Economies of Scale: Some Statistical Evidence

Frederick T. Moore

Quarterly Journal of Economics, Volume 73, Issue 2 (May, 1959), 232-245.

ECONOMIES OF SCALE: SOME STATISTICAL EVIDENCE*

By FREDERICK T. MOORE

I

Statistical evidence bearing on the existence of economies of scale in industry is, for the most part, sketchy and incomplete, although the logic of the economic and technical origins of such economies has been extensively developed. Reasons for this lack of statistical evidence are not hard to find; detailed cost studies of different sizes of plants are a *sine qua non* for analysis of the problem, yet such studies are difficult to obtain. Of necessity engineering information on technical possibilities for substitution among inputs must be combined with the mechanism of choice provided by economic calculations of cost. As Chenery has pointed out, the number of combinations of inputs which may be considered feasible by the engineer is much greater than the number observed in operation and studied by the economist; yet changes in relative prices alone will change the range of economically feasible combinations.[1]

In lieu of deriving production functions from technical data (which is what is actually required), engineers — and in particular chemical engineers — have experimented with various "rules of thumb" for estimating the capital cost of plants of different sizes or for estimating process equipment costs. One such rule of thumb which has found some acceptance is the ".6 factor" rule. The uses claimed and achieved for this rule will be summarized in a moment. Although the engineers do not seem to think of it as shedding light on economies of scale of plant, the rule can be so interpreted and will be discussed from that point of view.

Studies of capital coefficients (i.e., the ratio of capital expendi-

* Paper presented at the American Statistical Association meetings Dec. 29, 1953.

1. H. B. Chenery, "Engineering Production Functions," this *Journal*, LXIII (Nov. 1949).

tures to increases in capacity) by federal government agencies, universities, and others as part of an interindustry research program provide the statistical material for another evaluation of economies of scale. The methodology and results of these studies can be compared with those above.

II

The envelope cost curve usually serves as the vehicle for a discussion of economies of scale; the succession of plant short-run cost curves may trace out a smooth envelope curve or it may be scalloped in various ways. A discussion along this line overlooks the ways in which plant expansions actually take place, however. Expansions of capacity may occur through: the building of completely new plants at new locations; separate new productive facilities (multiple units) which utilize existing overhead facilities such as office buildings, laboratories, etc.; the addition of new productive facilities which are intermingled with the old (the case of "scrambled" facilities); conversions of plants or processes from one product to another; or the elimination of "bottleneck" areas in a plant (the case of "unbalanced" expansion).

It is conceivable that the elimination of bottleneck areas in a plant will increase the capacity by a large amount (e.g., 50 per cent); if that be the case, it is necessarily implied that in other areas of the plant there is excess capacity which can be utilized once the bottleneck is broken. This in turn implies that the productive units in the plant are not divisible, since, if they were, the plant could have been producing the old output with a smaller scale and lower costs. Thus it is usual to attribute economies of scale primarily — if not solely — to the lack of divisibility of productive units. Economies are realized by moving in the direction of larger common denominators of equipment, i.e., where fewer units are operated at less than capacity.

Size of equipment and indivisibilities therein are significant variables for a study of scale, but they do not necessarily go hand in hand. In a copper smelter capacity may be increased by lengthening or widening the reverberatory furnace by small increments (thus increasing its cubic content). This ability to increase the size of a capital input by small amounts exists for a fairly wide selection of industrial equipment; in fact the usefulness of the ".6 rule" is really predicated on this occurrence. It has been noted by engineers that the cost of an item is frequently related to its surface area, while the capacity of the item increases in accordance with its volume. For that reason alone economies in scale may be achieved.

There is another matter which bears on this topic. Chamberlin has argued that it is not only divisibility but the aggregate amounts of inputs used that explain the existence of economies of scale.[2] As size increases, the inputs change qualitatively as well as quantitatively. Different types of inputs are employed at various scales. Changes in quality mean changes in efficiency. The *form* of the input changes as well as the amount. It will not do to call this a question of classification, and to say that the inputs are really distinct. The functions performed by the inputs are the same; the quality changes do not alter the case.

