

Publication Note: A slightly revised version of this paper has been published in *Economics and Technology*, (O.Grandstrand, ed.), Amsterdam: Elsevier, 1994.

Positive Feedbacks and Research Productivity in Science: Reopening Another Black Box

by

Paul A. David

*Professor of Economics, Stanford University,
and Visiting Fellow, All Souls College, Oxford.*

March 1992
Revised: July 1993

Acknowledgements

Comments and suggestions made by members of the Science, Technology and Economics Workshop at Stanford University (October 1992), the able research assistance of Phillip Lim, and the support of the Science and Society Program of the Andrew W. Mellon Foundation, all are gratefully acknowledged. This revision of a paper originally presented to the conference on Innovation and Technological Change in Mastrand, Sweden, in August 1992, was undertaken during the author's tenure of a Visiting Fellowship at All Souls College, Oxford.

Address correspondence to: All Souls College, Oxford OX1 4AL, England.
Fax:44+(0)865 + 279299; E-Mail:pdavid@herald.ox.ac.uk

1. Introduction: Economics and the Black Box of Science

In view of the widely acknowledged importance of the resource allocation decisions that impinge upon scientific research, and given the tendencies toward rampant disciplinary imperialism that economists as a group are supposed to harbor, it is really rather remarkable that the economics of science has remained so underdeveloped a field of inquiry.

Today more than ever before it is apparent that the scientific knowledge-base constitutes a productive asset that will have a strategic role in the future economic progress of the industrially advanced and the developing societies alike. The informational assets yielded by scientific research are intangible, of course. Yet, their role is comparable *in some respects* to the one played during the 19th and early 20th centuries by stocks of tangible reproducible capital in the form of machines, factories and transportation and communication infrastructure. From this perspective, to study the organizations and institutions that mobilize resources for scientific research, and which thereby govern where and at what pace accretions are made in the science knowledge-base, could be seen as tantamount to inquiring into the workings of "the 21st century's machine tool sector."¹ Economists who would most readily assent to the validity of that analogy, and even those who have been most assiduous in exploring the interior of the "black box" of technological research and innovation, nevertheless become very diffident (even to the point of recalcitrance) as they approach the adjacent territory of pure science.² They have offered no dispute to allowing sociologists of science to retain exclusive dominion over that "heart of darkness".

Consequently, by comparison with the mass of theoretical and empirical studies that deal with the economics of "applied" R&D activities (undertaken largely by the private sector in the U.S., and supported by company or government agency funding), only a small segment of the economics literature has concerned itself with the production and distribution of "basic" scientific knowledge by

¹ This simile has been suggested to me by Alfonso Gambardella.

² The focus of attention upon private sector R&D, to the comparative neglect of the economics of publically supported science, remains somewhat more rationalizable in the situation of the U.S. today than it is in western Europe. In the latter region the (public and private) not-for-profit research sector rivals the business sector as an employer of scientists and engineers -- being roughly on a par in absolute size with the U.S.'s not-for-profit research sector, but far smaller as an employer than the U.S. business R&D sector. During 1985-89 the share of the not-for-profit sector in national research employment in the European Community was (47 percent) almost twice as large as it was in the U.S.(26 percent). "Researchers" are defined following the OECD Frascati Manual as "scientists and engineers engaged in the conception or creation of new knowledge, products, processes, methods and systems." The "not-for-profit" sector includes Government, Higher Education, and Private Non-Profit research employments. See OECD (1992): Tables 28-31.

researchers working in university settings. Insofar as it does treat the subject at all, our discipline's approach has been simply to consider the organized quest for fundamental or basic scientific knowledge as an activity plagued by the same "public goods" problems that arise in the case of "applied" R&D. Only, perhaps, moreso. Kenneth Arrow's (1962) classic paper of 30 years ago suggested that for private agents to benefit from investing in "pure" science would be especially problematic, in view of the greater uncertainty surrounding the nature and implications of basic research findings, and the fact that one of the main uses of the information thus obtained would likely be as an input into future research.

One legacy of this venerable analytical tradition has been the tendency among economists to emphasize the greater unpredictability of scientific findings, and the wider ramifications of their importance for researchers in many sub-branches of science, as the hall-marks of "basic research". That has had the unfortunate consequence of muddling two quite different taxonomic principles.³ One would have us draw a line between basic and "applied" by reference essentially to epistemological criteria, that is, to differences in the nature of the scientific findings themselves. The other principle of classification suggests distinguishing between *research modes or processes*, rather than among *research outcomes*, according to whether the former are more or less conformable with the basic researcher's intention of rapidly increasing the stock of reliable knowledge.⁴ The latter approach, which directs attention to the distinctive organizational and institutional arrangements for the pursuit of scientific knowledge, has seemed to me the more fruitful of the two.⁵ In the discussion that follows, therefore, it is the social organization of the inquiry process (and the related motivation of the researchers) that I have foremost in mind when referring to "basic" research in science. This usage is similar in spirit to the National Science Foundation official definition which makes the test of the *sine qua non* of "basic research" the researchers' intentions or objective of gaining "a fuller knowledge or understanding of the subject under study rather than a practical knowledge thereof."⁶

³ There are, of course, notable exceptions, e.g., Rosenberg (1990), and Trajtenberg, Henderson and Jaffe (1992).

⁴ This purpose, in general, can readily be distinguished from the goal of maximizing the *private* wealth of of the researchers themselves, or that of their patrons and employers, by means of acquiring possession of novel scientific information.

⁵ See Dasgupta and David (1987) for the initial formulation of a functionalist interpretation of the institutions of open science (including the norms of behavior for members of the "Republic of Science") as those "designed" to maximize the rate of growth of the stock of scientific knowledge, whereas the institutions of proprietary research (sometimes described as belonging to the "Realm of Technology") are better suited to maximizing the flow of economic rents from the stock of scientific knowledge.

Going somewhat beyond the NSF, I take the pursuit of "practicality" to be associated with the existence of an immediately foreseeable private profit-seeking (or national security) application, which in turn militates strongly against conducting the research in question under the "open science" norms of disclosure and cooperation that are publically attested to by members of the community of university-based scientists.⁷

While the economic rationale for public subsidization of basic science has been elaborated around Nelson's (1959) and Arrow's (1962) discussion of "the appropriability problem",⁸ practical decision-making in the sphere of science policy has not been able to draw much in the way of specific and detailed recommendations from those insights, or from subsequent examinations of "market failure" pathologies involving the production and distribution of information. Nor has economics yet had much to say about which types of contractual arrangements, institutional forms, and organizational policies prove to be especially conducive (or inimical to) the processes of "knowledge transfer". Until very recently most economists studying the insides of the "black box" of private sector R&D have not thought it important to understand the variety of mechanisms that were responsible for the "spillovers" that theoretical analysis holds would result in industrial innovators reaping private benefits from the work of publically supported academic and governmental research organizations. Thus, these latter structures have for too long been treated as though they were not only "black boxes", but ones that bore the stencilling: "This May Belong to Pandora" -- and so were better left tightly sealed.⁹ Can it be surprising, then, that economic analysts have been unprepared to enter into the recent public policy discussions of how best to connect university with industry

⁶ See National Science Foundation (1985: p.221). Correspondingly, "applied research" refers to activities that aim to gain "knowledge or understanding necessary for determining the means by which a recognized and specific need may be met."

⁷ On the organizational and institutional distinction between open science and proprietary research, and the centrality of norms regarding information disclosure, see the discussion in Dasgupta and David (1987), (1988), (1992). It is not maintained that academic researchers perfectly adhere to the norms of full disclosure and cooperation, any more than it is suggested that private corporations will not conduct some research activities under organizational rules that are indistinguishable from those governing university research laboratories (see Rosenberg (1990).

