PAKISTAN JOURNAL
OF
SCIENTIFIC AND INDUSTRIAL RESEARCH

Vol. 12, No. 4 June 1970

Pakistan J. Sci. Ind. Res., **12,** 315–323 (1970)

THE OPTIMUM SIZE OF RESEARCH GROUPS FOR MAXIMUM EFFECTIVENESS

Part II.—A Theoretical Model, and its Correlation with the Two Basic Empirical Distributions

M. M. Qurashi

P.C.S.I.R. Laboratories, Peshawar

(Received June 1, 1969)

In part I of this paper, a statistical analysis of the sizes of various scientific research organizations was shown to yield two basic distributions for density of scientific effort versus their size, expressed as the number, N, of scientific officers in a research unit. In order to analyze the factors underlying these two distinct distributions, a theoretical model is here set up to represent the output of a research unit, consisting of N scientific officers, as a function of their mutual interactions. These interactions are represented by (i) an interaction parameter, m, (ii) a wasteful-effort parameter, α, and (iii) a critical size parameter, N_0, and the following equation is obtained:

$$\text{Per capita output} \propto 2\left(\frac{N}{N_0}\right)^m \Big/ \left\{1+\left(\frac{N}{N_0}\right)^{m+\alpha}\right\},$$

with $0<m\leqslant 1$ and $1<\alpha<2$. Using the postulate that density of scientific effort $\propto$ (per capita output)^n, this gives

$$\text{Density of scientific effort} \propto \left[2\left(\frac{N}{N_0}\right)^m \Big/ \left\{1+\left(\frac{N}{N_0}\right)^{m+\alpha}\right\}\right]^n,$$

which can be solved to fit the empirically derived distributions when $1.8-m\leqslant\alpha\leqslant 1.9$. Taking α in the middle of this range, i.e. $\alpha=1.83-\frac{1}{2}m$, we obtain $n\simeq 5\frac{1}{2}$ and $N_0=98\pm 10$, with $m=0.20\pm 0.05$ for one distribution and $m=0.6\pm 0.1$ for the second distribution.

The nearly constant value of N_0 shows that the theoretical model is a good approximation, and the two values obtained for m, viz. 0.2 and 0.6, indicate that the interdependence between workers is relatively small in case of agriculture research, as against the high value for industrial research. A similarly high value of m (~ 0.7) is obtained for defence research in the U.K., which again yields $N_0\simeq 100$. The common value of 98 ± 10 for N_0 provides us with a quantitative formulation of the Parkinsonian decrease of efficiency, viz.

$$1\Big/\left\{1+(N/N_0)^{m+\alpha}\right\} = 1\Big/\left\{1+(N/98)^{1.83+m/2}\right\}$$

and it is hoped to study the application of the corresponding output formula to various organizations in a later paper.

Introduction

The study of the sizes of research groups in a quantitative manner is a matter of considerable importance for determining their optimum sizes, so that the most effective utilization can be made of the available manpower and laboratory facilities. In part I of the present series[1] of papers, a statistical analysis was made of the data available for major research organizations[2] in the U.K., Canada and Pakistan, utilizing the concept of "density of scientific effort". The sizes of the institutes were measured in terms of N, the numbers of scientific officer class per institute; and the total number of scientific officers working in laboratories or institutes of a given size e.g. 30–59 scientific officers per laboratory, when reduced to a constant interval of 10 scientific officers (instead of 30 in this case) was used as a measure of this density of scientific efforts. The distribution curves were plotted for this density against N, for various organizations in the three countries.

The patterns for agriculture research (cf. Table 1) and industrial research appeared to be distinct, and were averaged separately, thus yielding the two mean distributions reproduced in Fig. 1. While that for agriculture research[2] (cf. Table 1 and Figs. 1(a) and 2(a)) has a single maximum at $N=28\pm 1$, falling to half-value at $N=9$ and 54, the distribution of Fig. 1(b) for scientific industrial research is bimodal, with main maximum at $N=69\pm 2$. This bimodal distribution can be considered as a composite of the two distributions shown by broken lines, the subsidiary distribution with maximum at $N=28$ being nearly identical with that of Fig. 1(a). In this way, it was found that all the individual distributions (e.g. Fig. 2 (b) and (c)) are made up of various proportions of the following two basic distributions:

(i) Max. at $N=28\pm1$; half-value at $N=9$ and 54.

(ii) Max. at $N=69\pm2$; half-value at $N=52$ and 95.

Even the size-distribution curve for the defence research establishments (Fig. 4 inset) appeared to conform to this pattern. It was accordingly decided to examine the mechanism operative behind these two distributions, and in the present communication we set out a theoretical model for the interactions in a scientific research organization, which leads to a formula for the per capita output of units of various sizes. This is then correlated with the empirical size distributions found previously.

Basis of the Theoretical Model

Starting with the law of diminishing returns, there has been considerable interest in studying the behaviour of organizations and units of various sizes, as well as the possible relations between size and efficiency. In recent years, Parkinson[3,4] has proposed the view that in a large number of cases, the growth of organizations occurs at a steady 5% to 6% per annum, regardless of the total quantity of work done, and in some cases, even this work is actually diminishing. Parkinson lays down two axioms[3] for this growth in size, viz. (1) "An official wants to multiply subordinates, not rivals", and (2) "Officials make work for each other"; and therefrom follows the building up of a pyramid with each official having, on the average, two immediate subordinates. In addition, for a great many activities, one must also consider a third factor, namely the aspects of cross-fertilization of ideas through

