PRODUCTIVITY DIFFERENCES AMONG SCIENTISTS: EVIDENCE FOR ACCUMULATIVE ADVANTAGE*

PAUL D. ALLISON AND JOHN A. STEWART
University of Wisconsin – Madison

American Sociological Review 1974, Vol. 39 (August): 596-606

The highly skewed distributions of productivity among scientists can be partly explained by a process of accumulative advantage. Because of feedback through recognition and resources, highly productive scientists maintain or increase their productivity, while scientists who produce very little produce even less later on. A major implication of accumulative advantage is that the distribution of productivity becomes increasingly unequal as a cohort of scientists ages. Cross-sectional survey data support this hypothesis for chemists, physicists, and mathematicians, who show strong linear increases in inequality with increasing career age. This increase is highly associated with a changing distribution of time spent on research. Another implication of accumulative advantage is also corroborated: the association among productivity, resources and esteem increases as career age increases.

Scientists differ enormously in the number of papers they publish. Lotka (1926) found that for a sample of physics journal papers published in the nineteenth century the frequency distribution of scientists by number of papers published could be approximately described by the function

$$F(n) = k/n^2$$

where n is the number of papers, $F(n)$ is the number of scientists publishing n papers and k is a constant. This distribution is highly skewed so that in a typical case less than six percent of publishing scientists produce about fifty per cent of all papers (Price, 1963). Lotka's "inverse square law" of scientific productivity has since been shown to fit data drawn from several widely varying time periods and disciplines (Leavens, 1953; Davis, 1941; Price, 1963; Britton, 1964). More recently it has also been suggested that if a scientist's productivity is measured by the number of citations to his work, the distribution is even more highly skewed (Hagstrom, 1968; J. Cole, 1970; Cole and Cole, 1972).

Explanations of these extreme differences in scientists' performances are generally of two types (Cole and Cole, 1973). The "sacred spark" hypothesis says simply that there are substantial, predetermined differences among scientists in their ability and motivation to do creative scientific research. More interesting sociologically is the hypothesis of "accumulative advantage" – that because of a variety of social mechanisms, productive scientists are likely to be even more productive in the future, while scientists who produce little original work are likely to decline further in their productivity. Although these hypotheses are not logically incompatible, it is nevertheless important to gauge their relative contributions to the structure of inequality in the scientific community. Later we will present evidence which provides a preliminary answer to this question. First, we will attempt to formulate these ideas more precisely in order to discuss their problems and implications.

The Sacred Spark

Few would deny that scientists differ greatly in their productive capacities, and that these differences are, to a large degree, determined before their careers begin. Yet, the sacred spark hypothesis in its simplest form fails to answer two critical questions, one theoretical and one empirical. First, is it plausible that scientific ability is distributed as unequally as publications? Since scientists are a

*We are indebted to Warren Hagstrom for providing the data analyzed here, and for advice and encouragement throughout our work on this project. For helpful comments on an earlier version of this paper, we thank Jonathan Cole, Lowell Hargens, Shiela Klatzky, Robert Merton, Aage Sørensen, Seymour Spilerman, H.H. Winsborough and Harriet Zuckerman. This research was supported in part by grants to Hagstrom from the National Science Foundation.

relatively small, rigorously selected elite who have undergone long and consistent socialization, we might expect them to be quite homogeneous on dimensions relevant to scientific activity. Yet, scientific productivity is distributed more unevenly than income among individuals (Britton, 1964). Second, why do measures of intellectual ability or personality nearly always show very low correlations with productivity (Bayer and Folger, 1966; Taylor and Barron, 1963)? It could be argued that adequate measures of scientific ability have not yet been developed, but this remains to be demonstrated.

Shockley (1957) has proposed a mathematical model which seems to answer both questions. He suggests that productivity is the result of many "mental factors," such as the ability to find important problems, technical ability and persistence. The crucial feature of this model is that the mental factors determine productivity *multiplicatively* rather than additively. This implies that the resulting distribution of productivity will be more skewed than any of its determinants, and will tend toward lognormality as the number of determinants increases (Aitchison and Brown, 1957). As an extension of Shockley's model, we suggest that as tasks become less routine, more characteristics of the individual become relevant to the performance of task, and hence the distribution of task performance becomes increasingly skewed. This would account both for the highly skewed distribution of productivity in science, a relatively non-routine activity, and the low correlation of productivity with any one of its determinants.

Accumulative Advantage

Merton (1968) has identified what he calls the "Matthew effect" in science: "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." Essentially he argues that the overloading of the scientific communication system leads scientists to choose their reading matter on the basis of an author's preceding reputation, often further enhancing that reputation.