⁸ See Mowery (1983), Mowery and Rosenberg (1989).

⁹ For several recent and welcome efforts to measure spillovers from academic research to company financed R&D, and, more broadly, to begin systematic quantitative studies of the interdependence between university and industry research activities, see Jaffe (1986), Jaffe (1989), Mansfield (1991), Jaffe, Trajtenberg and Henderson (1992). On the limitations of the standard economic approach to measuring the social rate of return on basic research, see David, Mowery and SteimueLLer (1992).

research, and so promote the rapid application and commercial exploitation of scientific discoveries and new techniques produced by academic researchers?

The state of affairs just described is both unsatisfactory and unnecessary. It represents numerous missed opportunities for intellectual satisfaction as well as public service, opportunities to investigate inherently fascinating social phenomena, and thereby, it is to be hoped, to contribute to the formulation of more enlightened policies affecting the funding and conduct of scientific research. The present essay is addressed more immediately to the first of these opportunities than to the second, in the belief that the cause of better science and technology policies eventually will be well served if many others can be induced to join in developing a "new economics science".¹⁰

Rather than casting my net broadly on this occasion, however, I have chosen to reexamine the nexus between the institutional structures of public patronage and collegiate reputation-based reward systems that are characteristic of modern academic science communities on the one hand, and, on the other hand, a particular set of quantitative patterns observable in the allocation of basic research inputs and outputs in many scientific fields. The robust empirical phenomena to which I am referring are *the marked stratification of the research community in regard to reputational status and access to research resources, and the extremely unequal productivity of research scientists as measured by their publication rates*. These phenomena are well-documented and have been studied extensively by sociologists and "scientometricians", among whom there has emerged a broad consensus that the allocation of reputations and resources in scientific communities are subject to the operation of dynamic processes characterized by one or more forms of self-reinforcement, or positive feedbacks.

Although thoroughly familiar to a whole generation of researchers in the sociology of science, the empirical study of scientific productivity and the associated explanatory hypotheses that have been proposed as theories of "cumulative advantage",¹¹ may be surprisingly novel even for economists who can claim some professional concern with the allocation of resources in science and technology. A reconsideration from the economist's perspective of the major points of agreement to which this line of inquiry has led, therefore, would seem to be a task worth undertaking. Yet, this way of putting

¹⁰ For related work in this vein, see Dasgupta and David (1992), which synthesizes the approach of modern industrial organization economics (specifically, the analysis of behavior of agents engaged in games of incomplete information) with insights and findings drawn from the literature of the sociology of science associated with Robert K. Merton (1973) and his students; David (1991) draws upon ideas from modern agency theory to propose an explanation for the historical emergence of open science institutions in western Europe.

¹¹ Processes of "cumulative advantage" were first articulated by Merton (1968, reprinted as Ch.20 of Merton 1973) under the label of "Matthew Effect, and subsequently generalized as a central explanatory principle by Merton and his students: Cole and Cole (1973), Zuckerman and Merton (1972), Zuckerman (1977), Gaston (1978). See Allison, Krauze and Long (1982), and Merton (1988), for more recent reviews and further references.

it gives the misleading impression that the subject properly can be regarded as thoroughly settled. The sociology of science, having shed very considerable light upon the causes of inequality in scientific productivity, nonetheless, has left some explanatory work still to be done -- perhaps in conjunction with its nascent sister-discipline, the economics of science. Furthermore, the proper interpretations of some of the more widely accepted explanatory schemes turn out not to be agreed upon in every detail, which is to say that some of the underlying mechanisms remain incompletely specified, possibly statistically underidentified, and unevaluated as to their relative quantitative importance. In addition, the alternative interpretations that vie for acceptance in this literature would carry rather contradictory implications for appraisals of the allocative efficiency of the institutions of modern science, and hence for science policy-making. Here is a worthy challenge for the new economics of science.

Should another reason to read further on be needed, one more can be supplied: the sociological and bibliometric literature upon which the following sections has drawn contains a considerable amount of pioneering work on stochastic processes with positive feedback, and therefore merits the attention from economists who share the growing interest in the general phenomenon of *path dependence* in the dynamics of social systems.¹²

2. Productivity Inequalities in Science and their Economic Significance

From the perspective of the new economics of science, the problem to be taken up here concerns the allocative efficiency of the institutional structures that characterize "the Republic of Science".¹³ More precisely, it concerns the way in which a collegiate reputational-reward system and various derivative resource allocation mechanisms, such as those involving the award of grants on the basis of peer evaluations within open science communities, affects the utilization of research talent for the production of knowledge communicated in scientific publications. This problem presents itself in a stark form when one simply considers how enormous are the differences among trained academic researchers in the number of papers they publish. The phenomenon of concentration in scholarly productivity has been studied very thoroughly in respect to the natural sciences: most scientists publish a very limited number of papers over the course of their entire career, whereas a small proportion of them publish a great number. The skew in the frequency distribution of scientists by number of papers published is sufficiently pronounced that one may say that the 10 to 15 percent of

¹² On path-dependence in economics, see e.g., David (1985), David (1988), David (1992); Arthur, Ermoliev and Kaniovski (1987), Arthur (1989), Arthur (1990), Durlauf (1991).

¹³ See Polanyi (1962) for the origin of the phrase in quotes, and Dasgupta and David (1987, and 1992) for discussion of its significance.

scientists who constitute the productivity (and prestige) elite are responsible for about half of all the science produced.¹⁴

2.1 Publication and the Measurement of Scientific Production

Of course, to be able to speak of production and productivity meaningfully, one must have in mind some measureable "product". It is recalled that a sign bearing the legend "Work, Finish, Publish" hung in the laboratory of Michael Faraday, the great British scientist of the nineteenth century.¹⁵ The publication, and typically the "paper" is the tangible product of research in academic science organizations.

One of the reasons why academic researchers are so obsessed with publication is that (as the work of Robert K. Merton (1973) and his students has emphasized) the reward structure of the science community has been set up in such a way that it elicits compliance (even from the reluctant) with the norms of "openness" -- the speedy disclosure of new findings and their submission to others for verification, application, and extension. Recognition of one's contributions and consequent collegiate reputation, or esteem in the eyes of one's scientific colleagues, are the key currency of the open science reward system. To this are tied the academic researcher's material rewards, such as salary and job tenure, and access to the human resources and physical facilities that scientists typically need to produce published results.

Given the centrality of publications to the reward system that guides the allocation of resources, it is hardly surprising that sociologists of science, and the bibliometricians they inspired, have devoted much thought and effort to measuring the productivity of scientists and the reputational impact of those "contributions" by using data on scientific publications and the published citations made thereto. Suffice it to say that this research has established that publication counts are strongly correlated with "impact" as measured by peer appraisals (including peer evaluations of applications for research grants), and that citation counts are strongly correlated with scientists' reputations and the prestige of their honorific awards.¹⁶

¹⁴ See Price (1963), and below (section 2.2) for further discussion of subsequent studies that have found this pattern obtains in every scientific discipline examined, and for the U.S. and all other countries whose scientific outputs (publications) have been studied. Andrews (1979) finds it holds for research groups as well as individuals.

¹⁵ See Cole and Zuckerman (1984: pp. 218).