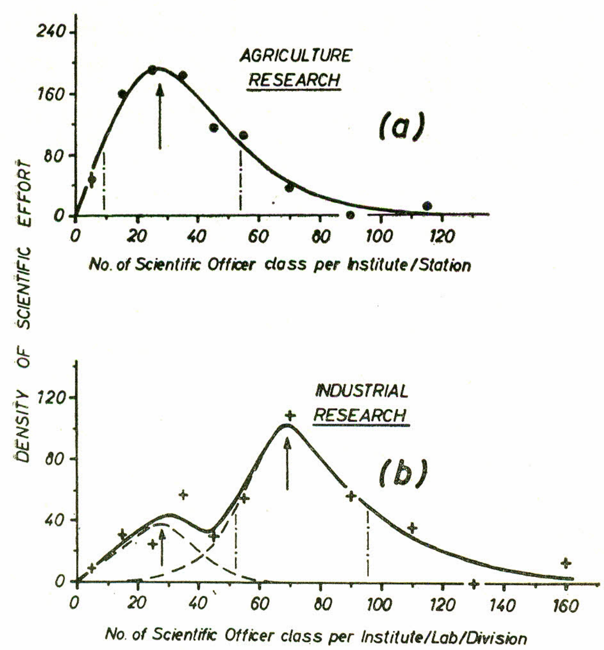

Fig. 1(a).—Mean plots for distribution of density of scientific effort among institutes of various sizes for agriculture research in U.K. and Canada, showing the small scatter about the smooth mean curve, which has its peak at $N=29$ scientific officers per institute. (b) Corresponding mean distribution for industrial scientific research in U.K., Canada, and Pakistan, showing the main peak at $N=69$, and a subsidiary maximum at $N=28\pm2$. The broken-line curves show the two component distributions.

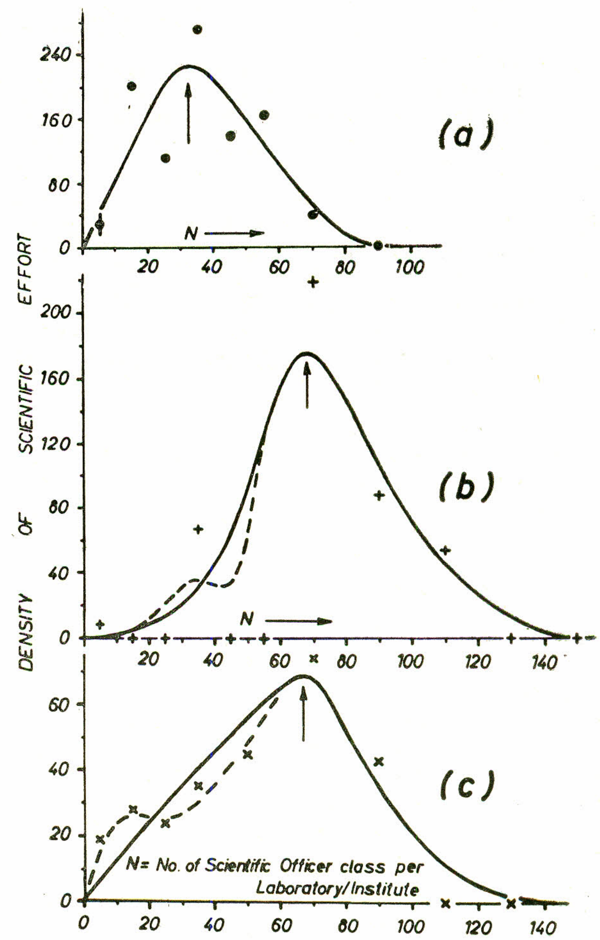

Fig. 2.—Some separate distributions of density of scientific effort among the institutes and laboratories of (a) the Research Branch of the Canadian Department of Agriculture, (b) the National Research Council of Canada, and (c) the Pakistan C.S.I.R. and A.E.C.

TABLE 1.—BREAKDOWN OF SIZES OF INSTITUTES AND STATIONS IN AGRICULTURE RESEARCH.

No. of scientific officer class or equivalent	U.K. Agriculture Research			Canada Agriculture Research			Mean density of scientific effort for agriculture research
	No. of institutes and units	Total no. of scientfic officer class in the range	Scientific effort per interval of 10 officers	No. of institutes and stations	Total no. of scientific officer class in the range	Scientific effort per interval of 10 officers	
Less than 10	$20\times\frac{2}{3}$	67	67	$6\times\frac{2}{3}$	29	29	48
10— 19	8	120	120	14	200	200	160
20— 29	11	269	269	4	111	111	190
30— 39	3	100	100	8	269	269	184
40— 49	2	95	95	3	137	137	116
50— 59	1	50	50	3	163	163	106
60— 79	1	71	36	1	77	38	37
80— 99	0	0	0	0	0	0	0
100—139	1	103	26	0	0	0	13
140—179	0	0	0	0	0	0	0

Note.—The data for units with less than 10 scientific officers have been multiplied by 2/3, because many of these do only tests, trials or data collection.

mutual interactions of the workers, which would be particularly important for scientists or other thinkers. The growth in time as well as the per capita productivity of a unit of any particular size would of course depend on the resultant of these *three* factors, whose relative magnitudes and importance would vary with the nature and type of activity. While both the axioms of Parkinson may operate fully in case of desk workers, it is possible that axiom I has a somewhat lesser significance for creative activities and scientific research.

For setting up a model to analyze the empirical distributions of sizes of research units found above, we may in the first place assume that, in any more or less *stable system* where equilibrium has been reached, the distribution will be such that the highest density of effort is in the particular units having the highest level of efficiency, and *vice versa.* It would then follow that the distribution curve of scientific effort versus size would broadly resemble the curve for efficiency (i.e. output per worker) against size, and be approximately proportional to some power n of this efficiency. The index n can be expected to be fairly large, perhaps around 4 to 8, because (as a general rule) a relatively small drop in efficiency would be associated with a relatively large decrease in probability for the unit concerned. Thus,

$$\text{Density of scientific effort} \propto (\text{per capita output})^n \qquad (1)$$

We may expect n to be indicative of the degree of "cost conciousness" of the organization studied. The problem then becomes one of setting up a plausible model, giving the actual output of a unit of N research workers as a function of N and their mutual interactions.