The accumulative advantage hypothesis, as discussed by the Coles (1973), is apparently a generalization of the Matthew effect to include productivity as well as recognition. The process can be viewed as consisting of two feedback loops in which recognition and resources are intervening variables. First, scientists who have been recognized as having made significant advances will be motivated to maintain or increase their recognition by additional publications, and will be influenced by their colleagues' expectations that they repeat or exceed those achievements. Second, beyond these direct effects, recognition usually implies increased access to resources which facilitate research: money, time, competent assistants, stimulating colleagues, easy access to useful information, etc. In addition to their instrumental value, such resources will be directly rewarding as concrete indicators of scientists' esteem and, like recognition itself, will have positive effects on productivity. In contrast to this picture of spiraling success, the scientist who publishes little or whose work is not recognized is likely to become discouraged with research, especially when he cannot get the resources to carry it out.[1]

What evidence is there to support this hypothesis? Cross-sectional studies consistently show that productivity is strongly associated with recognition (Cole and Cole, 1967) and with key resources (Hagstrom, 1968). This establishes at least the possibility of the reciprocal causal process postulated here. More to the point, the Coles (1967) traced the early publication history of 120 physicists, and found that those whose early papers were often cited continued to be highly productive, while those who received few citations declined significantly in their output.

If there is indeed a process of accumulative advantage in science, several questions remain to be answered. Already noted is the need to determine the importance of accumulative advantage relative to preexisting differences. Second, does the process operate throughout scientists' careers, or does it only affect the initial years, losing importance as research activity becomes stabilized? Finally, of great theoretical interest is the question of whether the process is the outcome of efficiency or inefficiency in the communication and reward system of science. Merton seems to say that the Matthew effect results from an inefficient system in which scientists get more or less

[1] For an extensive discussion of the career contingencies resulting in accumulative advantage, see Cole and Cole (1973: Chapter 9).

recognition than their work really deserves. The Coles (1973), on the other hand, argue both that accumulative advantage may result from giving each scientist his due, and that the resulting social inequities may be highly functional for scientific progress.

Mathematical Formulations

The sacred spark and accumulative advantage hypotheses can be seen as special cases of the principles of heterogeneity and reinforcement discussed in the literature of stochastic models. Heterogeneity models assert that individuals have different propensities to perform certain actions. Reinforcement models assert that an individual's propensity to perform a certain action changes systematically each time he performs that action. A mathematical formulation of accumulative advantage is possible, and would enable us to deduce empirical implications of the hypothesis. Simon (1957), for example, has proposed a stochastic model to explain scientific publication, and Bartholomew (1967, p. 128) discusses a leaving-process model which can be easily applied to scientific productivity.[2] Nevertheless, an indefinite number of such models could be constructed, and those simple enough to be tractable usually require unrealistic assumptions. Without opting for a particular model, we note that models involving positive reinforcement generally imply two things: (1) highly skewed distributions, and (2) increasing dispersion with time (Spilerman, 1970). The first implication helps to explain the observed distributions of productivity. The second supports the intuitive notion that accumulative advantage ought to result in increasing inequality, i.e., that small differences become magnified into large ones. Moreover, it suggests a way to test the reinforcement hypothesis since heterogeneity models do not imply changes over time.

Stochastic reinforcement models tell us nothing about the causal process, however, and additional implications can be drawn if we specify an underlying causal model. For example, the reciprocal relationships among the variables discussed earlier can be specified as a set of simultaneous equations, where P(t) is productivity, E(t) is esteem, and R(t) is resources, all at time t.

$$P(t+1) = a_1 + b_1 E(t) + b_2 R(t+1) + e_1$$

$$R(t+1) = a_2 + b_3 E(t) + e_2$$

$$E(t+1) = a_3 + b_4 P(t) + e_3$$

With certain constraints on the b_i coefficients to reflect positive reinforcement, this system implies that the variances of productivity, esteem, and resources will all increase over time, and that the covariances between them will also increase.

In the remainder of this paper, we will attempt to test these corollary hypotheses, focusing primarily on the predicted increase in the dispersion of productivity. Essentially, then, we are testing the reinforcement or accumulative advantage hypothesis, reserving the heterogeneity hypothesis as a residual explanatory principle.

METHODS

Our central hypothesis is that the distribution of productivity among scientists becomes increasingly dispersed with the passage of time. Since persons continually enter and leave the population of scientists, the ideal method would be to measure the variation in productivity for one or more cohorts at several time points during their career history. Lacking longitudinal data, our strategy is to simulate time-series data by dividing a cross-sectional population into several strata by career age. These strata are thus taken to represent a single cohort in its passage through time. Whatever differences are observed among the age strata can be logically decomposed into two sorts of differences: (1) life-course differences, which are the effects of biological and social aging, and (2) cohort differences, which are differences between cohorts at comparable points in their life history. Since our hypotheses only deal with life course differences, we must assume that cohort differences are negligible.[3]

The data come from probability samples of U.S. scientists in university departments offering advanced degrees in biology, mathematics,

[2] These models are treated in greater detail in Allison and Stewart (1973).