¹⁶ See Cole and Cole (1973), Gaston (1973), Zuckerman (1977), Cole, Rubin and Cole (1978), Cole and Zuckerman (1984); Narin (1976) reviews 24 papers which, on balance, supported the conclusion that publication and citation count indicators of productivity were positively correlated with peer evaluation and other measures of scientific impact, while Broadus (1977: pp. 310-313) concluded a similar review of the literature with the conclusion that despite some inconsistencies there was "a strong relationship between citation counts and

Those are widely accepted conclusions about *the research process* rather than about research *substance*; they relate primarily to the significance of publication counts and citations in the social and economic organization of research carried on within the Republic of Science. The validity of these measures as objective indicators of substantive research output is quite a different matter. As Cole and Zuckerman (1984: pp. 238) declare: "there is no agreement among sociologists and historians of science on the substantive significance of unequal rates of citation." Indeed, the use of citation frequency as an indicator of scientific (scholarly) quality or substantive importance is fraught with controversy, and not only due to the open hostility of the reaction among scientists to the early enthusiasts for citation-count analysis.¹⁷ Much of what we know of the difficulties and drawbacks surrounding the counting of publications and citations are the fruit of the careful bibliometric and "scientometric" studies inspired by this line of inquiry in the sociology of science.¹⁸

Some of the most problematic aspects of these measures of scientific "output" have entered the body of common knowledge, and therefore need only be briefly noted:

- (1) The "publication", and even the "paper" are arbitrary and largely undefined units of measurement.
- (2) Publications may vary in quality, as may the venues in which they are exposed to the public -- publishing houses and journals vary in their minimum standards, in their repute, and in their readership.
- (3) Citation practices may be sloppy, citation indexes are often misled by homonyms, and the leading databases used do not report any but the first author -- which can introduce not only correctable biases due to alphabetization of authors, but more subtle and uncorrected biases due to the existence of different name-ordering conventions in different branches of the sciences.
- (4) The quality of a scientist's contributions is not accurately gauged by weighting publications with citation frequency-based measures of scientific journals' "impacts", nor by raw citation counts, because citation data itself is so problematic.

And so forth, and so on. Which is not to minimize the difficulties, nay, the potential dangers in casual reliance upon these quantitative indicators of scientific "output", but simply to reassert the point that the robust patterns that emerge repeatedly in studies of scientific publication/productivity are not readily discounted as sheer artifacts of measurement. To a closer consideration of the nature, and the potential economic significance of those patterns we may therefore turn.

2.2 The Concentration of Research Productivity in Science

other measures of evaluating excellence". Bensman (1982) provides a fine overview of the bibliometric literature on these issues.

¹⁷ See, for example Wade (1975), in response to Cole and Cole (1973).

¹⁸ For references and further discussion, see e.g., the chapter by Stephan and Levin, among other contributions in van Raan (1988).

Three bald facts about the productivities of scientists as measured by unweighted publication counts can command our attention here.¹⁹ Firstly, the distribution of individual researchers' productivities typically is extremely unequal, being characterized -- in each of many different areas of science -- by a long upper tail in the frequency distribution of the number of papers published in a specified time interval. Secondly, the identities of the scientists occupying the upper and lower fractiles of the productivity distributions remain very stable over much of the lifetime of a cohort of scientists (adjusted for premature death), and the distribution of lifetime productivities is even more left-skewed than those for sub-intervals. Thirdly, and possibly related to the immediately foregoing phenomenon, in true cohort data (and also in synthetic cohort data for scientific sub-disciplines that have been "settled", rather than undergoing significant transformations in research paradigms), the dispersion of current period publication rates is found to increase over the professional life of the cohort.

2.2.1 Skew-Distributions

That the distribution of scientific output is extremely skewed, with the consequence that most of the papers published in a given area of specialized inquiry are the work of a small, highly prolific minority, is scarcely a novel finding. It first was documented over 60 years ago in a landmark paper by Alfred J. Lotka (1926), who was seeking "to determine, if possible, the part which men of different calibre contribute to the progress of science." To that end he compiled the number of entries in the *Chemical Abstracts* decennial index of 1907-14, just for those authors whose names began with "A" or "B". On the basis of that, and a similarly constructed dataset from Felix Auerbach's (1910) history of publications in the whole range of physics covering the period until 1900, he formulated the inverse square distribution known as "Lotka's Law": if k is the number of scientists who publish 1 paper, then the number publishing n papers in the same field is $f(n) = k/n^2$, so that if 100 scientists each have published 1 paper, 25 will have published 2, 11 will have 3 to their credit, 4 will have 5, etc. This rule generates a "left-skewed" frequency distribution, with a long right-hand tail and a single mode at 1 paper. The concentration of publications is very pronounced, inasmuch as approximately two-thirds of the scientists in the field publish but a single paper, while the upper decile of productive scientists accounts for approximately half of the aggregate output of papers in the field in question.

Derek deSolla Price (1963), who pioneered in the quantitative study of science and scientists' productivity, focused attention upon the invariance of the latter measure of concentration by

¹⁹ A fourth empirical regularity, which deserves separate discussion and so has been left for treatment on a subsequent occasion, is the puzzling "fact" shown by more than 50 studies of scientists in various fields: gender matters in scientific productivity. Women scientists publish less than men, although the distributions by publication counts also are highly skewed. See Cole and Zuckerman (1984), and Zuckerman, Cole and Bruer (1991) on this subject. An economist's tentative approach to understanding this phenomenon has been offered by David (1992b).

formulating a variant of Lotka's Law that permits quick calculation of the size of the most prolific subset of any population of scientists who collectively are responsible for half of the total publications. According to "Price's Law", one half of the total output of papers published by a population of P scientists will be the work of the $(P)^{1/2}$ most productive members of the population.²⁰ Of course, the important thing is not the exact distributional form describing the distribution of scientific productivities. While Lotka's Law has held up remarkably well over time, so much so that some bibliometrician's have been tempted to regard it as an invariant feature of academic research communities wherever they exist (see Bensman (1982: p. 280-281), other statistical distributions, such as the negative binomial distribution, have been shown to provide good fits to observed publication count data.²¹ More generally, it has been shown that such empirical distributions are well described by a general class of "skew" or hyperbolic distribution functions that Herbert Simon (1957: Ch.9) characterized and showed to be given by the Beta Function.²² The Beta Function, which is not to be confused with the Beta frequency density function, figured centrally in the statistical work of Price (1976) on cumulative advantage processes, as is noted below.

2.2.2 Stratification and Persisting Hierarchies of Productivity

The second "big fact" deserving emphasis, and the property of the empirical distributions of scientists by publication activity that renders the latter especially intriguing for social scientists, is that these left-skewed distributions reflect an underlying, *intertemporally stable productivity stratification* of the population. After all, left-skewed distributions in and of themselves are not a rarity; in economics one encounters such quite frequently, in the size distributions of firms, in income distributions, wealth distributions, and many other dimensions of economic performance.²³ It is well

²⁰ Thus, in a group of 100 scientists, the most prolific 10 percent (i.e., the upper decile of the productivity distribution) will have produced half of the total. See Allison, Price, et al.(1976) on "Price's Law".

²¹ See Allison, Long and Krauze (1982:p.620) for references.

²² The Beta Function, otherwise known as Euler's First Integral, may be written as:

$$B(a,b) = B(b,a) = \int_0^1 x^{a-1} (1-x)^{b-1} dx = [(a-1)!(b-1)!]/[(a+b - 1)!].$$

²³ It may be interesting to note that the distributions of current scientific outputs (measured by published papers) show greater degrees of inequality that do modern income distributions, in which, for example, the top 5 percent of income recipients in the Social Security Population during 1951-1969 accounted for 21% of the income received (see Brittain 1972, reproduced in Williamson and Lindert (1980:Appendix F). The distribution of cumulated publications (over the lifetimes of individual surviving scientists) is much less concentrated than is the distribution of wealth-holding, however. See, e.g. Thurow (1975). The latter is not

known that the multiplicative interaction of randomly distributed variates will generate left-skewed distributions; Gibrat's Law tell us that a population of firms whose individual growth rates are normally distributed will become lognormally distributed in size (see Atchison and Brown (1957)). Under the operation of such a process individuals, or firms would be found to be transiting through the whole size distribution over time, and the composition of the upper and lower tails of the distribution would gradually turn over.