In general it has been our experience in working with files of information on individual plant expansions in a number of industries, that the complementary character of capital goods in a large expansion is quite marked. A large increase in capacity usually involves the plant in expenditures on all productive equipment, not just on selected items. This does not mean that fixed proportions are the rule; flexibility in the use of particular pieces of equipment is common. However, the isoquants probably tend to be more angular and less flat, as they would be in the case of easy substitution between inputs. (N.B. See the case of pipelines below for the opposite case.) Among other reasons, economies of scale arise because the proportions among inputs change as scale of plant changes, although the proportions are variable within certain limits. In other words, the "scale line" may have "kinks" in it as the size of the plant expands. The kinks indicate the points at which quality and quantity changes in inputs alter the proportions in which they tend to be used.

III

The ".6 rule" derived by the engineers is a rough method of measuring increases in capital cost as capacity is expanded. Briefly stated the rule says that the increase in cost is given by the increase in capacity raised to the .6 power. Symbolically,

$$C_2 = C_1 \left(\frac{X_2}{X_1}\right)^{.6}$$

Here C_1 and C_2 are the costs of two pieces of equipment and X_1 and X_2 are their respective capacities.[3] The rule has been adduced from

2. E. H. Chamberlin, "Proportionality, Divisibility and Economies of Scale," this *Journal*, LXII (Feb. 1948).

3. For a description of the .6 rule see R. Williams Jr., " 'Six-Tenths Factor' Aids in Approximating Costs," C. H. Chilton, " 'Six-Tenths Factor' Applies to Complete Plant Cost," both in *Data and Methods for Cost Estimation*, A Collection

the fact that for such items of equipment as tanks, gas holders, columns, compressors, etc., the cost is determined by the amount of materials used in enclosing a given volume, i.e., cost is a function of surface area; while capacity is directly related to the volume of the container. Consider a spherical container. The area varies as the volume to the 2/3 power, or in other language, cost varies as capacity to the 2/3 power. If the container is cylindrical, then, by the same analogy, cost varies as capacity to the .5 power, if the volume is increased by changes in diameter, and if the ratio of height to diameter is kept constant, cost varies as capacity to the 2/3 power. From a consideration of these factors the .6 rule has been developed.

Now consider an alternative and generalized form of the .6 rule

$$E = aC^b$$

where E is capital expenditures, C is capacity and a and b are parameters. So long as $b < 1$, there are economies in capital costs. These economies should not be interpreted as being identical with economies of scale since variable costs must also be considered in the latter case; however, there are some indications that labor, power, and utilities costs also decrease with increased scale while the costs of materials embodied in the final product remain constant. These indications are tentative and not demonstrated by statistical evidence in the cases which follow, so that the ensuing discussion on the evidences of economies of scale must be qualified.

Originally the .6 rule was applied to individual pieces of equipment or processes. A reasonable argument can be made for its validity in those cases; however, the regression line for the formula above cannot be indefinitely extrapolated. There are several reasons for this. In the first place an extrapolation of the line may lead to sizes of equipment which are larger than the standard sizes available or in which stresses beyond the limits of the material are involved. Nelson points out that in building fractionating towers, an economical limit is reached at about 20-foot diameters since beyond that point very heavy beams are necessary in order to keep the trays level.[4] Second, in some industries expansion takes place by a duplication of existing units rather than by an increase in their size, e.g., in aluminum reduction where several pot lines are constructed rather than enlarg-

of Articles from *Chemical Engineering*, 1952. Also see R. S. Aries and R. D. Newton, *Chemical Engineering Cost Estimation*, Chemonomics, 1951; W. L. Nelson, *Cost-imating*, reprints of articles from *The Oil and Gas Journal*.

4. Nelson, *loc. cit.*

ing individual pots. If the rule is to be applied at all it is safest to limit its use to the range of capacities found in the observations.

The .6 rule when applied to complete plants runs into difficulties not encountered on individual equipment. Some expenditures are relatively fixed for large ranges of capacity, for example the utilities system in the plant, the "overhead" facilities, plant transportation, instruments, etc. Complicated industrial machinery does not necessarily exhibit the same relationships between area (cost) and volume (capacity) as do simple structures like tanks and columns. Furthermore, for both items of equipment and complete plants, the gradations between sizes are not necessarily small. Indivisibilities in size are a real factor in some cases; an illustration from the crude pipeline industry will be discussed later.