[3] A thorough discussion of the complexities of cohort analysis can be found in Riley, et al. (1972: Chapter 2). We have generally followed their terminology, although we sometimes use the term "career-course differences" as equivalent to "life-course differences."

chemistry, and physics.[4] Of an original sample of 2,250 there were 1,947 respondents to mailed questionnaires or telephone interviews. Additional data were gathered from *American Men of Science* (10th and 11th editions) for both respondents and non-respondents. Other data sources for the sample include the Cartter (1966) measurements of the prestige of university graduate departments, and the 1966 *Science Citation Index*. More complete descriptions of this survey can be found in Hagstrom (1967 and 1974).

Two measures of productivity are used – the self-reported number of research publications in the previous five years and the number of citations to all published works of the respondent. Publication productivity was measured as a response to the following question: "How many articles reporting your original research results have you published in the past five years (not counting abstracts)?" To estimate the reliability of the responses, we randomly selected fifty chemists from our sample, and compared their self-reported number of publications with counts from *Chemical Abstracts* for the same five-year period. The correlation was .94, suggesting that the data are adequate for our purposes.

Citation counts are probably a more reliable measure of productivity, but there are at least two reasons to question their validity. First, citations can be interpreted as indicating either the quality of a scientist's work or the recognition received for that work, an important ambiguity since the accumulative advantage hypothesis treats these as distinct variables. Second, although we measure publications for only the five years preceding 1966, citations in 1966 may be to works published at any previous time. Despite these problems, we believe the citation measure is a useful check on the publication measure, especially since the theory predicts increasing dispersion in citations on either interpretation. The difference in time coverage creates minor difficulties which will be discussed as they arise.

To measure the dispersion of the productivity distribution, we use the Gini index which is considered one of the best measures of inequality (Alker and Russett, 1964). Based on the Lorenz curve, the Gini index has a value of zero when all scientists are equally productive, and approaches unity when all the output is produced by a single scientist.[5] The advantages of the Gini index include its clear geometrical interpretation, its sensitivity to all areas of the distribution, and its relationship to a possible underlying distribution – the lognormal. Aitchison and Brown (1957) have shown that Gini index is monotonically related to the variance of the lognormal distribution.

RESULTS

In outline, our analysis takes the following form. We first divide the sample into eight age strata by the number of years since the Ph.D. in order to observe the simple relationship between social aging and the inequality of publications and citations. These results corroborate the hypothesis of increasing inequality. We then consider alternative explanations which might invalidate or restrict the generality of this result, such as interfield heterogeneity, demographic changes, and cohort differences. In general, these alternatives do not appear to threaten the validity of the initial findings. Finally, we attempt to specify the reinforcement process in greater detail by examining additional empirical implications and by a path analysis of smaller age groupings.

Increasing Inequality

The Gini indices for all disciplines and age strata are shown in Table 1. For the total sample, there is a clear and substantial rise in inequality for both productivity measures from the younger to the older strata, strongly supporting the accumulative advantage hypothesis. To estimate the rate of increase in inequality over the career course, we computed the linear regression of the Gini indices on the number of years since the Ph.D.[6]

[4] Equal probability samples were taken in each of the four fields, but the sampling fractions varied slightly among the fields. In some of the analyses to follow, the four samples were combined without weighting to obtain sufficient numbers within each age stratum. Given the similar results across fields, it is unlikely that precise weighting would have altered our conclusions.

[5] The Gini indices reported here have been modified to adjust for maximum and minimum possible values which occur because of the discrete character of scientists, publications and citations (Allison and Stewart, 1973). These adjustments are very slight, and do not affect the substance of our results.

[6] In nearly every case, the youngest stratum is anomalously high. We excluded them from the regression analysis on the grounds that their five-year

Table 1. Publication and Citation Inequality by Years since Ph.D. and Field

Years since Ph.D.[a]	Total Sample	Biology	Mathematics	Chemistry	Physics
	Gini Index for Publications (Citations)				
36	.69 (.82)	.50 (.67)	.76 (.92)	.66 (.81)	.78 (.88)
27	.60 (.78)	.54 (.67)	.66 (.74)	.63 (.82)	.60 (.81)
19	.56 (.75)	.49 (.72)	.64 (.84)	.53 (.71)	.52 (.78)
14	.50 (.74)	.47 (.69)	.46 (.77)	.48 (.74)	.45 (.66)
11	.50 (.70)	.47 (.70)	.56 (.74)	.44 (.59)	.45 (.73)
8	.44 (.69)	.35 (.68)	.45 (.65)	.47 (.60)	.43 (.74)
4	.42 (.69)	.37 (.67)	.40 (.82)	.42 (.58)	.45 (.73)
2	.48 (.76)	.58 (.73)	.42 (.79)	.40 (.64)	.43 (.80)
Overall Gini	.54 (.78)	.48 (.72)	.58 (.83)	.53 (.76)	.51 (.78)
N	1922 (2172)	507 (537)	361 (453)	555 (628)	499 (554)
	Regression of Gini Indices on Years since Ph.D.[b]				
$b \times 10^2$	.835 (.447)	.480 (-.015)	1.113 (.450)	.811 (.835)	1.084 (.501)
a	.387 (.661)	.374 (.687)	.371 (.708)	.379 (.550)	.338 (.674)
r^2	.98 (.98)	.61 (.01)	.88 (.34)	.95 (.82)	.89 (.64)

[a]Mean for all those in each stratum.