That does not characterize the situation in science. The arresting observation therefore is not simply that there are unusually marked inequalities in "publication attainments" among scientists during a given interval of time, but, rather, that the pronounced productivity stratification in science existing at any moment reflects the persistence of a particular hierarchical ordering throughout most of the life of a cohort. Once scientists enter the current productivity elite, it is rare for them to exit from it in the next period; and the same holds at the lower extreme of the productivity distribution. Something more than simple randomness in transitions between productivity states must be at work behind the scenes to generate the these intertemporal stabilities in the statification patterns observed among scientists.

2.2.3 Increasing Concentration of Productivity

The third of our empirical phenomena is the tendency for publications of a cohort of scientists to grow increasingly concentrated as the cohort ages, piling up more and more on the curriculum vitae of its most prolific members. Thus, scale-invariant measures of dispersion in the distributions of current publication rates such as the Gini index, and the coefficient of variation, are found to increase with the ageing of a cohort of scientists in a given field. For example, synthetic cohort data on publications (and citations) of U.S. scientists, constructed by Allison and Stewart (1974) for the period up to 1966, show the Gini-coefficient of inequality rising more or less linearly with the elapse of time since receipt of the Ph.D. in Chemistry, Physics, and Mathematics.²⁴ More recent work reported by Allison, Long and Krauze (1982) has found the same phenomenon in longitudinal observations on true cohorts of chemists, biochemists.

The observed tendency towards increases in the concentration of productivity over the life of scientific cohorts can be paired with the previously noted "fact" regarding the temporal persistence of productive and reputational hierarchies within in the various field of research. It it provides a second hint that if there is an underlying stochastic process generating these distributions, it cannot be a

so surprising, if only because intellectual "wealth" in the form of publications is not a heritable form of property.

²⁴ Because these are synthetic cohort data, the pure cohort effects may be obscured by the effects of (period) changes in the structure of the discipline, such as were occurring in Biology, during the 1960s and early 1970s. By comparison the three fields mentioned in the text were much more "settled", and it is there that the synthetic cohort technique is likely to reveal the true direction of the cohort effects.

simple, stationary one involving the heterogeneity of researchers with respect to their scientific "talent" -- or at least those talents that would translate into the propensity to follow Faraday's injunction to "work, finish and publish". Something more than a stationary distribution of publication propensities, which latter could correspond to an underlying distribution of inherent abilities, would appear to be generating the increasing concentration in productivity within these cohorts of scientists. On a simplified view this suggests that the phenomena reviewed here may be understood as arising from the interplay between a heterogeneous population of researchers, and an environment whose reward system acts to reinforce and amplify the effects of initially limited differences in productive potential, thereby creating extremely pronounced disparities in realized productivity. To be sure, even the identification in this formulation of persisting differences among scientists with respect to their propensity to emit published findings as indicative of underlying differences in inherent ability, would be too facile a way of solving the problem addressed by Lotka (1926). Directly associating individual productivity with research talent would provide a trivial to answer the question of what portions researchers "of different calibre" contributed to the "progress of scientific knowledge" in the form of published papers.

But, whether or not Alfred Lotka fully appreciated the complexity of the task of satisfactorily answering the question he posed for himself, the growth of the scientific establishment during the intervening sixty years, and its influence upon modern societies undoubted has made the issues at stake of much greater urgency.

3. Explanations and Interpretations: "Sacred Spark", "Matthew Effect" and "Cumulative Advantage"

Why should there be such great inequality in scientific productivity, and in the recognition, fame and other rewards that are heaped upon the prolific few? Are those prizes necessary incentives or simply a kind of economic rent that is being reaped by the exceptionally talented, or, perhaps, the exceptionally lucky few? Does it really reflect a highly unequal distribution of underlying talents, or compulsions to not only do science but to put it into print? Are we justified in supposing that the prevailing system of allocating recognition and material resources on the basis of performance-based reputations is managing to identify the most talented researchers and put the wherewithal to do science at their disposal? If the pronounced differences among scientists in their productivity corresponds to the variation of ability among them, then the institutions and social norms comprising the reward system might be viewed to be functioning well, in the sense that the highly productive members of the population have been picked out of the mass and given the resources to realize their potential -- even though superfluous psychic and tangible rents may be enjoyed by members of the honor-bedecked elite.

Even leaving aside the ambiguities surrounding use of publications, or citation-weighted publications, as measures of useful additions to scientific knowledge (and they are not easily put aside, it must be said), does the prevailing organizational regime in modern science -- a regime that at

each step awards individual and group "successes" with access to the means for future successes -- constitute the most "productive" allocative system that one might have? What warrant is there for supposing that the available scientific talent is being efficiently utilized by a system that results in almost 80 percent of all papers being attributable to only one third of the trained scientists who embark upon academic careers? Is there a problem of serious under-utilization of available talent, or, of inadequately selective recruitment? Such has been implied on past occasions, by the suggestion that the progress of basic science (in the form of published findings) really would not be slowed substantially if a considerable reduction occurred in the numbers of university scientists -- provided, of course, that the bulk of the less productive researchers were the ones who were screened out.²⁵ Alternatively, the problem may merely be one of a serious failure of the attribution mechanisms upon which the science reward system, and the quantitative measures of productivity rely. They both are evidently inadequate when it comes to indicating the extent to which published "contributions" credited to the prolific elite depend upon the complementary efforts and abilities of the supporting cast -- that myriad of trained scientific co-workers whose performance as research associates, postdoctoral students, technical and administrative assistants, and as teachers responsible for preparing the new cadres of researchers, will not be evaluated on the basis of publications in the scientific journals and the citations made thereto.

In one way or another, each of the foregoing barrage of loaded questions invites a reconsideration of the main lines of explanation that have been advanced for the inequality in scientific productivity. No resolutions of the issues are to be expected here. The far more modest aim is to take stock of what has been learned, and what problems remain to be addressed before one can arrive at an explanatory structure whose interpretation and policy implications are reasonably unambiguous.

3.1 The "Sacred Spark" Hypothesis and the Ability Distribution Problem

One of the first explanatory proposals to be taken up was whether the extreme left-skewed distributions in scientific publications, and in measures of quality-adjusted output based on citation frequencies, could simply be reflecting substantial *predetermined* differences among scientists in their ability and motivation to do creative scientific research. If, as proponents of the so-called "sacred spark hypothesis" contend, there are some rare individuals of great talent, a central part of the resource allocation task for science policy makers might be reduced to designing mechanisms to find the upper quartile of individuals who will do 75 percent of the useful work, without wasting too much

²⁵ This rather casual suggestion, made by Cole and Cole (1973: p. 234) in the context of discussing the extreme concentration of productivity measured by publication- and citation-counts, quite understandably was dismissed as "egregious" by those who read it as a potential policy recommendation.

funding on the "sparkless" mass. Three problems are encountered in attempting to dispose of the matter so simply:

(1) The first problem is that it is not too plausible that among professional scientists abilities should be distributed as unequally as are publications. Consider that these people are not a random draw from the population at large; rather, they constitute a pre-selected educational elite who have undergone a long and consistent socialization, motivating them to engage in original research and to emulate their mentors by racing against their colleagues for the prize of "priority".

(2) Further, in studies undertaken in the early 1960s (see Bayer and Folger 1966, Taylor and Barron 1963), measures of intellectual ability or personality were found nearly always to show very low correlation with scientific productivity. Efforts to save the "sacred spark hypothesis" gave rise to models of scientific productivity as requiring many distinct mental factors, which, if they interacted multiplicatively to determine productivity (as in William Shockley's 1957 model) could generate an output distribution more skewed than any of the determinants and weakly correlated with any one of them.