In spite of these obvious limitations, estimates of the value of b in the formula

$$\log E = \log a + b \log C$$

have been made for a number of industries or products. These estimates are apt to be best for industries: (1) which are continuous-process rather than batch-operation; (2) which are capital-intensive; and (3) in which a homogeneous, standardized product is produced, so that problems of product-mix do not intrude to muddy the definition of capacity. The industries which best meet these criteria are the chemical industries (including petroleum), cement, and the milling, smelting, refining, and rolling and drawing of metals. These are the industries for which statistical estimates of b have been made, and for which some explanation of economies of scale has been supplied.

IV

Chilton has estimated values for b for thirty-six products in the chemical and metal industries.[5] In three cases the value was greater than 1 but in only one of the cases was it so much larger as to be suspect. In the other thirty-three cases the values ranged from .48 to .91. The average value of b was .68 and the median .66, so that Chilton concluded that the .6 rule was reasonable even when extended to complete plants rather than individual pieces of equipment. Some of the values of b which Chilton obtained are shown on page 237. The petroleum industry is well represented in the sample; several processes and one example of complete refineries are shown.

From the point of view of statistical appraisal of these results, it is unfortunate that the error in the regression equation and the

5. Chilton, *loc. cit.*

standard error of *b* are not shown, although from a visual inspection of a few of the products it would appear that the correlations are very high. Nevertheless, it would be valuable to be able to apply a *t*-test

Product	Value of *b*
Magnesium, ferrosilicon process	.62
Aluminum ingot	.90
TNT	1.01
Synthetic ammonia	.81
Styrene	.53
Aviation gasoline	.88
Complete refinery, including catalytic cracking	.75
Catalytic cracking, topping, feed preparation, gas recovery, polymerization	.88
Topping and thermal cracking	.60
Catalytic cracking	.81
Natural gasoline	.51
Thermal cracking	.62
Low-purity oxygen	.47–.59

to the *b*'s to determine, for example, whether they differ significantly from 1. If they do not, the evidence on the existence of economies of scale in those industries would be shaky. It is reasonable to suppose that the values of *b* above .85 (approximately) are perhaps the ones most open to question.

The Harvard Economic Research Project directed by Professor Leontief has made estimates of these "scale factors" for a different selection of chemical products.[6] Their results agree in general with those above, although the range of values found is greater (.2 to an aberrational value of 4.2), and the weighted average for fifteen products is also higher than that found by Chilton. A selection of these values is as follows:

Product	Value of *b*
Aluminum sulfate from bauxite	4.2
Calcium carbide	.8
Carbon black, furnace process	.6
Carbon black, thermal decomposition	.2
Soda ash, Solvay process	.7
Styrene, from benzene and ethylene	.9
Sulfuric acid, contact process	.8
Synthetic rubber, Buna S	1.1

The average for fifteen products (weighted by the U. S. Census values of shipments in 1947) was .8. The scale factors above were computed from very small samples. Of the fifteen products studied, eight were

6. Harvard Economic Research Project, *Capital Coefficients for the Chemical Industry* (hectographed report), May 1952, Table III.

based on two observations; two were based on three observations; two on four observations; one on five and two on six. On the other hand, most of the observations were derived at least in part from engineering data or were checked for type of process and completeness of design and equipment by engineers conversant with the industry. Nevertheless, the results must be viewed with skepticism. Furthermore, even in the cases in which there were the most observations (e.g., carbon black with six plants), the range of variation of equipment costs was considerable; the correlations do not appear to be very high. It is obvious that there are other factors such as location of the plant, product grade, etc., which affect capital expenditures; the data have not been adjusted to account for these factors so that the test of scale is not without ambiguity.[7]

Under contract to the Bureau of Mines, the Petroleum Research Project, Rice Institute, has made a study of capital coefficients for crude oil and natural gas pipelines; one part of this study involved the derivation of a production function for pipelines and an investigation of economies of scale.[8]

The two basic inputs of importance in the construction of a pipeline are the line pipe and the pumping stations, or, more accurately, the amount of hydraulic horsepower. The two inputs may be combined in a variety of ways to achieve any given capacity (which is defined as barrels per day of "throughput"). Any given throughput can be carried by substituting additional horsepower for a certain number of inches of (inside) diameter of pipe. Obviously, a pipe of smaller diameter involves less line pipe costs but also requires more expenditure on horsepower. For example, a throughput of 125,000 barrels per day (60 SUS oil over 1,000 miles) can be obtained by any of the following combinations of pipe and horsepower:

(Outside) Diameter of pipe	Horsepower (approximate)
30	2,000
26	4,000
22	8,500
18	22,500
16	37,500

Other combinations of pipe diameter and hydraulic horsepower can be derived for throughputs greater or less than 125,000 barrels per day.