[b]Computations exclude youngest stratum.

(Statistics are presented in Table 1 with a graphic illustration for the total sample in Figure 1.) Again considering only the total sample, the very close fit around the regression lines (with r^2's of .98 in both cases) indicates a uniform, linear increase in inequality over the age strata. This suggests that accumulative advantage has important effects throughout scientists' careers, and is by no means limited to the early stages.

Assuming that the increase in inequality is attributable to accumulative advantage, what are its effects relative to preexisting differences? The regression intercept for publications (.39), which may be interpreted as the estimated Gini index at the time of receipt of the Ph.D., is slightly over half that for the oldest age stratum (.69), suggesting that accumulative advantage and preexisting differences contribute equally to publication inequality by the end of scientists' careers. However, since at any given time younger scientists greatly outnumber older ones, the relative contribution of accumulative advantage to the level of productivity inequality in the entire population (.54 in our sample) is considerably less.[7] Of course, as the Coles (1973) argue, accumulative advantage

publication counts were not meaningful, given a mean of only two years since the Ph.D.

[7]The level of inequality of publications for our total sample is considerably less than that found in article counts based on samples from abstracting services. As previously noted, Price (1963) reports that about six percent of the scientists produce fifty percent of the publications; we find that about thirteen percent produce that proportion. Britton (1964) reports a Gini index of .7 for his sample of astronomy papers, whereas our overall Gini is only .54. These differences can be partly explained by the fact that our sample includes only scientists at graduate departments. On the other hand, the differences would be substantially greater if we made our sample comparable to theirs by excluding scientists who published no articles at all.

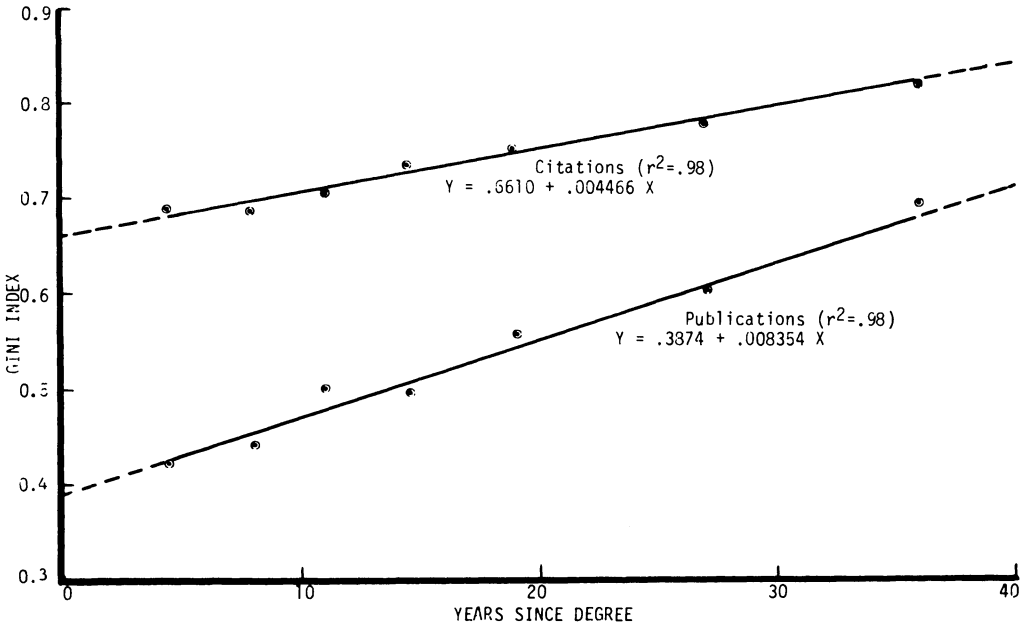

FIGURE 1. Linear Regression of Gini Indices on Years Since Ph.D. for Seven Oldest Strata.

also operates during the educational process, so that inequality at the start of the career cannot be attributed entirely to biological differences.[8]

Note that the Gini indices for citations are much higher than those based on publications for the total sample and for each age group. This result is reasonable if we assume that the number of citations a scientist receives is a product of the number of papers he publishes and the average quality of those papers. Such a multiplicative function will tend to be more highly skewed and have a greater dispersion than either of its two determinants (see earlier discussion of Shockley's model).[9] Although the level of inequality for citations is greater than for publications, the rate of increase over the career course appears to be somewhat lower. This difference is artifactual; the Gini index cannot be interpreted as an equal interval scale over its entire range because it approaches unity asymptotically. Thus, since citation inequality is already high, small increments have more substantive significance than at lower points on the scale.[10]

Alternative Explanations

We have established a strong relationship

[8] It is also unlikely that all the inequality at graduation is due to accumulative advantage. If we extrapolate our regression lines backward in time, we still find substantial inequalities at birth. Although of questionable meaningfulness, this result suggests that either there are important innate differences among scientists, or else accumulative advantage is stronger prior to the start of the scientific career.

