Cowles Foundation Paper 154
Reprinted from Quarterly Journal of Economics, 74, 1960

SOME ASPECTS OF RETURNS TO SCALE IN BUSINESS ADMINISTRATION*

By MARTIN J. BECKMANN

Entrepreneurial capacity is often regarded as the limiting factor in determining the size of the firm, as it undoubtedly is for the traditional one-man enterprise.[1] The modern corporation has in its hierarchy of executives a device for overcoming this hurdle. But, it is thought that such hierarchies operate under diseconomies to scale.[2] The diseconomies are of two kinds: delays in decision-making through bureaucratic "sluggishness"; and increasing cost of administration per worker.[3] Thus it is concluded, that even where increasing returns prevail in production activities, at some level — possibly beyond that which is required at the present — the diseconomies of administration will catch up with the economies of production. There exists then an optimal size of the firm at which both effects are balanced.

Before accepting this conclusion it may be well to examine the supposed diseconomies of scale in management by means of a theoretical analysis of the underlying structure of administrative hierarchy. Consider a firm whose production activities show constant returns to scale. In the long run the scale of plant and the number of workers are strictly proportional to each other and to output, and we may use the number of workers as our measure of the size of the

* This paper owes much to J. Marschak's earlier "Remarks on the Rationale of Leadership," Cowles Foundation Discussion Paper 6 (Oct. 1955) and to his stimulating comments. I am greatly indebted also to C. B. McGuire, to A. S. Manne, and to T. C. Koopmans for helpful suggestions.

1. N. Kaldor, "The Equilibrium of the Firm," *Economic Journal*, XLIV (1934), 60–76. F. H. Knight, *Risk, Uncertainty and Profit* (Boston: Houghton Mifflin, 1921), pp. 286–311. A. Robinson, "The Problem of Management and the Size of Firms," *Economic Journal*, XLIV (1934), 242–57.

2. Cf. G. Stigler, *The Theory of Price* (2d ed.; New York: Macmillan, 1952) pp. 139–40; M. M. Bober, *Intermediate Price Theory* (New York: Norton, 1955).

3. Some writers believe that a further inevitable diseconomy of size is *inefficiency* bred by "Say's Law": that a supply (of offices) creates its own demand (for sham work). An amusing version of this proposition has gained popularity under the name of Parkinson's Law: "Work expands so as to fill the time available for its completion" (D. F. Parkinson, "Parkinson's Law," *The Economist*, London, Nov. 19, 1955, pp. 635–37). In theory, at any rate, there is no basis for the assertion that inefficiency is a necessary attribute of organizational size. For business firms it may be argued, on the contrary, that the relatively efficient firms will tend to be the large ones. We have therefore disregarded this somewhat elusive type of inefficiency in the analysis that follows.

firm. Assume that employees are arranged in a hierarchy of well-defined administrative levels. Our interest will be centered on the amount by which the administrative cost per worker increases, when the size class of the firm is raised through addition of one administrative level. This increment of administrative cost per worker will be called the marginal administrative cost per worker.

Notation

Level of workers	0
Any administrative level	$m = 1, \ldots n$
Level of top executive	n

n will also designate the size class of the firm: An n-level firm has n administrative levels above the level of workers.

Number of employees at the mth level (of course $L_n = 1$)	L_m
Wage rate of employees at mth level	W_m
Wage and salary bill of an n-level firm	C_n
Number of workers of an n-level firm	$E_n = L_0$
Wage of a worker	$W = W_0$

Diseconomies in large-scale organizations must show up in the behavior of the marginal administrative cost per worker. What is the law of increase of this cost? How decreasing are the decreasing returns to scale in business administration?

In Section 1 we shall estimate marginal cost per worker under fairly general assumptions. More structure will be assumed in Section 2 and a sharper bound will be obtained. In Section 3 delays are also considered, and our conclusions are stated.