The Theoretical Model

One might expect that, if there is no wasteful effort, the output could be represented by $\phi(N)$, a monotonously increasing function of N, which represents the integrated numerical effect of the N workers. However, it is well-known that the useful output of any small unit increases at first with the number of workers, but as the unit gets bigger and bigger, a stage of diminishing returns sets in, and the output per worker in fact diminishes, as first laid down in the last century's economics, and later elaborated by Parkinson.[3] Thus, in general mathematical terms, the total output of any unit of N workers can be written as

$$\text{Output} = \phi(N) \times \text{per capita efficiency} \qquad (2)$$

where the per capita efficiency denotes the ratio of actual output per worker to the ideal output if the fact of diminishing returns were *not* operative.

The nature of $\phi(N)$ merits careful consideration. If, for example, the work is of a nature centring around particular individuals, and where there are little or no cross-stimulations or interdependence, then

$$\phi_1(N) = \text{Constant} \times N \qquad (3a)$$

whereas under conditions when each worker is linked up with, supported, and perhaps mentally stimulated by the work of *all* the others, $\phi(N)$ will be proportional to $N \times N$, i.e. N^2, so that

$$\phi_2(N) = \text{Constant} \times N^2 \qquad (3b)$$

In a typical modern research unit, there would probably be intermediate degrees of interaction, although cases are imaginable where an even higher order of mutual interaction than equation 3b may operate. In general, we may put, as a first step,

$$\phi(\mathcal{N}) = \text{Constant} \times \mathcal{N}^{m+1} \quad (4)$$

where $0 < m < 1$, and get

$$\text{Output} = \text{Constant} \times \mathcal{N}^{m+1} \times \text{per capita efficiency} \quad (5)$$

Here the case of $m = 0$ clearly corresponds to zero cross-stimulation.

The variation of per capita efficiency with $\mathcal{N}$ is *not* easy to plot in detail, but we know that the proportion of wasteful effort resulting from men "making work for each other" should increase faster than $\mathcal{N}^{m+1}$ at the very least, but more probably as $\mathcal{N}^2$, or even as $\mathcal{N}^{2+m}$. The argument for $\mathcal{N}^2$ is that each person also generates a wasteful fraction of work proportional to $\mathcal{N}^1$ *for everyone else*, so the total wasteful fraction is proportional to $\mathcal{N}^1 \times \mathcal{N}$, i.e. $\mathcal{N}^2$. Thus, one can (as a first guess) represent the efficiency by the function

$$\text{Per capita efficiency} = 1/\{1 + (\mathcal{N}/\mathcal{N}_0)^{m+\alpha}\} \quad (6)$$

with $\mathcal{N}_0$ a constant, and $1 < \alpha \leqq 2$, which should certainly be valid for $\mathcal{N} < 2\mathcal{N}_0$ or thereabouts. The expression 6 has the characteristics that

(i) it is unity for $\mathcal{N} \to 0$,
(ii) it drops slowly with increasing $\mathcal{N}$ at first, and
(iii) it tends asymptocially to 0, as $\mathcal{N} \to \infty$.

We thus obtain from equations 5 and 6, the expression

$$\text{Output} = \text{Constant} \times \mathcal{N}^{m+1}/\{1 + (\mathcal{N}/\mathcal{N}_0)^{\alpha+m}\} \quad (7)$$

which approaches an upper limit of (constant $\times \mathcal{N}_0^{1+m}$) for very large $\mathcal{N}$, when $\alpha = 1$, but tends to 0 like $1/\mathcal{N}$ when $\alpha = 2$, and we shall attempt to determine α more precisely from the empirical curves. Now we can divide by $\mathcal{N}$, and write

$$\text{Output per worker} = \text{Constant} \times 2\left(\frac{\mathcal{N}}{\mathcal{N}_0}\right)^m / \{1 + (\mathcal{N}/\mathcal{N}_0)^{\alpha+m}\} \quad (8)$$

where the constant has been multiplied by $\frac{1}{2}\mathcal{N}_0^m$, which is itself a constant parameter. If we limit ourselves to $1 < \alpha \leqq 2$, then equation 7 makes the total output *decrease* like $(\mathcal{N}_0/\mathcal{N})^{\alpha-1}$ for large $\mathcal{N}$. This would correspond to Parkinson's description[4] of the *increase* of Admiralty staff as the strength of the British Navy *decreased* in the period 1914 to 1928. We may first restrict ourselves to a discussion of equations 7 and 8 with $\alpha = 1$ and 2, noting that, as an alternative to equations 7 and 8, it can be argued that, if the wasteful interactions leading to equation 6 always vary as $\mathcal{N} \times \mathcal{N}$, i.e. $\mathcal{N}^2$, then the fall in efficiency should also be a function of $(\mathcal{N}/\mathcal{N}_0)^2$. This would lead to the equation

$$\text{Output per worker} = \text{Constant} \times 2\left(\frac{\mathcal{N}}{\mathcal{N}_0}\right)^m / \{1 + (\mathcal{N}/\mathcal{N}_0)^2\} \quad (8a)$$

which really corresponds to equation 8, with $\alpha = 2 - m$.