(3) Yet, a third problem remains, as has already been intimated. A pre-determined distribution of abilities can perhaps account for the static, cross-sectional pattern of productivity differences in the population, and in cumulated lifetime productivities. But, it won't explain the fact that the degree of inequality -- and, more generally, scale-invariant measures of the dispersion in productivities -- has been found to increase systematically over the life of cohorts of scientists in many different fields. Some of the early studies showed this by constructing synthetic cohorts from survey data on publications and citations, using age since Ph.D. to form the cohorts: reference has already been made to the almost linear increase in the Gini coefficients computed for publications (and for citations in parentheses) for mathematicians, chemists, and physicists, by Allison and Stewart (1974), and to the confirmation of this phenomenon's appearance in true cohort data, by Allison, Long and Krauze (1982).

3.2 Merton's "Matthew Effect" and Recognition in Science

Robert K. Merton's classic (1968) paper on the role of collegial reputations in the reward system characteristic of communities of scientists, called attention to one aspect of the process that resulted in very pronounced status differentiation. This was

"the accruing of greater increments of recognition for particular scientific contributions to scientists of considerable repute and the withholding of such recognition from scientists who have not yet made their mark."

Merton dubbed this "the Matthew Effect", by allusion to the passages of New Testament gospel according to St. Matthew that read

"For unto everyone that hath shall be given, and he shall have abundance; but from him that hath not shall be taken away even that which he hath." (Matthew 13:12 and 25:29)

In his original formulation of the Matthew Effect, Merton emphasized the disproportionately great credit for their contributions (and even for the contributions of others associated with them in publications) that was bestowed upon scientists who had attained some eminence. He suggested that busy scientists, not being able to read everything that was being published in their field, might be rationally allocating their time by paying special attention to studying the works of colleagues on the basis of the preceding reputations of the papers' authors. Another, certainly not mutually exclusive explanation for "name salience" and "halo effects" in citation practices, and one that economists may find no less convivial than Merton's, would put more emphasis on the role of "signalling" by third parties. The latter, when deciding who or what to cite, may be concerned to demonstrate that they are conversant with the reputational ranking of people in a specific area of science. The demonstration allows such considerations to color their choices of whose work should be cited as an authority on a particular point, or whose proposals or journal submissions should be recommended in a peer review process. Judges and referees in this model have "preferences" for research proposals, or reports, based on their inherent attributes; but, they also care about the approbation or disapprobium received from their colleagues regarding their evaluation of the "track record" and reputation of the authors of the proposals or papers under review. Consequently, certain citation conventions may arise that have a partly fortuitous history, but which, once they become established, command conformity nonetheless.²⁶

3.3 Cumulative Advantage--Recognition, Reputation, Reinforcement and Rationality

Merton and his students -- Harriet Zuckerman (1972), and Jonathan and Stephen Cole (1973) -- quickly generalized the original idea behind the Matthew Effect by proposing that self-reinforcing processes affected productivity as well as recognition in science. In this broader form, the Matthew effect became synonymous with the "cumulative advantage" hypothesis. The sociologists' argument, according to a recent statement of it (Allison, Long and Krauze (1982)) goes like this:

"Scientists who are rich in recognition find it easier to get the resources that facilitate research: grants, free time, laboratories, stimulating colleagues, capable students, etc. They are also encouraged by their colleagues to continue to invest time and energy in research.... As a consequence their research productivity is likely to increase -- or at least be maintained at high levels."

In contrast, scientists who receive little recognition for their research efforts get little in the way of resources and encouragement, and, it seems plausible enough to suppose that those who become discouraged will have still lower future productivity and diminished chances for recognition.

Notice something about the implicit psychological assumptions in these formulations: the sociological literature is quite explicit in viewing the positive feedback loop as passing from recognition to productivity via augmented effort. Moreover, it sees it as utterly plausible that raising

²⁶ This sort of explanatory structure for the formation and perpetuation of particular citation conventions, where numerous valid citations might be offered, resembles the coordination equilibrium theory of political consensuses offered by Kuran (1987).

one's esteem in the eyes of colleagues will translate into wanting to meet their raised expectations of exceptional achievements, thus motivating greater efforts to fulfill expectations -- even to the point of stress. I think we have all observed that a remarkably high proportion of our fellow scientists who have enjoyed reputational success complain frequently of being stressed by "overwork", and perhaps it's true that they are trying hard not to disappoint us.

Dyed-in-the-wool microeconomists, however, will be casting about at this moment for some less altruistic-sounding explanation of the phenomenon just described. Here is one that turns on considering the option value of maintaining a high reputational standing, however you happened to come by it. Suppose that an individual researcher's reputation derived in some measure from prior successes that owed much to exceptional "good luck" (perhaps the timing of their paper happened to be just right, because an audience had been prepared by others, or some transient external event had made the scientific contribution in question appear terribly "relevant" to real world problems). But while you may know this, you can hope that others will impute your fame to your innate abilities, which they can't observe directly in any case. And even if others can still recall the actual circumstances of your good fortune, you'll want to remind them of Pasteur's homily about the role of chance in science -- "Fortune favors the prepared mind", and all that. In order to collect the rents that accrue to those who are believed to be especially talented, the merely lucky scientists can be induced to work much harder than they expected to work, and harder than they would have worked had they not suffered the extraordinary luck to be mistaken for the extraordinarily gifted. Surely the folk caught on such treadmills are to be more pitied than envied; they are rational prisoners of their unanticipated success.

The other side of the irrational or rational compulsion to keep going is the irrational or rational decision to exit from a competition in which one believes oneself to be at a comparative disadvantage. A low success rate in having one's work published, recognized, and accepted among one's scientific peers is likely to lead to reduced future effort for any one of several reasons. Some individuals receiving this sort of negative reinforcement will, rightly or wrongly lose confidence in their earlier favorable appraisal of their own capabilities as research scientists, and consider shifting to alternative career paths. Others may remain convinced of their talents, and of the validity of their ideas, but despair of overcoming the handicaps under which a poor early showing in the competition to recognition and status will have placed them. Theories of "social learning" maintain that individual's reactions to events, or stimuli (either positive or negative), are influenced both by their prior experience with the same stimuli -- because cognitive processes influence the perception and retention of events of that type, and by their anticipations of the future consequences of particular responses.²⁷

²⁷ See discussion by J. R. Cole and B. Singer, "A Theory of Limited Differences: Explaining the Productivity Puzzle in Science," p. 190-191 in Zuckerman, Cole and Bruer (1991), where Albert Bandura's (1977) account of social learning theory is quoted as follows:

It should therefore be recognized that one process contributing to the intra-cohort growth of inequality in productivity is the observed tendency for an increasing fraction of a cohort's members to exit completely from research activity (and the publication of scientific papers) as they accumulate adverse experiences in that pursuit. By contrast, those who remain active are likely to keep producing at a virtually undiminished pace until they are quite senior in age. Where populations are heterogeneous, however, inferences about individual behavior from population averages can be quite misleading. To the extent that attrition from the ranks of the active is selective, and that those having a lower "intensity" (or per period expected rate) of publication are most likely to be among the first to withdraw, the empirical mean publication rate for the "survivors" would be pushed upwards from one year to the next -- even though each researcher's individual productivity intensity remained unchanged. This implies that indexes of concentration or inequality in current productivity computed for the still-active portion of the cohort exclusively (omitting those who have definitively ceased) might well exhibit stability and even trend downwards, whereas the growing gap between the mean productivity rates of the active and the inactive was contributing to the growth of the variance of productivity among the members of the cohort as a whole. Unfortunately, no detailed statistical analyses appear to have been made of extended longitudinal panel data on individual scientists' publication careers, from which it would be possible to determine the quantitative importance of the hypothesized selective attrition process.