1. In order to fix structural proportions in the hierarchy, three basic assumptions will be made.

Assumption 1.1

In a given enterprise the number of immediate subordinates per administrator — the "span of control" — is never less than a certain constant a which is greater than one.

$$L_m/L_{m+1} \geqq a > 1 \qquad (m = 0, \ldots n - 1) \text{ where } a > 1.$$

In order to rule out cost increases resulting from reorganization in the administration of a given work-force, we specify that nowhere in the hierarchy shall the number of subordinates to any position be reduced. This means that we also have

$$\Delta L_m \geqq a \, \Delta L_{m+1} \qquad (m = 0, \ldots n - 1)$$

for all changes ΔL_m in the number of employees as the size of the firm increases.

Assumption 1.2

An administrator's salary is greater than that of any immediate subordinate but by not more than a ratio of $b > 1$.

$$1 < W_{m+1}/W_m \leqq b \qquad (m = 0, \dots n - 1)$$

Assumption 1.3

The ratio of salaries from level to level is smaller than the span of control.

$$a > b$$

Assumption 1.1 is not to be confused with that of a "fixed span of control" nor Assumption 1.2 with that of a fixed ratio salary scale (see section 2 below).

For evidence that a tends to be larger than b one may compare the numbers and salaries of persons at the highest and lowest levels of a large firm and see that the spread of the numbers is greater than that of the salaries (see page 468).

To estimate the marginal cost per worker let ΔL_m denote the change in the number of employees at the mth level, as the size class of the firm is increased from n to $n + 1$. The change in the total wage and salary bill then is as follows:

$$\Delta C_n = \sum_{m=0}^{n} W_m \,.\, \Delta L_m + W_{n+1} \,.$$

By assumption 1.2

$$W_m \leqq b^m W$$

Therefore,

$$\Delta C_n \leqq W \,.\, \sum_{m=0}^{n} b^m \, \Delta L_m + W\, b^{n+1}$$

To estimate ΔL_m observe that

(1.1) $\quad \Delta L_n \geqq a - 1$

since there are now at least a administrators at level n where there was only one before.

By assumption 1.1

(1.2) $\quad \Delta L_m \geqq a \, \Delta L_{m+1} \qquad (m = 1, \dots n - 1).$

From (1.1) and (1.2) it follows that

(1.3) $\quad \Delta L_0 \geqq a^n \,.\, (a - 1)$

(1.4) $\quad \Delta L_0 \geqq a^m \, \Delta L_m.$

Consider now

$$\frac{\Delta C_n}{\Delta L_0} \leqq W \, \frac{\sum_{m=0}^{n} b^m \, \Delta L_m}{\Delta L_0} + W \, \frac{b^{n+1}}{\Delta L_0} \,.$$

Substituting (1.4) and (1.3) we have

$$\frac{\Delta C_n}{\Delta L_0} \leqq W\left[\sum_{m=0}^{n}\left(\frac{b}{a}\right)^m + \frac{a}{a-1}\left(\frac{b}{a}\right)^{n+1}\right]$$

$$= W\left[\frac{1-\left(\frac{b}{a}\right)^{n+1}}{1-\frac{b}{a}} + \frac{1}{1-\frac{1}{a}}\left(\frac{b}{a}\right)^{n+1}\right].$$

Since $b > 1$, finally

$$(1.4) \qquad \frac{\Delta C_n}{\Delta L_0} \leqq \frac{W}{1-\frac{b}{a}}.$$

We have obtained the rather surprising result that the marginal cost per worker remains below a definite bound no matter how large the organization. This does not imply, of course, that marginal cost is itself constant or asymptotically constant. But, it shows that under our assumptions marginal cost of size does not increase indefinitely, contrary to what is often asserted. It will be our next task to estimate the size of the bound and the manner in which it is approached. However, this will require more definite assumptions.[4]

The same bound (1.4) can be shown to apply also to the average wage and salary cost per worker.

$$(1.5) \qquad \frac{C_n}{E_n} \leqq \frac{W}{1-\frac{b}{a}}.$$

Another interesting bound is

$$\frac{\text{salary cost}}{\text{wage cost}} = \frac{\text{administrative costs}}{\text{nonadministrative costs}} \leqq \frac{1}{1-\frac{b}{a}}.$$

It may also be interesting to consider the *number* of employees per worker. An estimate is obtained by simply setting $b = 1$ in our

