AN ECONOMIC MODEL OF THE LIFE-CYCLE RESEARCH PRODUCTIVITY OF SCIENTISTS

A. M. DIAMOND, Jr.

Department of Economics
The Ohio State University (USA)

(Received August 30, 1983)

Scientific productivity is constant as a scientist ages according to recent studies relying mainly on quantity measures of productivity. An economic model of the life-cycle productivity of scientists is presented which implies that the number of citations made to a scientist's previous work will decline with age. The implication could be consistent with the finding of constant quantity output with age if the decline in quality (as measured by number of citations per article) is large enough.

Lehman is a still oft-cited study (27 citations from 1979–1981), argued that the productivity of scientists declines continuously after age 30. To defend his view, with regard to mathematicians, Lehman went through Cajori's History of Mathematics in order to note all datable contributions to mathematics by then-deceased individuals. Of the 938 contributions on his initial list, 583 were made by mathematicians who lived to 70 or beyond. By matching the contributions to the mathematician's age when the contribution was made, Lehman came up with roughly the following profile: Although Lehman gives 16 causes for the decline in productivity after 30, (including increasing "maladjustment in sex life"), his first four emphasize declining physical capacities.

Common wisdom still sides with Lehman, but the current consensus among the experts is that Lehman was wrong. Finding adumbration in the late 1950's papers of Dennis, the new view that productivity is constant with age is shared by Stern, Zuckerman and, most formidably, Cole. Although Cole makes a flawed attempt to measure changes in quality of articles with age, his most solid evidence is that the quantity of articles published remains constant. The evidence of Stern and Zuckerman also concerns quantity of publications and adds robustness to Cole's result. As a policy implication of his results, Cole concludes that we need not fear, as we would if Lehman were right, that the increasing average age of scientists over the next 20 years will reduce the nation's "scientific capacity".
No less surprised than others by Cole's findings are those economists whose human capital models imply that as the end of life approaches, the stock of human capital, and hence productivity, declines because no new investment offsets depreciation. Since most scientific research takes place in the university and since the objective function of the university is not well understood, economists are divided about whether scientists' research is directly demanded by the university or is human capital that increases the scientists' productivity in other directly demanded activities such as teaching, fund-raising, or administration. Whichever interpretation turns out to be most fruitful for other purposes, both have the same implication about the relation of research productivity to age: that research and productivity should increase up to some peak age and then decline.

A simple human capital investment model, following closely the work of Becker, will be adapted here to the situation of the research scientist in order to derive testable implications concerning the scientist's life-cycle productivity. If science is an enterprise aimed at providing general explanations of an ever-wider range of phenomena, then economics has a plausible claim to being the most scientific of the 'social sciences'. The model presented here is characteristic of many economic models in that it explains behavior in terms of income maximization under constraints. To the extent that the behavior of scientists is explainable in such terms, we have further illustration of the power of the economist's model of man. To the extent that the behavior of scientists is not explainable in such terms, then that will hearten those who view science as fundamentally different from other human activities in spirit and aim.

The version of the model that follows could be reformulated to provide the greater rigor that is possible with the mathematically moe sophisticated techniques of the calculus of variations or of optimal control theory. Based on past applications of these techniques, few if any, new empirical implications would result. The techniques are
therefore avoided, in the belief that the added mathematical complexity would be too high a price to pay for the increase in rigor.

The model has the scientist maximizing $V$, the discounted sum of his income in the current year, $i$, and all future years until retirement, $n$. Income in any year is the product of time spent working, $t_{wi}$, his stock of prestige capital, $K_i$, and the 'wage' rate per unit of human capital, $W_i$ (assumed constant). The discount factor is $1/(1 + r)$ where the internal rate of return, $r$, is assumed constant across periods. The preceding, expressed formally is:

$$V = \sum_{i=1}^{n} \frac{WK_{i}t_{wi}}{(1 + r)^i}$$

(1)

For the research scientist, 'work' consists of administration, lecturing, dissertation supervising and 'gate-keeping' activities such as refereeing. The human capital that increases his value in these activities is taken to be his professional prestige which, in a more complete model, would itself be a function of the importance of a scientist's past work to research being done at the current frontiers. A scientist's professional prestige can thus be proxied by the current citations that his peers make to all of the scientist's past works. The scientist's total time in year $i$ is labelled $T_i$ and is allocated between time spent in academic work activities, $t_{wi}$, and time spent producing publications that will yield citations, $t_{ki}$. The time constraint is thus:

$$T_i = t_{wi} + t_{ki}.$$  

(2)