4. Stochastic Models of Cumulative Advantage Processes

The generalization of the "Matthew effect" suggested a rationale for the success found in using certain statistical distributions to describe scientific publications and citations data. The early work of Derek de Solla Price (1976) in this connection holds special interest because it drew attention to the relevance and mathematical tractability of Polya Urn schemes as stochastic representations of positive feedback processes -- almost a decade before this class of models began to be thoroughly investigated and exploited as interesting and tractable representations of self-reinforcing economic processes.

4.1 Price's Characterization of the Matthew Effect as an Urn Process:

Price (1976) started by taking up an idea due to Polya, namely, of employing a model of sampling from the contents of an urn with "over-replacement" as a statistical description of "after-effects", or hysteresis.²⁸ Price first characterized Merton's Matthew effect as a "double-edged"

"In the social learning theory view, people are neither driven by inner forces nor buffeted by environmental stimuli. Rather, psychological function is explained in terms of a continuous reciprocal interaction of personal and environmental determinants. Within this approach, symbolic, vicarious, and self-regulating processes assume a prominent role.

²⁸ On Polya urn processes, see Feller (1968, vol.I: p. 119).

process that rewarded successes and punished failures. The model likens the individual to an urn, and supposes that fate has in storage for each scientist an urn containing balls of two colors -- let them be red and white -- mixed in proportions representing the probabilities of success (red) and failure (white) to generate a type of event -- say, a scientific publication -- on any given draw. If the urn was sampled with exact replacement, the composition of its contents would remain fixed and the individual's chances of success and failure would not vary. But, if it was sampled with *over-replacement*, so that whichever color ball was drawn would be returned along with some additional number of balls of the same color, the chances on successive draws would change as an after-effect of the previous history in a direction that would *reinforce* the cumulative experience of the past. A given population of identical individuals who started out with the same chances of success would soon become heterogeneous in their respective probabilities of success. If more than one ball of the same color is added back with one drawn, the situation will be as shown in Figure 1: depending upon the "ground state" (the initial identical composition of the urns in the population), per trial probabilities of success for some part of the population will be tending to fall towards zero, whereas, for the rest, their probabilities of success will tend to move towards unity.²⁹ Due to a theorem of Polya's it is known that if exactly one ball of the corresponding color is added back on each draw, in the limit any probability of success in the range from zero to one can become established in a given urn, and the probability to which the urn actually converges will be determined entirely by luck (early) in the history of successive draws.

The whole population of urns would then represent a mixture of Poisson processes, each individual having a mean probability of publication success. This was what Derek Price wanted to show, because it is a well-known property that the cumulative number of successes generated by the resulting mixture of Poisson's follows the negative binomial distribution. And the latter distribution had been found quite satisfactory as an alternative to Lotka's Law in describing scientific publications.

Price (1976) then went on to modify the classic Polya urn model, on the argument that it was not evident how failures to publish, being "non-events," would be marked and punished. Therefore, he characterized a one-sided process of cumulative advantage in which nothing changed after a publication failure, but the chance of future success was improved by each success. From this modification he derived a new family of distribution functions describing a population of urns, which he called "Cumulative Advantage Distributions." These had the virtue of approaching Lotka's Law, and Price's Law, as a special, limiting case.³⁰ where $f(n)$ is the number of scientists in a

²⁹ Economists recently have been introduced to the properties of Polya urn processes and their implications by Arthur (1989, 1990), and David (1985, 1987), building on the work of Arthur, Kaniovski and Ermoliev (1985).

³⁰ Price's (1976) general cumulative advantage distribution for the population of urns is written as:

population of size P who publish n papers; C is a constant of proportionality, m is a parameter, and $B(n, m + 2)$ is the Beta Function, as defined in note 22 (above). When $m = 0$ the function approaches results obtained with the Lotka (1926) inverse square rule as n increases. Price suggested that as a rule about two-thirds of scientists were only transiently present "on the research front", emitting an initial article on arrival but publishing nothing more. This would correspond to the case where the parameter $m = 1$ in the cumulative advantage distribution, in which the one third who would go on publishing would have an even chance of a second publication, and their success probability would thereafter approach unity asymptotically.

As elegant as is the derivation of the Cumulative Advantage Distribution by Price (1976), it remains fair to question his objection to the classic urn model's characterization of self-reinforcement as a double-edged process, in which failure to publish would impart a disadvantage, tending to lower the probability of successful publication in the following period. Perhaps in past eras, when the pace of research advances in many scientific fields was slower, and when individual researchers and groups of scientists were not obliged to compete with one another for scarce research funding from granting agencies in the fashion that has become the norm in academic science in the U.S., Britain, and elsewhere during the past several decades, it would be reasonable to characterize a failure to publish as "non-event" from which no disabling after-effects ensued. Under such conditions it could be supposed, as Merton (1973) asserted, that it was the competition among scientists for priority, acceptance and recognition of their ideas that the reward system elicited, and that animated the pursuit of knowledge.³¹ Nowadays, however, as Cole and Singer (1991:p.203-204) acknowledge, the competition of ideas has become augmented by and thoroughly intertwined with competition for funds, most visibly in fields where the minimum effective sizes of research teams and the costs of the physical apparatus have grown very large. The centrality of this two-sided form of competition in the lives of scientists, and especially among the stratum of prolific, highly productive researches located at the modern scientific institutions, is attested to by biographical reports and sociological studies ranging from large-scale surveys to studies of individual laboratories. Scientists' productivity is directly linked to keeping the laboratory running and thereby retaining key, highly trained personnel. The common perception is that continuity of operations would be seriously jeopardized by evident

$$f(n) = C P^m B(n, m + 2),$$

³¹ But, as David (1991) emphasizes, the quest for patronage, and the material support and time to undertake fundamental inquiries was not an insignificant aspect of the social organization of science even in the age of Galileo.

failures to deliver results projected in previous grant proposals, or other untoward events that would cause rejection of the next grant proposals submitted to the peer-review process. In short, an explicit delineation of critical points of in the competition among university-based researchers for scarce material resources is warranted by the direction in which the greatly expanded basic science enterprise has evolved, and this may offer some corrective to the tendency in sociological formulations of the cumulative advantage hypothesis to portray a career in science as a sort of sequential obstacle course, on which entrants who had been successful in clambering over earlier hurdles were better positioned to surmount the subsequent ones.

4.2 Search Processes Under Conditions of Positive Feedback

What we have been looking at in the foregoing are very stylized representations of a process of cumulative causation where small differences in initial achievements become magnified over time. Stochastic processes in which positive feedbacks dominate, are found to be at the heart of a more general analysis of "path dependent" dynamical systems of economic resource allocation, and other social processes. They seem to be especially relevant in the channelling of technological change and institutional evolution, as well as in the dynamics of self-fulfilling expectations and coordination games, in which a multiplicity of stable equilibria or "attractors" can be shown to exist. With such systems, quite small perturbations or shocks at early points on the path can be sufficient to play the role of "selection mechanisms", pushing the system into one or another basin of attraction.³²

In the present context what is involved are the positive feedbacks between achievements, and access to research facilities that raise the chances of scoring further achievements); it is a self-confirming process through which are constructed scientific reputations, the fundamental currency of the reward system of the "open science" research community. To the extent that opportunities to gain expertise are allocated on the basis of achieving early distinction, and there is an element of luck in the latter process, some individuals can gain a cumulative advantage that is not proportionate to their initial, inherent ability. This tendency will be amplified if the allocation of resources for research among investigators is influenced by consideration of their "pedigree and provenance" (who trained them, and where have they worked), and of their scientific "track records." The person who lost out in some early rounds will find it increasingly difficult, if not impossible to prove (to others, or even to himself) that he or she was just as good as, if not better than, the person who has now become a big research star.