7. Anne Carter, "Capital Coefficients as Economic Parameters: the Problem of Instability," *Conference on Research in Income and Wealth*, National Bureau of Economic Research, October 1953.

8. L. Cookenboo Jr., "Capital Coefficients for Crude Oil Pipelines and Natural Gas Pipelines," Petroleum Research Project, Rice Institute, Houston, Texas, June 1, 1953 (hectographed).

The isoquants relating diameter of pipe to hydraulic horsepower are of the usual form, convex to the origin, but they are relatively "flat," indicating a fairly easy substitution of these inputs for each other for any given throughput being considered.

Although the isoquants, in generalized form, appear as continuous curves which indicate that substitution possibilities may be considered in incremental amounts, in fact there are discontinuities because pipe comes in standard sizes only. The most commonly used sizes for crude oil trunk lines have (outside) diameters of 8, 10, 12, 14, 16, 18, 20, 24, 26, and 30 inches. Inside diameters have a greater range of variation since wall thickness is also variable, but the number of sizes is not infinite; consequently, there are discontinuities in the production function.

The study of pipelines indicates clearly that economies of scale exist in the industry. Marginal physical product increases up to about 200,000 barrels per day and for larger throughputs the marginal returns appear to be approximately constant. However, because of the discontinuities in the production function the line indicating increasing returns to scale may not cut the isoquants at points representing real alternatives in terms of line pipe size and horsepower. Furthermore, as the size of pumps increases, the cost per horsepower definitely decreases so that although the marginal physical product tends to be constant above 200,000 barrels per day, the capital costs per unit may continue to fall if larger pumps are used. Although there are other costs to be considered, many of them are invariant with respect to throughput and are associated only with the length of the line so that they need not be considered for this problem.

V

Some selected industries in the minerals area have been studied using data obtained from records of plants built during World War II and during the mobilization period beginning in 1950. The records of the Defense Plant Corporation ("Plancors") and of applications of firms for rapid tax amortization ("TA's") contain information on specific expenditures for capital equipment and the increase in capacity which was expected. In order to obtain reasonably homogeneous data, observations selected for study were limited to completely new plants and large "balanced additions." Unbalanced expansions (the elimination of bottlenecks) were eliminated from consideration. This increase in sample homogeneity was thus accomplished at the expense of sample size; small samples were the rule rather than the exception.

However, in partial compensation, each of the plants was studied intensively; the expenditures were classified by type and compared as between plants and processes within plants; in short, every precaution was taken qualitatively to increase the homogeneity of the data. In final form two statistics were presented for each plant: (1) the total capital expenditure (secured as the sum of individual expenditures on equipment and facilities); and (2) the capacity increase secured. These were then correlated using a linear function of the logarithms (i.e., in the form indicated above in this paper). The results in general corroborated those discussed above. In almost all cases the scale factor was less than 1. The industries covered are as follows:

A. *Alumina:* Complete and detailed information was available on only two plants, both using the combination Bayer process for production of alumina; on both of the plants (Baton Rouge, Louisiana and Hurricane Creek, Arkansas) the engineering designs and flow sheets as well as the engineering rated capacities were available. Scale factors for the complete plants and for particular process equipment in the plant were computed.

Plant or Equipment	Values of b
Total plant and equipment	.95
Total equipment	.93
Boiler shop products	.85
Construction and mining machinery	.24
Industrial furnaces and ovens	.98
Pipe and fittings	1.13

The value of b for the total plant corresponds closely to that secured by Harvard. The range of values secured for the process equipment is particularly interesting. The chief machinery complex in the plant exhibits very marked economies of scale, while the value for pipe and fittings indicates diseconomies of scale. It appears that the larger size plant (which has a yearly capacity of 778,000 tons compared with 500,000 tons for the other) can use machinery more efficiently but the connections among the units (piping, etc.) must become substantially more expensive in order, for example, to utilize fully a group of evaporators, mills, or filter presses. An analysis of the engineering flow diagram of the plant tends to confirm this deduction.

It also appears that short-run costs fall as output is expanded. Operating costs, including raw materials, operating labor, allocable share of overhead, and interest on working capital, for the Baton Rouge plant have been estimated for three different levels of output.