The stock of prestige in year $i$, $K_i$, is given by the accounting identity:

$$K_i = K_{i-1} + g(t_{ki-2}, K_{i-2}) + \delta K_{i-1}$$

(3)

where $\delta$ is the depreciation rate (assumed constant) of prestige and $g(t_{ki-2}, K_{i-2})$ is the production function for prestige. (The motivation for the $i - 2$ subscript is that roughly two years pass before a publication begins to be cited.)

In this model there are two choice variables in each period. Since there are $n$ periods, there are $2n$ choice variables over the entire span. Maximizing income (1) subject to constraints (2) and (3) is equivalent to maximizing the following Lagrangean:

$$L = \sum_{i=1}^{n} \frac{WK_{i}t_{wi}}{(1 + r)^i} + \mu_i(T - t_{wi} - t_{ki})$$

(4)

where $\mu_i$ is the marginal value of time. Maximizing (4) results in the $2n$ equilibrium conditions:

$$\frac{\partial L}{\partial t_{wi}} = \frac{WK_i}{(1 + r)^i} - \mu_i = 0 \quad i = 1, \ldots, n$$

(5)
A. M. DIAMOND JR.: ECONOMIC MODEL OF PRODUCTIVITY

\[
\frac{\partial L}{\partial t_{ki}} = \sum_{j=1+2}^{n} \frac{W_{t_{wj}}}{(1+r)^j} \frac{\partial K_j}{\partial t_{ki}} - \mu_i = 0 \quad i = 1, \ldots, n. \tag{6}
\]

Since in equilibrium the marginal value of time, \( \mu \), must be equal in all its uses, (5) and (6) can be combined to yield:

\[
0 = \frac{W_{K_i}}{(1+r)^j} \sum_{j=1+2}^{n} \frac{W_{t_{wj}}}{(1+r)^j} \frac{\partial K_j}{\partial t_{ki}} \quad i = 1, \ldots, n. \tag{7}
\]

Equation (7) shows that time spent producing publications, \( t_k \), would decline for two reasons. The first is that the marginal return would decline as the years left diminished. The second assumes that the production function of human capital is non-neutral (i.e., \( \frac{\partial K_j}{\partial t_k} \) does not increase sufficiently with increasing \( K \) to offset the effect of increasing \( K \) on the value of time at work). In that case, time spent producing publications would decrease as the cost of investment, in income foregone, increased. Thus equation (7) implies the following age profiles for investment time, \( t_k \), and prestige stock, \( K \): The assumption of a constant 'wage' rate for each unit of prestige stock, implies that the life-cycle profile for income will have the same concave shape as the profile for \( K \). On the standard assumption that added human capital in a period is a concave function of the time spent acquiring it in the period, the first profile implies the following profile for human capital invested (i.e., prestige):

![Diagram](image)

**Fig. 2.**
If a scientist's prestige is proxied by the current number of citations to his work, then the final profile implies that number of citations should decline with age. At first glance this result might seem inconsistent with the empirical findings of Cole and others about the constancy of quantity produced. Certainly it does contradict the generalization Cole makes from his findings that the rate of scientific advance need not be slowed by an aging of the scientific population. The specific findings of Cole, however, might be consistent with the model if scientific productivity had two components: quantity and quality. Then, one component (quantity) might remain nearly constant over most of the life-cycle provided the other (quality) declined sufficiently to reduce to over-all investment in prestige.