For $m > 0$, and $\alpha > 0$, the expression 8 can be seen to be zero for $\mathcal{N} = 0$, and to rise rapidly to a maximum at $\mathcal{N}^{\alpha+m} = \left(\frac{m}{\alpha}\right) \times \mathcal{N}_0^{m+\alpha}$, after which it drops asymptotically to zero, as shown in the bottom of Fig. 3(a) and (b) for $m = 0.5$, 1.0 and 1.5, with $\alpha = 1$ and 2, respectively. The precise nature of the asymptotic drop to zero in an actual system depends on the value of α, and the graphs of the above expression should provide a fairly good representation of the variation in output of organizations for values of $\mathcal{N}$ upto $2\mathcal{N}_0$, or even $4\mathcal{N}_0$. Recalling the previous argument of equation 1, we get as a good approximation,

$$\text{Density of effort} \propto (\text{output per worker})^n$$

$$= \text{Constant} \times \left[2\left(\frac{\mathcal{N}}{\mathcal{N}_0}\right)^m \Big/ \left\{1 + \left(\frac{\mathcal{N}}{\mathcal{N}_0}\right)^{m+\alpha}\right\}\right]^n \quad (9)$$

This has a maximum value of ~ 1 at $\frac{\mathcal{N}}{\mathcal{N}_0} = \left(\frac{m}{\alpha}\right)^{1/(m+\alpha)}$, which is $\simeq m + \frac{m}{3}(1-m)$ for $\alpha = 1$ and $0 < m < 1$, and $\simeq m + m(\frac{2}{3} - m)$ for $\alpha = 2$. Numerical values of $(\mathcal{N}/\mathcal{N}_0)_{max}$ at this maximum are given in Table 2 for $m = 0.1$ to 1.2, and $\alpha = 1$ and 2.

Values of the Three Parameters, α, m and $\mathcal{N}_0$

We can now try to fit this expression 9 to the empirical distribution curves of Figs. 1 and 2 for the Agriculture Research and other research organizations. In the first place, we see that we can derive reasonable values of $\mathcal{N}_0$ and m in this way. Clearly, the maxima in Figs. 1 and 2, occurring at $\mathcal{N} \simeq 28$ and 69, would fit in with $\mathcal{N} = 100$ and $0.2 < m < 0.8$, or with $m = 0.4$ and $\mathcal{N}_0$ varying from 50 to 150. So, we may now examine the behaviour of the function 9 with

TABLE 2.

m		0.1	0.2	0.3	0.4	0.5	0.6	0.8	1.0	1.2
$\left(\frac{N}{N_0}\right)_{max}$	$\alpha=1$	.124	.261	.396	.52	.63	.73	.88	1.00	1.09
	$\alpha=2$	.240	.351	.44	.51	.575	.63	.72	.795	.855

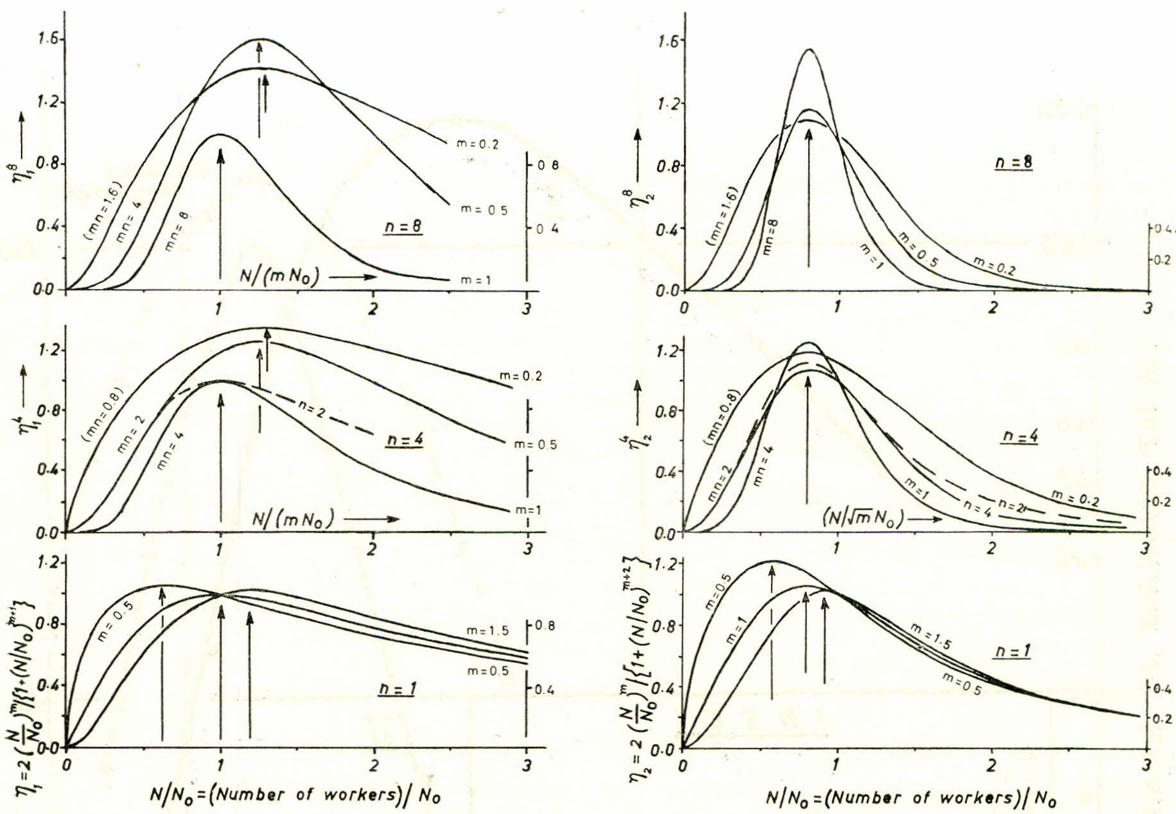

Fig 3 (a) and (b).—The behaviour of the porposed theoretical formula for per capita output, η, and of $(\eta)_n$ for $n=4$ and 8 and various values of the other parameters, α and m. The abscissa in case of η is the ratio N/N_0, while for $(\eta)^n$, which is proportiona to the density of effort, the abscissae are taken as $N/(mN_0)$ and $N/(\sqrt{m}N_0)$ for $\alpha=1$ and 2, respectively, in order to make the position of the maxima nearly independent of the value of m.