[9] Merton's Matthew effect could also account for the greater inequality of citations. One could argue that eminent scientists are cited more often than their work deserves on the basis of intrinsic merits and, conversely, that lesser-known scientists do not receive full credit for their work.

[10] The differences in slopes disappeared when we used the standard deviation of the logarithms of citations or publications as a measure of inequality. Although closely related to the Gini index, this measure does not have an upper limit. Another reason to expect citation inequality to increase more slowly is that citations may be to work published throughout scientists' careers, and thus are less sensitive to recent changes in scientists' publication activity. We suspect that if we only counted citations to works published in the previous five years, the rate of increasing inequality for citations would be *greater* than for publications. Furthermore, a survey of coauthored publications in the *Directory of Graduate Research* (American Chemical Society, 1965) indicated that the more productive professors were more likely to be a junior author than were the less productive professors or the associate or assistant professors. This exercise of noblesse oblige in name ordering patterns (Zuckerman, 1968) will also reduce citation inequality in the older age strata.

between career age and productivity inequality, which we would like to attribute to an underlying reinforcement process. First, however, we will consider alternative explanations which might entirely account for this relationship or restrict its applicability to a smaller segment of the scientific community. Some can be ruled out by controlling for relevant variables or by examining other empirical implications. For some alternatives, lacking essential data, we can only argue their implausibility based on unsystematic observation of scientists' careers.

Differences between fields. Given that the four fields show important differences in the mean publication rate (Hagstrom, 1967), it is plausible that the increasing inequality is due to changing proportions of scientists in the four fields over the age strata. Moreover, it is possible that increasing inequality does not occur in one or more fields. This can easily be checked by disaggregating the sample and performing the same analysis within each field, as we have done in Table 1. The first alternative appears to be refuted – for both citations and publications there is still an increase in inequality for all fields, with the exception of citations among biologists.

The markedly lower regression slopes for biologists have important theoretical implications, although the small N's in each cell indicate a need for caution in interpreting this result. For publications, the rate of increase in inequality for biology is only about half that observed for the other fields; for citations the slope is actually a small negative value. (The r^2's show a corresponding drop.)

We believe these anomalous results can be explained as resulting from inefficiency in the communication and reward systems of biology relative to the other three fields. There is remarkable agreement among both natural and social scientists that biology ranks substantially below chemistry and physics in the degree of consensus over issues of theory and methodology (Lodahl and Gordon, 1972). We can expect this lack of consensus to be manifest in disagreements and "errors" in evaluating individual scientists. Second, biology lacks a core system of journals containing articles of central importance, a fact reflected in the high rate of reprint circulation among biologists (Hagstrom, 1970). Third, the formal academic structure of biology is so fragmented that biologists with similar interests may be located in any of several different departments (Hagstrom, 1967). Surely these factors limit an individual's ability to obtain discipline-wide recognition for his work.

According to Zuckerman and Merton (1972), the Matthew effect should operate "with special force" in this situation of communication overload and lack of clear criteria for evaluating research. Yet citation inequality, which ought to be a good indicator of the Matthew effect, shows no increase at all for biologists. All of this tends to support the Coles' (1973) contention that accumulative advantage is strongest where individuals are rewarded according to their merits. Clearly, however, more comparative research is required to corroborate this conclusion.

Composition changes.[11] A troublesome alternative is the possibility of selective attrition and recruitment during the career course of particular cohorts. Although scientists seldom leave science for entirely different occupations, there is a moderate amount of shifting between scientific disciplines (Harmon, 1965). If such inflows and outflows occur selectively, they could produce increasing inequality. We cannot completely reject this possibility, yet we think it more reasonable to argue the reverse – that attrition and recruitment will reduce inequality over the career course rather than increase it. In such highly competitive occupations as science, we can expect that the least productive persons will be those most likely to leave the occupation or move into less competitive fields. Similarly, those who move into a new field are likely to be more productive than those who have left it. Inter-field mobility, therefore, will magnify differences between fields while reducing differences within them. Although we cannot test this possibility for outflows, we were able to control for inflows by computing the Gini index over age strata for only those scientists whose current field was the same as their Ph.D. field. The results were quite similar to those for the total sample, and we conclude that inflows have little impact on the distribution.