And in one sense -- namely the comparison of researchers in regard to their talents *ab initio*, at the outset of their respective careers -- the losers may be entirely justified in their doubts about the reigning stars of their discipline. In fact, Robin Cowan (1991) recently has given a rigorous demonstration for this, in an analysis of multiple-armed bandit games under conditions of positive

³² See e.g., David (1975:Intro.; Ch. 1), (1985), (1988); Arthur (1983, 1989, 1990).

reinforcement of the performance of the "arms" that get played. So, let us for a moment imagine the grant awards process of the U.S. National Science Foundation and the National Institutes of Health, or of the Research Advisory Councils in Great Britain -- or, more abstractly the research resource allocation process -- in the role of the optimizing "player", and the different scientists (or their laboratories and research institutes) as the many "arms" of the slot machine that can be played. There are then three features of the general structure that make the results of Cowan's (1991) formal model applicable to the present context:

(1) complete information is not available to the "player" about the success (or payoff) probabilities that initially characterize the "arms" of the slot machine;

(2) choices are made sequentially, on the basis of previous observable outcomes, to support one rather than another among a set of rival researchers (or parallel research teams);

(3) the results of a team's efforts at each stage of the sequence are in some part a matter of chance but the odds become more favorable as the group is given the opportunity to acquire research experience.

Thus, we could make use of the theorem that says that the structure of the optimal *ex ante* strategy for the funding agency will be to bet on alternative competitors for a number of rounds before picking the one that emerges with the best calculable chances of future success. At that point a "star" will have been born, and the agency would want to concentrate support upon that individual researcher or group leader thereafter. For, under the conditions stated, the accumulation of experience made possible by past funding will have rendered him or her the currently best bet.

But, a second applicable proposition about the results in these circumstances is this: implementation of this optimal strategy inevitably will lead the manager of research to make "stars" of some comparatively dull researchers, and to consign others of potentially greater luminosity to comparative oblivion. The reason is simply that a run of pure luck, especially during the early stages of a career, can create a presumption in favor of further backing, and, hence greater opportunity to further improve the lucky individual's research capabilities. But, as luck would have it, those capabilities when realized may be considerably more limited than those of other, initially less fortunate scientists.

5. Conclusion--An Agenda for Further Work

Entertaining, and possibly illuminating as all this may be, these simple uses of the Polya urn model -- in which everyone starts off on the same footing and becomes progressively more differentiated -- will not take us as far as we need to go. They may be able to generate properly skewed productivity distributions for populations research scientists observed over a finite life, but they cannot explain an increasing degree of inequality over the life of a cohort. It is intuitively reasonable, and has been demonstrated formally by Allison, Long and Krauze (1982), for example, that one needs to introduce some heterogeneities in the rates at which individual's undergo positive

(and possibly negative) reinforcement of their productive propensities, in order to get scale invariant measures of the degree of inequality to increase over the life of a cohort of researchers.

It is not necessary, however, to suppose that the only relevant source of heterogeneities has to be the difference in inherent abilities, as the sacred spark theory insists. Noise in the initial perception of abilities on the part of agents who allocate resources to researchers may also be part of the story. But there are still other sources of relevant heterogeneity in a population of researchers who are not, after all, the non-volitional arms of a giant NSF slot machine. The researchers, like the funding agency, are having to make choices about whether to make more efforts or to give up seeking to build a publication reputation and spend time writing proposals for funding; and they, too, are proceeding under uncertainty while "learning" (in the sense of forming expectations) about their future prospects of success in such endeavors. On the basis of those revised perceptions and the rewards anticipated for alternate career paths, they will be making decisions as to whether it is worth persevering in the competition for recognition of their ideas, and for the resources to continue their investigation.

The dynamic system of overlapping generations of researchers who have complementary and competitive relationships with other researchers cannot be adequately analyzed using purely reductionist techniques. Its aggregate performance cannot be predicted (except in very special circumstances), and is not explained by adding up the productivities of a "representative" individual researcher over the course of a complete scientific career. Further empirical work, of a character more fine-grained than could be undertaken by the pioneer generation of quantitative sociologists of science, will be required to articulate the many distinctive subprocesses of the competition for research resources, the production of findings, preparation of papers, competition for space in reputable professional journals, and to identify important positive and negative feedback mechanisms, as well as sources of "exogenous" heterogeneities in the population of scientists and the groups that they form. But evaluation of the importance of the various subprocesses to the functioning of the whole requires putting the pieces together and some stylized models that will permit undertaking experiments with stochastic simulation techniques to discover how they interact, and how sensitive is the behavior of the entire structure to alterations in the specific features of its components.³³

Such a program of research seems essential if we are to significantly advance our understanding of the workings of the existing resource allocation regimes in science as a whole, and its various sub-disciplines. Without a firmer grasp of the quantitative behavior of these systems, we are likely either to succumb to the cautious conservatism of leaving well enough alone (however wasteful it may be), or undertaking disruptive and ill-considered reforms in an effort to exert control over a social and economic structure that we have failed to properly understand but, nonetheless, perceive to be terribly important to our future welfare.

³³ Such a program of research has been initiated by the author, in collaboration with Roland Maude-Griffen; specifics of the approach and some preliminary findings were reported in David (1992b).

References

Aitchison, J. and J. A. C. Brown (1957), The Lognormal Distribution, With Special Reference to Its Uses in Economics, Cambridge: Cambridge University Press.

Allison, Paul D., J. Scott Long and Tad K. Krauze (1982), "Cumulative Advantage and Inequality in Science," American Sociological Review, 47(October): pp.615-625.

Allison, P. D., D. J. deS. Price, B.C. Griffith, M. J. Moravcsik, and J. A. Stewart (1976), "Lotkas Law" a Problem in Its Interpretation and Application," Social Studies of Science, 6,: pp. 269-276.

Allison, Paul D. and John A. Stewart (1974), "Productivity Differences Among Scientists: Evidence for Accumulative Advantage," American Sociological Review, 39 (August): pp.596-606.

Andrews, F. M., ed. (1979), Scientific Productivity: the Effectiveness of Research Groups in Six Countries, Cambridge: Cambridge University Press.

Arrow, Kenneth J. (1962), "Economic Welfare and the allocation of Resources for Inventions," in R. R. Nelson, ed., The Rate and Direction of Inventive Activity: Economic and Social Factors. National Bureau of Economic Research-Universities Conference. Princeton, N.J.:Princeton University Press.

Arthur, W. Brian (1983), "On Competing Technologies and Historical Small Events: The dynamics of Choice Under Increasing Returns," International Institute for Applied Systems Analysis, Working Paper WP-83-90.

Arthur, W. Brian (1989), "Competing Technologies and Lock-in by Historical Small Events", Economic Journal, 99(394):pp.116-131.

Arthur, W. Brian (1990), "Positive Feedback in Economics," Scientific American, September.

Arthur, W. B., Yu. M. Ermoliev, and Yu. M. Kaniovski (1987), "Path Dependent Processes and the Emergence of Macro-Structure," European Journal of Operations Research, 30.

Auerbach, Felix (1910), Geshichtstafeln der Physik, Leipzig: J.A. Barth.

Bandura, Albert (1977), Social Learning Theory, Englewood Cliffs, N.J.: Prentice Hall.

Bayer, Alan E. and John Folger (1966), "Some Correlates of a Citation Measure of Productivity in Science," Sociology of Education, 39 (Fall):pp. 381-389.

Bensman, Stephen J.(1982), Bibliometric Laws and Library Usage as Social Phenomena," Library Research,4 pp. 279-312.

Broadus, R. N. (1977), "The Application of Citation Analyses to Library Collection Building," Advances in Librarianship, 7: pp.299-335.

Cole, Jonathan R. and Stephen Cole (1973), Social Stratification in Science, Chicago: The University of Chicago Press.