$n=4$ and 8, taking $m=0.2$, 0.5 and 1.0, so as to cover the probable range of 0.2 to 0.6. Fig. 3(a) and (b) show the twelve graphs of this function plotted against $N/(m\,N_0)$ for $\alpha=1$ and $N/(\sqrt{m}\,N_0)$ for $\alpha=2$, with the above-mentioned values of n and m, the maxima being now nearly coincident. It is at once apparent that the shape of the graph before the maximum is largely determined by the product, $m\times n$, while the tail is determined by α and n, cf. the broken line curves for $mn=2$, $n=2$. Approximate values of $(m\times n)$ for any particular graph can be estimated from the part near the origin (before it reaches half the peak value), and are found to lie between 1 and 4 for fitting the graphs for the data in Figs. 1 and 2. From this, we may deduce that if n has a fixed value, then $n\geqslant 4$, for m to be less than 1. Furthermore, if $\alpha=2$, then $n=5\pm 1$ and $m=0.2$ to 0.8 (cf. Fig.3 (b)) apparently yield fair agreement with the empirical graphs of Figs. 1 and 2. On the other hand, the calculated graphs for $\alpha=1$ do not reconcile readily with the empirical data.

For a more precise determination of α, m and n for the two basic distributions of Fig. 1, either we can make use of the points at which the curves of

Fig. 1 fall to half their peak values to define three quantities A, L, and R, and thence estimate m and n as a first approximation, or better still plot the two curves of Fig. 1 on "log–log" paper as shown in Fig. 4, when the logarithmic graphs are seen to become linear for small N, as well as for large N. It can be shown from equation 9 that

log (density of effort)

$$= n \log 2 + mn \log \frac{N}{N_0} - n \log \left\{ 1 + \left(\frac{N}{N_0}\right)^{m+\alpha} \right\}$$

which for small N becomes

$$n \log 2 + nm \log \frac{N}{N_0} - n\left(\frac{N}{N_0}\right)^{m+\alpha},$$

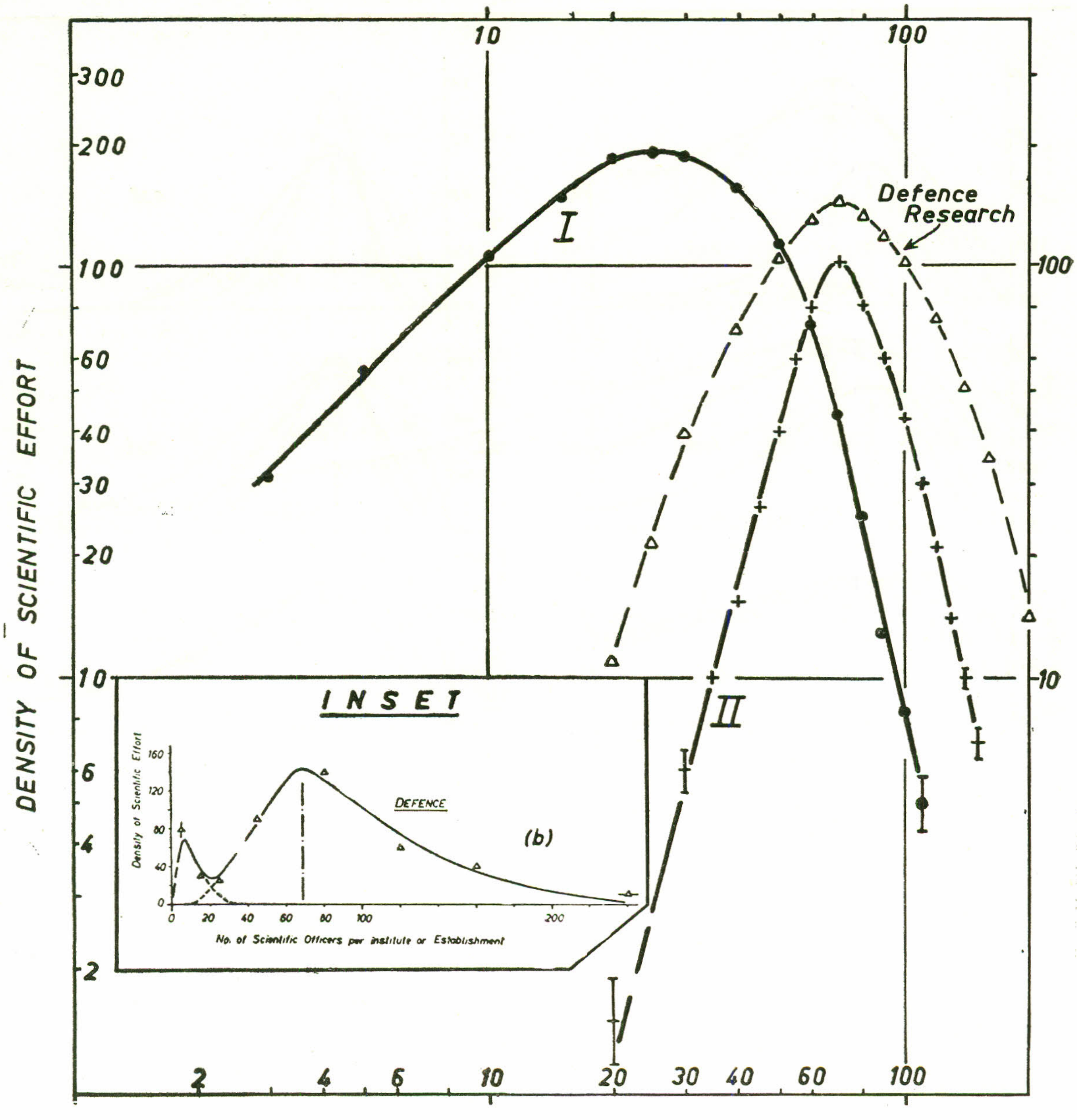

Fig. 4.—Logarithmic plots of the two basic empirical distributions of density of scientific effort (cf. Fig. 1), showing the approximately constant slopes for $N \to 0$ and for large N. The solid circles are for the agriculture research distribution with peak at $N=28$, and the crosses are for the scientific industrial research effort having its peak at $N=69$. The third graph (hollow triangles) is for the defence research in U.K., the actual distribution for which is reproduced in the inset.

