIG. 5. The distribution of growth rates of (a) physics and (b) mathematics subfields 1800-1900 indicated through numbers of papers (see also Fig. 2). Sources: Catalogue of Scientific Papers (transformed into databases by Wagner-Döbler and Berg, 1996, 1999), and own calculations.

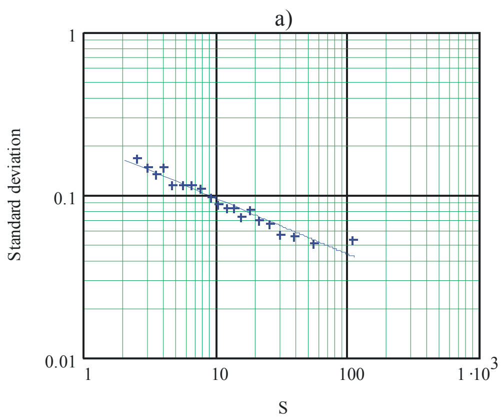

FIG. 6. Standard deviation of subfield growth rates, plotted against subfield size S: (a) physics and (b) mathematics 1800-1900. Sources: See Fig. 5. The 20 points of both disciplines represent classes of the same weight: in physics they include 203 growth rates, in mathematics 121. Their values on the S axis are geometric means of classes' minimum and maximum value of S. The lines were obtained by linear regression.

The conditional probability distributions $p(g \mid S)$ of the growth rates of physics and mathematics subfields, in both cases divided into three classes of equal size, look more like normally and not like Laplace distributed: with the exception of the class of large physical subfields all KS test values for the Gaussian distribution are smaller than for the Laplace distribution (cf. Figure 7). In the scaled plot (Figure 8) they nonetheless „collapse" quite well onto one single graph, in analogy to the results of Stanley et al. The results of using „active contributors" instead of publications as indicators are not shown here because of lack of space. They are in agreement with the results presented here with publications. Here we obtained for physical subfields $\beta = 0.51 \pm 0.03$ (25 classes with 147 rates), and in mathematics $\beta = 0.57 \pm 0.03$ (16 classes with 121 rates). For both disciplines the growth rates have Laplace-like distributions $p(S)$ and conditional distributions $p(g \mid S)$ (in three classes), which have smaller KS test values for the Gaussian distribution.

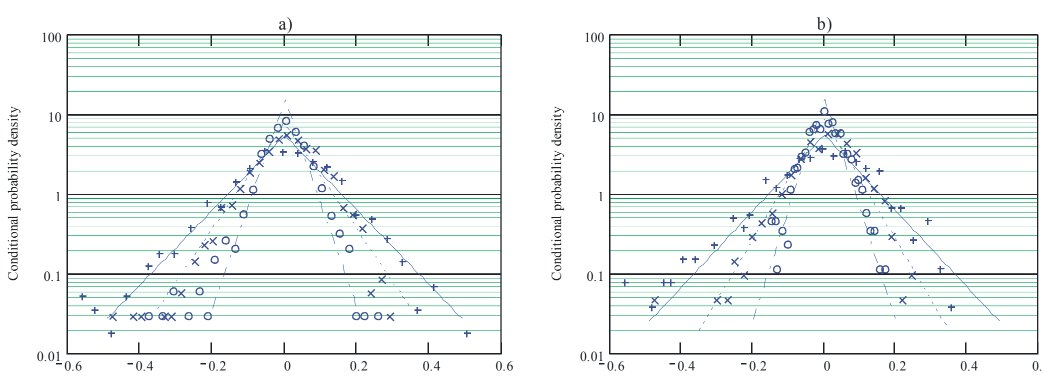

FIG. 7. Conditional probability distributions $p(g \mid S)$ of (a) physics and (b) mathematics

subfields growth rates *g* of three classes in each discipline. The classes include each about 1360 rates in physics and about 800 rates in mathematics.

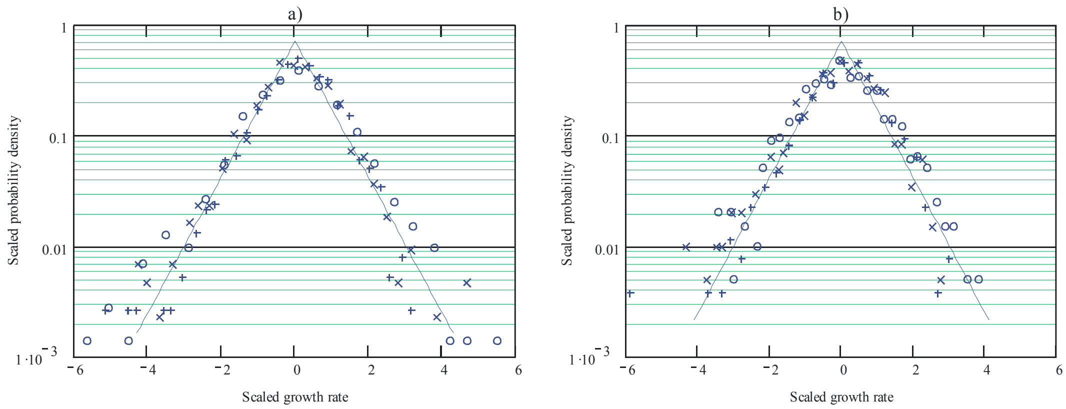

FIG. 8. Scaled plots of the distributions of Fig. 7: (a) physics, (b) mathematics.