4. Without specifying the administrative structure in detail, one may assume instead that as salaries w increase by a factor of b the number of recipients l decreases at least by the factor $1/a$. In continuous terms this means that the elasticity $\frac{d \log l}{d \log w} \leqq -\frac{a}{b}$. It can then be shown that the wage bill for administration per worker is $\leqq \frac{W}{\log \frac{a}{b}}$, in analogy to the previous model.

previous formula, yielding $\dfrac{1}{1-\dfrac{1}{a}}$. To obtain the number of administrators per worker the number one — the worker himself — must be subtracted. Thus, in an organization where each administrator has not less than a immediate subordinates the total number of administrators per worker is not more than $\dfrac{1}{a-1}$ no matter how large the organization.

There is some empirical evidence in support of (1.5) that the administration cost per worker is independent of the scale of operations in the Electric Power Industry of this country.[5]

2. An extreme case is that in which the bounds a and b are attained at each level, that is:

$$L_{m+1}/L_m = a, \; W_m/W_{m+1} = b,$$

and therefore,

$$L_m W_m/L_{m+1} W_{m+1} = b/a.$$

We shall want to study the approximation of marginal cost to the ceiling (1.5) in this case.

With regard to b, the factor of steepness in the salary scale, there exists some empirical evidence for such a geometric progression.[6] The situation is more complex with respect to a, the so-called span of control.[7] While it is irrelevant for our purposes that this parameter varies from industry to industry,[8] it must be noted that it also varies from level to level in many cases. Thus, "Students of Management have found [the span of control] to exist usually at numbers of 4 to 8 subordinates at the upper level of organizations and from 8 to 15 or more at the lower levels."[9] However, the recommendation of "experts on management" is that the span of control should be held

5. J. McNulty, "Administrative Costs and Scale of Operations in the U. S. Electric Power Industry — A Statistical Study," *Journal of Industrial Economics*, V (1956), 30–44.

6. A special part is played by the presidential salary. It is often out of line with other salaries in the hierarchy, but for a given industry may nevertheless tend to be proportional to the size of the firm. Cf. H. Simon, "The Compensation of Executives," *Sociometry*, Vol. 20 (1957), pp. 32–35 and references given there.

7. Cf. Simon, *Administrative Behavior* (New York: Macmillan, 1947), pp. 26–28; M. Polanyi, *The Logic of Liberty* (Chicago: University of Chicago Press, 1951), pp. 114–22, 170–84.

8. C. H. Koontz and C. O'Donnell, *Principles of Management* (New York: McGraw-Hill, 1955).

9. *Ibid.*, p. 88.

between narrow limits and near 4,[1] the typical span of control in military organizations. Apart from its simplicity the case of a constant span of control has therefore at least some theoretical interest.

A little arithmetic now produces explicit expressions for the number of workers and the total wage bill as well as the marginal cost per worker in an organization of size n.

$$E_n = a^n$$

$$C_n = W \cdot \frac{a^{n+1} - b^{n+1}}{a - b}.$$

$$\frac{\Delta C_n}{\Delta E_n} = \frac{W}{1 - \frac{b}{a}} \left[1 - \frac{b-1}{a-1} \left(\frac{b}{a}\right)^{n+1} \right]$$

The percentage difference from our previous estimate $\frac{W}{1 - \frac{b}{a}}$ is therefore $100 \, \frac{b-1}{a-1} \cdot \left(\frac{b}{a}\right)^{n+1}$ per cent. We shall estimate the size of this term presently. The intuitive reason why marginal cost of size is asymptotically constant is, of course, this: While the administrative staff grows exponentially with the number of tiers in the hierarchy (given the span of control) so does the number of workers at the ground level. The ratio of both tends rapidly to a limit.

That b is well below a may be guessed by looking at large corporations. A company employing in the neighborhood of 100,000 workers has approximately 6 levels in its administrative hierarchy: foreman, engineer, manager, general manager, vice-president, president. The average "span of control" is then 6.8. But the salary of a president is 100 rather than 100,000 times that of a worker, yielding an average value of $b = 2.5$.