A similar problem arises from the fact that the sample includes only university scientists, while a large proportion of scientists work in

[11] Because of the questionable validity of citations as a measure of productivity, noted earlier, we restrict our attention in the remainder of the paper to the distribution of publications.

industry or government (the proportion varying greatly by field). Although the proportion of scientists who are in universities is fairly stable over cohorts (Harmon, 1965), there is considerable movement between these sectors throughout the career span (National Research Council, 1968). Again, however, it seems reasonable to argue that whatever selective movement occurs will reduce inequality over the career-course rather than increase it. The young university scientist who has published little is likely to move to industry, while the highly productive industrial scientist is likely to move to a university.

Cohort differences. We interpret the increase in productivity inequality over the eight age strata as representing the pattern a single cohort would exhibit over its career. Certainly, if all cohorts started their careers at the same level of inequality and increased their inequality at the same rate, then cross-sectional data stratified by age would precisely mirror the longitudinal changes. Unfortunately, an identical pattern could be produced if older cohorts in the sample had begun their careers at high levels of inequality while younger cohorts started out at lower levels, each maintaining the same level throughout their careers. It is difficult, however, to think of changes in the structure or environment of science that might have produced such a pattern.

One fundamental change in science is that each new cohort of scientists tends to be much larger than the one that preceded it. Price (1963) argues that more extensive recruitment to science implies that many less talented individuals are entering scientific careers. Similarly, Zuckerman (1970) hypothesizes that "increasing recruitment into science will enlarge the differences between the most and least talented." If correct, however, this hypothesis would imply the reverse of what we have observed, with older age strata showing more equality than younger ones. Either the hypothesis is false, or else some opposing process, such as accumulative advantage, has a strong enough effect to obscure the expected result. Thus, our analysis may actually *underestimate* the increase in inequality that would be observed longitudinally for a single cohort.

Intervening Processes

Having considered several alternatives, we conclude that accumulative advantage is the most plausible explanation for the increase in productivity inequality over the age strata. As further evidence, we now examine the relationship between productivity inequality and the distribution of resources, a variable which we have argued is a central component of the accumulative advantage process. We hypothesize that resources become more unequally distributed as scientists age and, more important, that the distribution of resources mediates the relationship between career age and productivity inequality.

Our measures of the distribution of resources are (1) the Gini index of the reported percentage of work-time spent on research, (2) the Gini index of the number of research assistants, and (3) the proportion who report that they "always" get the grants they seek. To test the hypotheses we divide the sample into thirty-three age strata by years since Ph.D., and compute the distribution measures within each age stratum. For these smaller age strata, the regression slope of the Gini for publications on career age is only four percent less than that reported in our earlier analysis, indicating that the principal finding is not very sensitive to the size of the age groupings.

Our hypothesis is expressed as a path model in which resources act as intervening variables between career age and productivity inequality (Figure 2.) Although the Gini for research assistants and the proportion who got grants were both positively related to career age, they had virtually no effects on the Gini for publications and so were eliminated from the model. The Gini for research time, on the other hand, is strongly associated with career age and has a substantial effect on publication inequality. Indeed, it mediates slightly over half the total effect of career age on publication inequality. This result supports the contention of Zuckerman and Merton that "research role attrition" is an important feature of the aging process in science. Noting Harmon's (1965) finding that time spent on research declines with age, they argue that such changes in research activity occur differentially: "the more productive scientists, recognized as such by the reward system of science, tend to persist in their research roles," while less able scientists shift to other roles, primarily to administration.

The increase in productivity inequality with career age might simply reflect the fact that some scientists stop publishing at a certain point in their career, and thus accumulative advantage might be primarily an all-or-nothing

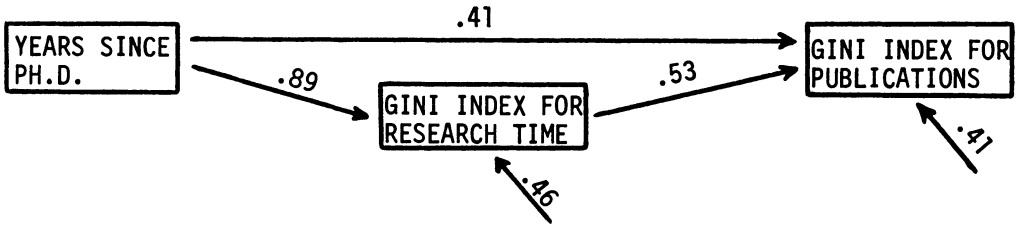

FIGURE 2. Path Model for 33 Age Strata of Scientists Relating the Inequality in Publications to Years Since Ph.D. and the Inequality in Research Time

process. To examine this possibility, we added to the previous model the proportion of scientists in each stratum who published nothing at all in the preceding five years. As shown in Figure 3, about 25 percent of the variance in publication inequality is accounted for by changes in this one segment of the distribution, thus mediating roughly half the total effect of career age. Moreover, *all* the effect of research time inequality is meditated through the proportion of non-publishers (the direct path was near zero), suggesting that role attrition leads some scientists to stop publishing entirely, but doesn't affect the output of those who continue publishing. Investigation at the individual level is needed to corroborate this result. Note further that when the proportion of non-publishers is included in the model, inequality of research assistants has a small indirect effect on publication inequality. Moreover, as the proportion of scientists within a stratum who are biologists goes up, the inequality in the distribution of research assistants diminishes.