To see somewhat more rigorously how this might happen, consider the simple case:

\[ g_i = N_i Q_i \]  

(8)

where \( g_i \) is new citations in period \( i \), \( N_i \) is number of articles in period \( i \) and \( Q_i \) is average citations per article in period \( i \). If \( N \) and \( Q \) were each simple exponential functions of time in period \( i \) (henceforth the \( i \) subscripts will be suppressed) then:

\[ N = t_n^a \]  

(9)

\[ Q = t_q^b \]  

(10)

where \( t_n \) and \( t_q \) equal \( t_k \), the total time spent investing:

\[ t_k = t_n + t_q \]  

(11)

If \( t_k \) can be taken as determined by an earlier stage of the maximization process, then the scientist’s problem at this stage may be interpreted as how to maximize his prestige
(NQ) subject to the time constraint (11). Solving for the factor demand equations results in:

\[ t_n^* = \frac{at_k}{a + b} \]  

\[ t_q^* = \frac{bt_k}{a + b} \]  

Substituting (12) into (9) and (13) into (10) yields:

\[ N = \left( \frac{at_k}{a + b} \right)^a \]  

\[ Q = \left( \frac{bt_k}{a + b} \right)^b \]  

If \( t_k \) decreased over the life cycle then from (14) and (15) the product NQ must decrease. If, as Cole et al. believe, N is constant over the life cycle, then Q must diminish and the ratio \( \frac{N}{Q} \) must increase. To see what this implies about the relative time intensity of production of quantity and quality express the ratio as:

\[ \frac{N}{Q} = \frac{\left( \frac{at_k}{a + b} \right)^a}{\left( \frac{bt_k}{a + b} \right)^b} \]  

and then differentiate by \(-t_k\):

\[ \frac{d}{d-t_k} \frac{N}{Q} < 0 \quad \text{as} \quad a < b \]  

So as the time spent investing decreases, an increase in the ratio \( \frac{N}{Q} \) implies that \( a < b \), which means that quality is more time intensive than quantity. The important result is that if quality is more time-intensive than quantity, then an overall decline in research productivity with age is consistent with the constant quantity output found by Cole, provided that quality declines sufficiently with age.

The simple economic model of citations presented here has provided plausible and testable implications about the life-cycle profile of citations to a scientist's work. Most notably, the model implies that: (1) the life-cycle salary profile is concave, (2) the life cycle citation profile is concave and (3) either the quantity of output or the average quality per article...
(or both) decline with age. The model thus has important bearing on the controversial relationship between age and scientific productivity. One avenue for future research is to test the model using longitudinal data on research scientists. A new data set containing information on citations, age and salaries for scientists and mathematicians at Berkeley and the University of Illinois at Urbana is now being analyzed in order to test both the implications of the model and the robustness of previous empirical generalizations. Another possible avenue for future research would be to investigate the relative time intensity of various scientific activities in order to predict more precisely how participation in these activities will change with age.

Notes and References

of the model for time to promotion without necessarily maintaining the specific distributional implications for papers production" (p. 172, italics in original). Despite Price's concurrence with Cole et al., the current consensus about age/productivity profiles is not total. For a sample of 20 past presidents of the American Economic Association, Eagly found that number of articles peaked at 6 per year at age 47 and declined to about 2 per year at age 67 (R. Eagly, Contemporary Profile of Conventional Economists, History of Political Economy, 6, No. 1, (1974) 88). For a sample of 5079 scientists, Bayer and Dutton compare 6 functional forms of the quantity/age relation of find out which one provides the 'best fit'. Since they require that all the betas be significant for the model that 'fits best', they rule out from the start the possibility that productivity is constant with age (in which case, the constant, but none of the betas would be significant). Of the remaining 6 functional forms, they find that a 4th degree polynomial of age (which they call the spurt-obsolescence function) provides the best fit (A. E. Bayer, J. E. Dutton, Career Age and Research-Professional Activities of Academic Scientists, Journal of Higher Education, 48, No. 3 (May/June, 1977) pp. 259–282).

[5] Noting correctly that "longitudinal data are superior to cross-sectional data in measuring the effects of age on scientific productivity," (p. 964) Cole uses longitudinal data for a sample of 497 mathematicians who received their Ph. D.'s between 1947 and 1950. For five year periods between 1950 and 1975, Cole finds that mean citations (in 1975) to articles written in each period varied only slightly around one citation per article (p. 695). From this, Cole concludes that quality of output does not decline with age. Cole's method can be criticized on several grounds. Perhaps the most important of these is that in the early periods the sample includes work by low-quality mathematicians who later stop publishing. Since their work diminishes the group mean in the early periods and since the group mean remains constant over all periods, it follows that the mean quality of work for those mathematicians who continue to publish must decline with age.