With these values let us now estimate the size of the term by which the incremental cost per worker differs from constancy. We have

$$\frac{(b-1)}{(a-1)} \cdot \left(\frac{b}{a}\right)^{n+1} = \frac{1.5}{1.8} \left(\frac{2.5}{5.8}\right)^{n+1} \doteq 0.258\,(0.368)^{n+1} =$$

1.28 per cent for $n = 2$.

1. L. Urwick, "Axioms of Organization," *Public Administration Magazine* (London) 1935, pp. 348–49, as quoted by Koontz and O'Donnell, *op. cit.*, p. 88.

For medium- and large-sized firms administrative cost per worker is thus very close to a constant, if the assumed values of the span of control and the salary progression are at all realistic. For these values the "wage-multiplier" turns out to be

$$\frac{1}{1-\frac{b}{a}} = 1.74.$$

3. Let us consider also the matter of delays incurred in decision-making on the assumption that each administrator makes an equal number of decisions. This if anything is an unfavorable assumption because it tends to overstate the amount of work assigned to the top levels. We are interested in the average delays to decisions. The number of decisions made at each level is proportional to the number of executives at that level. Thus the proportion of decisions that reach the presidential level is 1 divided by the size of the administrative staff,

$$\frac{1}{\frac{a^{n+1}-1}{a-1}} = \frac{a-1}{a^{n+1}-1} \doteq \frac{a-1}{a^{n+1}}.$$

Let $\bar{t}$ be the average time for passing on a decision. Then, allowing for the possibility that a task may arise at the production level, be decided by the president, and then handed back down again to the production level, the *maximal* average delay is $2\,n\bar{t}$. Now at the second highest level a times this number of decisions are made with a maximal average delay of $2(n-1)\bar{t}$, etc. The average delay of all decisions is thus no more than

$$2\,\bar{t}\,\frac{\sum_{m=0}^{n} m\left(\frac{1}{a}\right)^{m}}{\sum_{m=0}^{n}\left(\frac{1}{a}\right)^{m}} = 2\,\bar{t}\left(\frac{1}{a-1} - \frac{n+1}{a(a^{n+1}-1)}\right) < \frac{2\,\bar{t}}{a-1}.$$

Not only is the average delay limited independent of the size of the organization, but for medium- and large-sized firms it is very nearly constant, the term $\frac{n+1}{a(a^{n+1}-1)}$ being less than 1 per cent for $a = 6$ and $n \geqq 1$. The bound $\frac{2\,t}{a-1}$ can be shown to apply to any organization for which the number of subordinates per person is never less than a constant a.

To sum up the argument: It is by no means true that marginal cost of administration increases indefinitely with size, rather it is bounded at a level only a few times that of workers' wages. It is furthermore possible that marginal cost of administration per worker is practically constant, and this requires no more than that $\frac{b}{a}$, that the ratio of the salary scale factor to the span of control be (practically) constant. If it is granted that salaries are scaled in a constant ratio — and some empirical evidence points in that direction[2] (with the possible exception of the president's compensation which really is in a separate category but again may be in an essentially constant proportion to the size of the firm) — then the problem of the returns to scale in administration revolves on the behavior of the span of control. Specifically, diminishing returns to scale would result only when the average span of control declines systematically with an increase in the number of administrative levels. It might be asked, whether the analysis might be pushed farther by a theoretical examination of factors affecting the span of control. If Parkinson is right, then the total time spent on staff conferences, interoffice memoranda, committees, etc., goes up more than proportionately to size, thus in effect reducing the average span of control. To this an economist might reply that there are real economies of scale via division of labor, specialization of tasks, improved staff work and organization which must reflect themselves in an increasing span of control. Obviously these questions cannot be settled by theoretical arguments alone. What are the facts? We do not know and further research on this point would be illuminating. What we have shown is that diminishing returns are inconsistent with a constant span of control. A constant span of control implies the existence of very nearly constant returns to scale in administration. In industries characterized by constant spans of control and by constant returns to scale in production, firms are equally efficient over an entire range and there is then no unique optimal size of the firm.

Brown University

2. Simon, *Administrative Behavior*, *op. cit.*