Other Empirical Implications

An important feature of accumulative advantage as we have interpreted it, is that resources and the motivation to publish flow to those scientists with high esteem in the scientific community, and that esteem flows to those who are highly productive. This suggests that as a cohort of scientists ages, the fit between productivity, resources, and esteem should steadily improve, enabling us to make an additional assessment of the reinforcement hypothesis. Specifically, we expect that the covariances (within age strata) between resources and esteem and between esteem and publications are an increasing function of career age (across age strata). The simultaneous equation system presented earlier formally implies these results. (We use the covariance rather than the correlation because the latter has an upper limit of one, making linear changes unlikely, and is not independent of the variances, which are also expected to increase with career age.)

Most of these hypotheses are supported by the results in Table 2. Here we measure esteem by the log of the number of citations, and resources by percentage of work-time spent on research, size of research team (number of research assistants plus number of postdoctoral fellows), the ACE (Cartter, 1966) rated quality

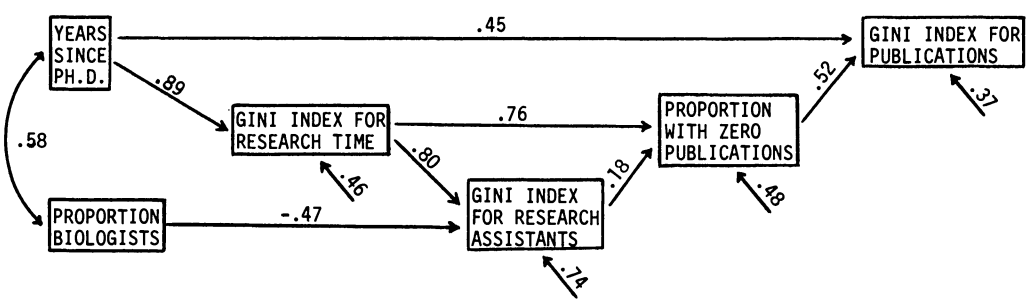

FIGURE 3. Path Model for 33 Age Strata of Scientists Relating the Inequality in Publications to Years Since Ph.D., the Proportion Biologists, and Selected Intervening Variables.

Table 2. Product-moment Correlations between Years since Ph.D. and Selected within Strata Covariances, 33 Age Strata of Scientists

Within Strata Covariances between . . .	Correlation with Years since Ph.D.
Log of citations and log of publications	.81*
Log of citations and research time	.78*
Log of citations and departmental quality	.72*
Log of citations and research assistants	.47*
Log of citations and ease of grants	.10

*Significant at the .01 level.

of department faculty and reported ease of getting grants. We calculated the appropriate covariances within each of the thirty-three age strata, then computed the correlations between these covariances and the years since degree, across the age strata. All but one of the correlations are in the expected direction and significant at the .01 level. The correlation of .78 between career age and the covariance between citations and research time supports the Zuckerman and Merton (1972) hypothesis that research-role attrition is associated with lack of recognition. Furthermore, we find that the correlation between years since degree and the covariance between citations and publications is very high (.81), even though there are artifactual reasons for it to be lower.[12]

SUMMARY AND CONCLUSION

The hypothesis of accumulative advantage in science implies that the distribution of publications and citations will exhibit increasing inequality as a cohort passes through its career. In the absence of cohort differences, this same pattern should be observed in a cross-sectional sample stratified by age. Our data corroborate this consequence for three disciplines: physics, chemistry and mathematics. Specifically, when productivity inequality is measured by a Gini index of publications or citations, there is a substantial, nearly-linear increase from younger to older career-age groups.

Biologists, on the contrary, show virtually no change in citation inequality and only a small increase in publication inequality. We suggest that this anomaly can be explained by the relatively low consensus and poor communication in biology, both of which inhibit the efficient allocation of rewards and resources according to merit.

We have also presented evidence that the increasing inequality of publications can be partly explained by an intervening process of role re-allocation in which the distribution of time spent on research becomes more unequal.

Finally, we found that the fit between scientists' resources, productivity, and esteem improves over the career course.

This evidence in support of the accumulative advantage hypothesis does not disconfirm the heterogeneity hypothesis, however. In fact, our analysis still leaves a major source of inequality unaccounted for: that which is observed among the youngest age strata. Either a much stronger reinforcement process operates during the educational career, or else there are basic underlying differences in scientists' abilities or motivation to do productive research.