and for large N is

$$n \log 2 - \alpha n \log n \frac{N}{N_0} - \left(\frac{N}{N_0}\right)^{m+\alpha}.$$

Thus, the slope of this "log–log" plot will tend to the value $(m \times n)$ for very small N, and to the value $(-n \times \alpha)$ for very large N. More precisely,

$$\text{Slope} = nm - n(m+\alpha)\left(\frac{N}{N_0}\right)^{m+\alpha} \Big/ \left\{1 + \left(\frac{N}{N_0}\right)^{m+\alpha}\right\}$$

$$= -\alpha n \frac{(N/N_0)^{m+\alpha}}{1+(N/N_0)^{m+\alpha}} + \frac{mn}{1+(N/N_0)^{m+\alpha}}$$

$$= - \frac{\alpha n}{1+(N/N_0)^{m+\alpha}} + \frac{nm}{1+(N/N_0)^{m+\alpha}} \quad (10)$$

For $N << N_0$, this gives as before

$$(\text{Slope})_{N\to 0} = nm, \quad (11a)$$

while for $N \sim N_0$, we obtain

$$(\text{Slope})_{N\sim N_0} = \{-\alpha n + mn(N_0/N)^{m+\alpha}\} \div \{1 + (N_0/N)^{m+\alpha}\}.$$

whence

$$\alpha n = -(\text{Slope})_{N\sim N_0} \times \{1 + (N_0/N)^{m+\alpha}\} + mn(N_0/N)^{m+\alpha} \quad (11b)$$

Noting that the "log–log" plots of Fig. 4 extend up to $N = 120 \pm 20$, and that $N_0 \sim 100$, we take $N_0/N \simeq 0.8$, and get $(N_0/N)^{m+\alpha} \simeq 0.7$, and so find, as a first approximation,

$$\alpha n \simeq -1.7 \times (\text{slope})_{N\sim N_0} + 0.7 \times (\text{slope})_{N\to 0}. \quad (12)$$

Equations 11 and 12 can now be used to estimate the quantities $(m \times n)$ and $(n \times \alpha)$ from the "log–log" plots of Fig. 4 for the two basic distributions of Fig. 1. We thus find that

(i) for the distribution with peak at $N = 28$,

$(\text{Slope})_{N\to 0} = (mn)_{\text{I}} = 1.00 \pm 0.05$,

$(\text{Slope})_{N=100} = -4.6 \pm 0.3$, whence from equation 12, $(\alpha n)_{\text{I}}$

$= +4.6 \times 1.7 + 0.7 \times 1.00 = 8.5$, and

(ii) for the distribution with peak at $N = 69$,

$(\text{Slope})_{N\to 0} = (mn)_{\text{II}} = 3.8 \pm 0.2$,

$(\text{Slope})_{N=140} = -5.0 \pm 0.3$, whence from equation 12, $(\alpha n)_{\text{II}}$

$= +5.0 \times 1.7 + 0.7 \times 3.8 = 11.2$.

This yields a mean value of 9.8 ± 1.3 for αn. It follows, taking $\alpha = 2$, that $n \simeq 9.8/2 = 4.9$, whence we get as a first approximation for the two distributions,

$$m_{\text{I}} = \frac{1.00}{4.9} = 0.20, \text{ and } m_{\text{II}} = \frac{3.8}{4.9} = 0.79,$$

and this with the formula $(N/N_0)_{\max} = (m/\alpha)^{1/(m+\alpha)}$ gives us the values

$$(N_0)_{\text{I}} = \frac{2.8}{0.35} = 80, \text{ and } (N_0)_{\text{II}} = \frac{69}{0.71} = 97 \quad (14a)$$

By feeding these values of N_0 back into equation 11(b), we can obtain more accurate values for αn, viz.

$$\alpha(n)_{\text{I}} = 4.6 \times 1.61 + 0.61 \times 1.00 = 8.0,$$
$$\alpha(n)_{\text{II}} = 5.0 \times 1.35 + 0.35 \times 3.8 = 8.1,$$

which gives us as a second approximation,

$$n = \frac{\alpha n}{2} = \frac{8.0 \pm 0.0}{2} = 4.0. \text{ This yields}$$

$$\left.\begin{array}{l} m_{\text{I}} = \dfrac{1.00}{4.0} = 0.25, \text{ and } m_{\text{II}} = \dfrac{3.8}{4.0} = 0.95, \\ \text{whence we get} \\ (N_0)_{\text{I}} = \dfrac{28}{0.40} = 70, \text{ and } (N_0)_{\text{II}} = \dfrac{69}{0.78} = 88. \end{array}\right\} (14b)$$

The third approximation gives

$$n = \frac{\alpha n}{2} = \frac{7.1 \pm 0.0}{2.} = 3.55 \pm 0.0, \text{ whence}$$

$$\left.\begin{array}{l} m_{\text{I}} = \dfrac{1.00}{3.55} = 0.28, \; m_{\text{II}} = \dfrac{3.8}{3.55} = 1.07 \\ \text{and } (N_0)_{\text{I}} = \dfrac{28}{0.425} = 66, \; (N_0)_{\text{II}} = \dfrac{69}{0.82} = 84 \end{array}\right\} (14c)$$

The successive values of n, m and N_0 converge steadily to the ultimate values of $n = 3.3$ and

$m_{\text{I}} = 0.30 \qquad m_{\text{II}} = 1.15$

$(N_0)_{\text{I}} = 64 \qquad (N_0)_{\text{II}} = 83$

When this solution is attempted for $\alpha = 1.5$, it is found (cf. Appendix) that the successive ap-

TABLE 3.—VALUES OF n, m AND N_0 OBTAINED FOR THE TWO DISTRIBUTIONS OF FIG. 1 FOR VARIOUS VALUES OF α.