Cole, Jonathan R. and Burton Singer (1991), "A Theory of Limited Differences: Explaining the Productivity Puzzle in Science," Ch. 13 in The Outer Circle: Women in the Scientific Community, H. Zuckerman, J. R. Cole and J. T. Bruer, eds., New York: W. W. Norton.

Cole, Jonathan R., L. Rubin, and Stephen Cole (1978), Peer Review in the National Science Foundation: Phase One of a Study, Washington, D.C.:National Academy of Sciences.

Cole, Jonathan R. and Harriet Zuckerman (1984), "The Productivity Puzzle: Persistence and Change in Patterns of Publication of Men and Women Scientists," in Advances in Motivation and Achievement, Vol.2: pp. 217-258.

Cowan, Robin (1991), "Tortoises and Hares: Choice Among Technologies of Unknown Merit," Economic Journal, 101 (July): pp. 801-814.

Dasgupta, Partha and Paul A. David (1987), "Information Disclosure and the Economics of Science and Technology," in G. Feiwel, ed., Kenneth Arrow and the Ascent of Economic Theory, New York: MacMillan.

Dasgupta, Partha and Paul A. David (1988), "Priority, Secrecy, Patents, and the Economic Organization of Science and Technology," CEPR Publication No.127, Stanford University, March. Forthcoming (revised) in Science in Context.

Dasgupta, Partha and Paul A. David (1992), "Toward a New Economics of Science," CEPR Publication No. 320," Stanford University. October. Revised and forthcoming in Research Policy, 1994.

David, Paul A. (1975), Technical Choice, Innovation and Economic Growth: Essays on American and British Experience in the Nineteenth Century, Cambridge, Eng.: Cambridge University Press.

David, Paul A. (1985), "Clio and the Economics of QWERTY", American Economic Review, 75(2), May:332-37.

David, Paul A. (1987), "Some New Standards for the Economics of Standardization in the Information Age," Ch.8 in Economic Policy and Technology Performance, P. Dasgupta and P. Stoneman, eds., Cambridge: Cambridge University Press.

David, Paul A. (1988), "Path-Dependence: Putting the Past into the Future of Economics," Institute for Mathematical Studies in the Social Sciences Technical Report 533, Stanford University. November.

David, Paul A. (1991), "Reputation and Agency in the Historical Emergence of the Institutions of 'Open Science'", Center for Economic Policy Research Publication 261. Stanford University, Stanford, California. May.

David, Paul A. (1992a), "Path-Dependence and Predictability in Dynamic Systems with Local Network Externalities: A Paradigm for Historical Economics," in Technology and the Wealth of Nations, D. Foray and C. Freeman, eds., London: Pinter Publishing.

David, Paul A. (1992b), "Positive Feedback, 'Matthew Effects', and Productivity Puzzles in Basic Science," The 1992 Harry Johnson Lecture, presented to the London Meeting of the Royal Economic Society. March 29-1 April.

David, Paul A., David R. Mowery, and W. Edward Steinmueller (1992), "Analyzing the Economic Payoffs from Basic Research," Economics of Innovation and New Technology, 2(4): pp. 73-90.

Durlauf, Steven N. (1991), "Non-Ergodic Economic Growth and Path Dependence in Aggregate Output Fluctuations," American Economic Review, 81 (May).

Gaston, Jerry (1978), The Reward System in British and American Science, New York: Wiley and Sons.

Feller, William (1968), An Introduction to Probability Theory and Its Applications. 3rd Edition. New York: John Wiley & Sons, Vol.2 (Vol.1, 1966).

Jaffe, Adam (1986) "Technological Opportunity and Spillovers of R&D: Evidence from Farms' Patents, Profits and Market Value," American Economic Review, 76(5): pp. 984-1001.

Jaffe, Adam (1989), "Real Effects of Academic Research," American Economic Review, 79(5): pp.957-70.

Jaffe, Adam, Manuel Trajtenberg and Rebecca Henderson (1992), "Geographic Localization of Knowledge Spillovers, as Evidenced by Patent Citations," Quarterly Journal of Economics.

Kuran, Timur (1987), "Preference Falsification, Policy Continuity and Collective Conservatism," Economic Journal, 97 (September): pp. 642-655.

Lotka, Alfred J. (1926), "The Frequency Distribution of Scientific Productivity," Journal of the Washington Academy of Sciences, 16, No.12: pp.317-323.

Mansfield, Edwin (1991), "Social Rate of Return from Academic Research," University of Pennsylvania, Department of Economics Working Paper.

Merton, Robert K. (1968), "The Matthew Effect in Science," Science, 159(3810), January 5: pp. 56-63.

Merton, Robert K. (1973), The Sociology of Science: Theoretical and Empirical Investigations, N. W. Storer, ed., Chicago: Chicago University Press.

Merton, Robert K. (1988), "The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property," Isis, 79: pp. 606-623.

Mowery, David R.(1983), "Economic Theory and Government Technology Policy," Policy Sciences, 16.

Mowery, David R. and Nathan Rosenberg (1989), Technology and the Pursuit of Economic Growth, Cambridge: Cambridge University Press.

Narin, Francis (1976), Evaluative Bibliometrics, Cherry Hill, NJ: Computer Horizons.

National Science Foundation (1985), Science and Technology Data Book, Washington, D.C.: Government Printing Office.

Nelson, Richard R. (1959), "The Simple Economics of Basic Scientific Research," Journal of Political Economy, 67:pp.297-306.

OECD (1985), Main Science and Technology Indicators, Paris:O.E.C.D.

Polanyi, Michael (1962), "The Republic of Science: Its Political and Economic Theory," Minerva, 1(1): pp. 54-73.

Price, Derek J. deSolla (1963), Little Science, Big Science, New York:Columbia University Press.

Price, Derek J. deSolla (1976), "A general Theory of Bibliometric and Other Cumulative Advantage Processes," Journal of the American Society for Information Science, 27, Nos. 5/6: pp. 292-306.

Rosenberg, Nathan (1990), "Why Do Companies Do Basic Research With Their Own Money?" Research Policy, 19: pp. 165-174.

Shockley, William (1957), "On the Statistics of Individual variations of Productivity in Research Laboratories," Proceedings of the Institute of Radio Engineers, 45 (March): pp. 279-290.

Simon, Herbert A. (1957), Models of Man: Social and Rational, New York: J. Wiley & Sons.

Taylor, Calvin W. and Frank Barron, eds. (1963), Scientific Creativity: Its Recognition and Development, New York: J. Wiley and Sons.

Thurow, Lester (1975), Generating Inequality, New York: Basic Books.

Trajtenberg, Manuel, Rebecca Henderson and Adam Jaffe (1992), "Ivory Tower Versus Corporate Lab: An Empirical Study of Basic Research and Appropriability," NBER Working Paper No. 4146, National Bureau of Economic Research, Cambridge MA. August.

Wade, Nicholas (1975), "Citation Analysis: a New Tool for Science Administrators," Science, vol.88, No.4187): pp.429-32.

van Raan, A. F. J., ed. (1988), Handbook of Quantitative Studies of Science and Technology, Amsterdam: North Holland Publishing Co.

Williamson, Jeffery G. and Peter D. Lindert (1980), Two Centuries of American Inequality, New York: Cambridge University Press.

Zuckerman, Harriet (1977), Scientific Elite: Nobel Laureates in the United States, Chicago: The University of Chicago Press.

Zuckerman, Harriet and Robert K. Merton (1972), "Age, Aging and Age Structure in Science", in M. W. Riley, M. E. Johnson and A. Foner (eds.), A Sociology of Age Stratification, New York: Russell Sage.

Zuckerman, Harriet, Jonathan R. Cole and John T. Bruer, eds. (1991), The Outer Circle: Women in Science, New York: W. W. Norton.