[12] As we stressed before, our publication measure only includes scientists' work for the previous five years, while the citation measure may include citations to the life-work of each scientist. On this basis, one would expect citations and publications to be less highly associated for older scientists simply because their recent publications make up a smaller segment of all their cited works, as evidenced by a correlation of −.70 between scientists' career age and the mean year of publication for their works cited in 1966.

REFERENCES

Aitchison, J. and J.A.C. Brown.
1957 The Lognormal Distribution. Cambridge University Press.

Alker, Hayward R., Jr., and Bruce M. Russett
1964 "On measuring inequality." Behavioral Science 9 (July): 207.

Allison, Paul D. and John A. Stewart
1973 "Productivity differences among scientists: evidence for cumulative advantage." Madison: University of Wisconsin Department of Sociology, unpublished paper.

American Chemical Society
1965 Directory of Graduate Research. Washington, D.C.

Bartholomew, David J.
1967 Stochastic Models for Social Processes. New York and London: John Wiley and Sons.

Bayer, Alan E. and John Folger
1966 "Some correlates of a citation measure of productivity in science." Sociology of Education 39 (Fall): 381-9.

Britton, John P.
1964 "The productivity of scientists: a prelude to manpower studies." Mimeographed paper. Yale University.

Cartter, Allan M.
1966 An Assessment of Quality in Graduate Education. Washington: American Council on Education.

Cole, Jonathan R.
1970 "Patterns of intellectual influence in scientific research." Sociology of Education 43 (Fall): 377-403.

Cole, Jonathan R. and Stephen Cole
1967 "Scientific output and recognition: a study in the operation of the reward system in science." American Sociological Review 32 (June): 377-90.
1972 "The Ortega hypothesis." Science 178 (October 27): 368-75.
1973 Social Stratification in Science. Chicago University Press.

Davis, Harold T.
1941 The Analysis of Economic Time Series. Principle Press.

Hagstrom, Warren O.
1967 "Competition and team work in science." Final Report to the National Science Foundation for research grant GS-657, mimeographed. University of Wisconsin, Madison.
1968 "Departmental prestige and scientific productivity." Paper presented at the 63rd Annual Meeting of the American Sociological Association, Boston, Mass., August 27.
1970 "Factors related to the use of different modes of publishing research in four scientific fields." Pp. 85-124 in Carnot E. Nelson and Donald K. Pollock (eds.), Communication Among Scientists and Engineers. Lexington, Mass: Lexington Books.
1974 "Competition in science." American Sociological Review 39 (February): 1-18.

Harmon, Lindsey R.
1965 Profiles of Ph.D.'s in the Sciences: Summary Report on Follow-up of Doctorate Cohorts, 1935-60. Washington: National Academy of Sciences – National Research Council, Publication 1293.

Leavens, Dickson H.
1953 "Communication from Dickson H. Leavens." Econometrica 21 (October): 630.

Lodahl, Janice Beyer and Gerald Gordon
1972 "The structure of scientific fields and the functioning of university graduate departments." American Sociological Review 37 (February): 57-72.

Lotka, Alfred J.
1926 "The frequency distribution of scientific productivity." Journal of the Washington Academy of Sciences 16 (June 19): 317.

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

National Research Council, Office of Scientific Personnel
1968 Careers of Ph.D.'s: Academic Versus Nonacademic. Washington: National Academy of Sciences, Publication 1577.

Price, Derek J. de Solla
1963 Little Science, Big Science. New York: Columbia University Press.

Riley, Matilda White, Marilyn Johnson and Anne Foner
1972 Aging and Society: Vol. III. A Theory of Age Stratification. New York: Russell Sage Foundation.

Shockley, William
1957 "On the statistics of individual variations of productivity in research laboratories." Proceedings of the Institute of Radio Engineers 45 (March): 279-90.

Simon, Herbert A.
1957 Models of Man. New York: John Wiley and Sons.

Spilerman, Seymour
1970 "The causes of racial disturbances: a comparison of alternative explanations." American Sociological Review 35 (August): 627-49.

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

Zuckerman, Harriet
1968 "Patterns of name ordering among authors of scientific papers: a study of social symbolism and its ambiguity." American Journal of Sociology 74 (November): 276-91.
1970 "Stratification in American science." Pp. 235-57 in Edward O. Laumann (ed.), Social Stratification: Research and Theory for the 1970's. Indianapolis: Bobbs-Merrill.

Zuckerman, Harriet and Robert K. Merton
1972 "Age, aging and age structure in science." Pp. 292-356 in Matilda White Riley, Marilyn Johnson, and Anne Foner (eds.), Aging and Society: Vol. III. A Theory of Age Stratification. New York: Russell Sage Foundation.