α	2.0		1.8		1.7		1.6		1.5		1.4	
n	3.3	3.3	4.4	4.2	5.4	4.7	7.5	5.8	>10	7.3	>10~10	
m	0.30	1.15	0.23	0.90	0.18	0.80	0.13	0.66	<0.01	0.52	<0.1~0.38	
N_0	64	83	78	89	90	93	118	102	>150	116	>150~145	

proximations for the first distribution diverge to $n>10$, $m<0.1$ and $N_0>150$; and with $\alpha=1.0$, as also $\alpha=1.3$, the approximations diverge for both distributions. The results obtained with various values of α from $\alpha=2.0$ down to $\alpha=1.4$ are tabulated below, and it is seen that the smallest value of α for a convergent solution are 1.6 for the first distribution (with peak at $N=28$), and 1.4 for the second distribution. Comparing with the corresponding values of m in Table 3, we find that

$$\alpha_{\text{min.}}+m=1.76\pm0.03,$$

so that $\alpha\geq1.76-m$.

On the other hand, if $m<1$, then the maximum value of α appears to be ~1.9, and we may conclude that

$$1.76-m\leq\alpha\leq1.9 \qquad (15a)$$

Taking α in the middle of this range, we get the most probable value as

or
$$\left.\begin{aligned}\alpha&=1.83-m/2\\ \alpha+m&=1.83+m/2\end{aligned}\right\} \qquad (15b)$$

which is intermediate between the two considered probable in section 3.

Having fixed the value of α in the above fashion, we obtain the following values of the remaining parameters, n, m, and N_0 for the two distributions:

$$\left.\begin{aligned}n_{\text{I}}&=5.1,\ m_{\text{I}}=0.20,\ (N_0)_{\text{I}}=86\\ n_{\text{I}}&=6.5,\ m_{\text{II}}=0.58,\ (N_0)_{\text{II}}=109\end{aligned}\right\} \qquad (15c)$$

Interpretation and Conclusions

From these results, we can surmise that n and N_0 are constants, with mean values of

$$n=5.8\pm0.6 \text{ and } N_0=98\pm11, \qquad (15d)$$

while m has a value of 0.20 for the first distribution of Fig. 1 and 0.6 for the second distribution. This surmise is further borne out by the fact that a similar analysis of the U.K. data for defence research establishments yields

$$n_{\text{D}}=4.4,\ m=0.68, \text{ and } N_0=100,$$

all of which are in close agreement with the values found above.

Thus, we can justifiably conclude that:

(i) The expressions 8 and 9 proposed for the output per worker and the density of scientific effort are essentially correct, with $\alpha=1.83-m/2$ and $n\simeq5\frac{1}{2}$,

(ii) the size-parameter, N_0, is a constant with a value close to 100, and

(iii) the interdependence parameter, m, has a value of 0.2 for agriculture research, and a value of 0.6 to 0.7 for activities like industrial research and defence research.

It is hoped to examine the formula 8 for output more thoroughly and to study the yearwise variation in output of various organizations in a later paper. This would throw further light on the problems connected with possible strengthening and regrouping of existing research laboratories, etc., so as to keep the numbers of Scientific Officer class in each within the optimum limits of 28 and 69.

References

1. M.M. Qurashi, Pakistan J. Sci. Ind. Res., **12,** 1 (1969).
2. (a)S. Zuckerman, *The Management and Control of Research and Development* (H.M. Stationery Office, London, 1961; reprinted 1964), pp. 75, 112–121; (b) *Research Report*—1967, Canada Department of Agriculture, pp. 1–444; (c) *Review of the National Research Council of Canada* (Ottawa, 1968), pp. 232–33.
3. C.N. Parkinson, *Parkinson's Law and Other Studies in Administration* (The Riverside Press, Cambridge, Mass., U.S.A. 1957), pp. 4–6.
4. C.N. Parkinson, *ibid.*, pp. 7–10.

Appendix

Fitting of Equation 9 for α=1.5, 1.3, and 1.1.

Case (a): $\alpha=1.5$

We next try a smaller value of α, viz. 1.5, in order to examine if a stable, convergent solution is still possible. The first approximation gives us $n \simeq 9.8/1.5 = 6.6$, and thence

$$\left.\begin{aligned} m_{\mathrm{I}} &= 1.00/6.6 = 0.15, \text{ whence } (N_0)_{\mathrm{I}} = 28/0.24 = 116 \\ m_{\mathrm{II}} &= 3.8/6.6 = 0.57, \text{ whence } (N_0)_{\mathrm{II}} = 69/0.63 = 109 \end{aligned}\right\} \quad (16a)$$

The second approximation follows by feeding these values into equation 11(b), viz.