Discussion

We demonstrated that the same apparently universal statistical behaviour of growth rates is yielded for the growth dynamics of student enrolments and of scientific activity in the 19th century as in the works of Stanley et al. regarding university research at the end of the 20th century. An exception is the not Laplace but rather normally distributed probability distributions of university enrolments in Germany after 1870. A further difference is, that in nearly all cases analysed by us the conditional distributions $p(g \mid S)$ (constructed as in the Stanley papers) show also a more normal than a Laplace-like form. We think, however, that these are nonessential differences. In which cases Laplace distributions (as observed by Stanley et al., and in some cases also by us) appear will be discussed in a further paper.

Obviously, weak competitive structures of the 19th century lead to similar quantitative behaviour of growth dynamics as the strong competitive conditions, which are said to prevail at the end of the 20th century. Consequently, any interpretations of the 20th-century results cited above which refer to competitive structures of research activities are built on weak ground insofar (correct however the claims as such may be, and independent from our agreement with them). But we cannot find, in addition, that competition is defined precisely and that convincing indicators for it were presented so far, at least in scientometrics. Thus we interpret the characteristics of the growth dynamics as reflections of self-organization of complex institutions and communities. We intend to perform a comparative analysis of the growth dynamics of research activities before and after the „Wende“ in countries of the former socialist block. It has the aim to compare growth dynamics of the same system under quite different economic regimes.

Finally we point out that other quantitative similarities between the time-behaviour of economic activities and research activities over centuries (Wagner-Döbler, 1998), and economic activities and education system development over the last two centuries (Müller-Benedict, 1991) are apparent: their cyclical or quasi-cyclical character at least on a time scale of years and of decades, respectively. To be sure, the different characteristics of different phases of development of areas under examinations have to be paid more attention (cf. Sutton, 2002, p. 580). For mathematical logic, at least, Goffman and Harmon (1971) interpreted those different phases as recurring waves of epidemics.

Acknowledgement

We thank Prof. Dr. Walther Umstätter for interesting discussions on the subject of the paper.

References

Amaral, L.A.N., Gopikrishnan, P., Matia, K., Plerou, V., & Stanley, H.E. (2001). Application of statistical physics methods and concepts to the study of science & technology systems. *Scientometrics, 51*, 9-36.

Goffman, W., & Harmon, G. (1971). Mathematical approach to the prediction of scientific discovery. *Nature, 229*, 103-104.

Moed, H. F., & Luwel, M. (1999). The business of research. *Nature, 400*, 411-412.

Müller-Benedict, V. (1991). *Akademikerprognosen und die Dynamik des Hochschulsystems. Eine statistisch-historische Untersuchung.* Frankfurt/M.: Campus-Verlag.

Plerou, V., Amaral, L.A.N., Gopikrishnan, P., Meyer, M., & Stanley, H.E. (1999). Similarities between the growth dynamics of university research and of competitive economic activities. *Nature, 400*, 433-437.

Prahl, H.-W. (1978). *Sozialgeschichte des Hochschulwesens*. München: Kösel.

Puig, P., & Stephens, M.A. (2000). Tests of fit for the Laplace distribution, with applications. *Technometrics, 42,* 417-423.

Stanley, M.H.R., Amaral, L.A.N., Buldyrev, S.V., Havlin, S., Leschhorn, H., Maass, P., Salinger, M.A., & Stanley, H.E. (1996). Scaling behaviour in the growth of companies. *Nature, 379*, 804-806.

Tabah, A.N. (1999). Literature dynamics: Studies on growth, diffusion, and epidemics. *Annual Review of Information Science and Technology, 34,* 249-286.

Wagner-Döbler, R., & Berg, J. (1996). Nineteenth-century mathematics in the mirror of its literature. A quantitative approach. *Historia Mathematica, 23*, 288-318.

Wagner-Döbler, R., & Berg, J. (1999). Physics 1800-1900: A quantitative outline. *Scientometrics, 46*, 213-285.

Wagner-Döbler, R. (1998). Scientometric evidence for the existence of long economic growth cycles in Europe 1500-1900. *Scientometrics, 41*, 201-208.

Wagner-Döbler, R. (1998a). Kognitive Mobilität. Eine makroskopische Untersuchung der Wanderung von Wissenschaftlern zwischen Forschungsgebieten am Beispiel der

Mathematik. *Journal for General Philosophy of Science*, *29*, 265-287. - Revised English version: Wagner-Döbler, R. (1999). Cognitive Mobility. In P.S. Nagpaul, K. C. Garg, B.M. Gupta, et al. (Eds.), *Emerging Trends in Scientometrics*. *Essays in honour of Ashok Jain* (pp. 225-245). New Delhi: Allied Publishers. - 2nd ed. 2000.

A revised version of this contribution has been published as:

Frank Havemann, Michael Heinz and Roland Wagner-Döbler (2004). Growth dynamics of German university enrolments and of scientific disciplines in the 19th century: scaling behaviour under weak competitive pressure. Scientometrics, 60(3), 283-294.