$$\left.\begin{aligned} &(\alpha n)_{\mathrm{I}} = 4.6 \times 2.28 + 1.28 \times 1.00 = 11.8, \\ &\text{so that } n_{\mathrm{I}} = 11.8/1.5 = 7.9, \\ &\text{whence } m_{\mathrm{II}} = 1.00/7.9 = 0.127, \text{ and } (N_0)_{\mathrm{I}} = 28/0.22 = 127, \\ &(\alpha n)_{\mathrm{II}} = 5.0 \times 1.60 + 0.60 \times 3.8 = 10.3, \\ &\text{so that } n_{\mathrm{II}} = 10.3/1.5 = 6.9, \\ &\text{whence } m_{\mathrm{II}} = 3.8/6.9 = 0.55, \text{ and } (N_0)_{\mathrm{II}} = 69/0.62 = 111. \end{aligned}\right\} \quad (16b)$$

The third approximation gives

$$(\alpha n)_{\mathrm{I}} = 4.6 \times 2.84 + 1.48 \times 1.00 = 12.9,$$
$$\text{so that } n_{\mathrm{I}} = 12.9/1.5 = 8.6,$$

$$\left.\begin{aligned} &\text{whence } m_{\mathrm{I}} = 1.00/8.6 = 0.116, \text{ and} \\ &(N_0)_{\mathrm{I}} = 28/0.205 = 137, \\ &(\alpha n)_{\mathrm{II}} = 5.0 \times 1.62 + 0.62 \times 3.8 = 10.5, \\ &\text{so that } n_{\mathrm{II}} = 10.5/1.5 = 7.0, \\ &\text{whence } m_{\mathrm{II}} = 3.8/7.0 = 0.54, \text{ and } (N_0)_{\mathrm{II}} = 69/0.61 = 113. \end{aligned}\right\} \quad (16c)$$

The values of n_{II} and $(N_0)_{\mathrm{II}}$ can be seen to converage slowly towards $n_{\mathrm{II}} = 7.3$ and $(N_0)_{\mathrm{II}} = 116$, while those of n_{I} and $(N_0)_{\mathrm{I}}$ for the distribution with peak at $N = 28$ appear to diverge rapidly. This is to be expected when N_0 becomes much greater than the highest value of N, so that the factor $(N_0/N)^{\alpha+m}$ in equation 11(b) becomes increasingly greater than 1, as shown by the values of $(\alpha n)_{\mathrm{I}}$ in equations 16(b) and 16(c).

Thus, the fourth approximation gives us

$$(\alpha n)_{\mathrm{I}} = 4.6 = 2.67 + 1.67 \times 100 = 14.1,$$
$$\text{so that } n_{\mathrm{I}} = 14.1/1.5 = 9.4,$$
$$\text{whence } m_{\mathrm{I}} = 1.00/9.4 = 0.106, \text{ and } (N_0)_{\mathrm{I}} = 28/0.19 = 147.$$

Case(b) : $\alpha = 1.0$

Similarly, we examine the behaviour of n_{II} and $(N_0)_{\mathrm{II}}$ when $\alpha = 1.0$, starting with the first approximation, which yields $n_{\mathrm{II}} \simeq 10$, and therefore

$$m_{\mathrm{II}} \simeq 3.8/10 = 0.38, \text{ and } (N_0)_{\mathrm{II}} \simeq 69/0.50 = 138 \quad (17a)$$

Subsituting these values back into equation 11(b), we get for the second approximation,

$$n_{\mathrm{II}} = 5.0 \times 1.98 + 0.98 \times 3.8 = 13.6, \text{ whence}$$

$$m_{\mathrm{II}} = 3.8/13.6 = 0.28, \text{ and } (N_0)_{\mathrm{II}} = 69/0.37 = 186 \quad (17b)$$

which is already beyond the upper limit of the second distribution of Fig. 1.

The third approximation now gives

$$n_{\mathrm{II}} = 5.0 \times 2.44 + 1.44 \times 3.8 = 17.7, \text{ whence}$$

$$m_{\mathrm{II}} = 3.8/17.7 = 0.21_5, \text{ and } (N_0)_{\mathrm{II}} = 6.9/0.28 = 246 \quad (17c)$$

and it is clear that the solution diverges rapidly.

Case (c) : $\alpha = 1.3$

We may also try an intermediate value of α, viz., 1.3, to see whether the solution converges here. The first approximation is $n_{\mathrm{II}} = 9.8/1.3 = 7.5$, whence

$$m_{\mathrm{II}} = 3.8/7.5 = 0.51, \text{ and } (N_0)_{\mathrm{II}} = 69/0.59 = 117 \quad (18a)$$

Subsitituting these values into equation 11(b), we get for the second approximation,

$$(\alpha n)_{\mathrm{II}} = 5.0 \times 1.73 + 0.73 \times 3.8 = 11.4,$$

whence $n_{\mathrm{II}} = 11.4/1.3 = 8.8$, and therefore

$$m_{\mathrm{II}} = 3.8/8.8 = 0.43, \text{ and } (N_0)_{\mathrm{II}} = 69/0.53 = 130 \quad (18b)$$

The third approximation now gives

$$(\alpha n)_{\mathrm{II}} = 5.0 \times 1.88 + 0.88 \times 3.8 = 12.7,$$

whence $n_{\mathrm{II}} = 12.7/1.3 = 9.8$, and therefore

$$m_{\mathrm{II}} = 3.8/9.8 = 03.9_0, \text{ and } (N_0)_{\mathrm{II}} = 69/0.49 = 141 \quad (18c)$$

Similarly, the fourth approximation gives us

$$(\alpha n)_{\mathrm{II}} = 5.0 \times 2.01 + 1.01 \times 3.8 = 13.9,$$

whence $n_{\mathrm{II}} = 13.9/1.3 = 10.7$, and therefore

$$m_{\mathrm{II}} = 3.8/10.7 = 0.35_5, \text{ and } (N_0)_{\mathrm{II}} = 69/0.45_5 = 152, \quad (18d)$$

which shows clearly that this solution is also divergent.