Output	Operating Cost ($ton)
1,000 tons/day	$27.28
500 tons/day	29.63
300 tons/day	32.43

B. *Aluminum reduction:* The sample consisted of eight plants comprising a little less than half of the total in existence. Some of the results of the calculations are summarized in the following table.

Item	b	Sy	$\bar{r}$	$\overline{\sigma b}$
Total plant and equipment	.93	.038	.98	.06
Total equipment	.95	.021	.99	.03

A t-test applied to the values of b, testing it against the hypothesis $b = 1$, gave values of 1.17 for total plant and equipment and 1.67 for equipment. Using a 5 per cent critical probability level, neither of the values of b can be regarded as significantly different from 1, so that there is reason for questioning whether these values are really indicative of economies of scale in the industry.

This industry expands by introducing multiple pot lines rather than by an expansion in the size of individual process equipment so that it is possible that the results would be improved if samples stratified according to number of pot lines were used. This suggests, of course, that there is a "lowest common denominator" for total equipment in the plant, and that the equipment is simply duplicated in any expansion, so that economies of scale cease once the lowest common denominator has been reached.

In this industry there are two basic processes of production which are basically similar but which have different capital expenditures in certain process areas. In a pre-baked carbon plant the carbon anodes are manufactured separately and then used in the pots; in Soderberg plants the carbon anodes are continuously replenished in the pot, so that expenditures on pot lines are larger. A Soderberg plant substitutes larger initial costs on equipment for lower operating costs; therefore a consideration of scale necessarily involves an attention to short-run operating costs in deciding on the type of plant to be built.

C. *Aluminum rolling and drawing:* The sample in this industry consisted of four plants making rolled products and four making extrusions. The two types of operations were kept separate in the analysis. The results are summarized below:

Process	b	$\bar{r}$	$\overline{\sigma b}$
Aluminum rolling			
Total plant	.88	.95	.16
Equipment	.81	.93	.18
Aluminum extrusions			
Total plant	1.00	.99	—
Equipment	.92	.97	.13

The *t*-test applied to these results also fails to reveal values of *b* significantly different from 1; however, it is true here as in aluminum reduction, that there are limits to the size of rolls or dies and that multiple units are the usual way in which capacity is expanded.

D. *Cement:* The sample consisted of seven plants with a range in yearly capacity from 450,000 tons to 1,400,000 tons. For total plant the value for *b* was .77 and for equipment 1.06; the former value was not significantly different from 1 according to the *t*-test.

The major variable in the construction of a cement plant is the size (length and diameter) of the kiln. Fuel economy in firing the kilns is a prime objective, since fuel constitutes a large part of operating costs. Kilns and allied furnace equipment may be almost infinitely varied in size; however, since the primary purpose of the kiln is the holding of a cubic charge it was interesting to see if the .6 rule applied to kilns and to allied machinery in the cement plant.

Construction and mining machinery	.60
Furnaces and ovens (including kilns)	.73

These values accord well with the logic of the .6 rule.

E. *Tonnage oxygen:* The sample consisted of five plants ranging in capacity from 50 to 500 tons per day, and producing 95 per cent oxygen. The value for *b* was .63. There are significant changes in capital inputs and costs in one process area (air compression) as scale increases. The major cost item in this area is compressors. For plants of up to 100 tons per day it is most economical to use reciprocating compressors, while between 100–200 tons, there is a choice of reciprocating or centrifugal compressors, and above 300 tons axial flow compressors are more economical.[9] Not only the size, but, more particularly, the character of the capital input changes as the scale increases, and, since the horsepower-hours required per ton decrease as scale increases, there are distinct economies of scale in this process area of the plant. A value of $b = .54$ was computed for compressor costs; this bears out the deductions made from the information on compressor types used in various sizes of plants.

VI

All of the above is but a smattering of evidence on the existence of economies of scale or the lack thereof. From a purely statistical point of view it is discouraging to find no scale factors which test out significantly against the hypothesis of constant returns; yet the

9. "What Price Tonnage Oxygen," *Chemical Engineering*, July 1951, pp. 186–88.

samples are small, and above all it is not clear that a lack of homogeneity in the data does not vitiate the results. These are complex plants usually with a number of process areas. Some areas may be deliberately built with capacities larger than necessary in order to make easier any future expansion. If such is the case, the results are biased.

Although the formula may be applied to complete plants with useful result, it is clear that its application to particular pieces of equipment or process areas is apt to provide better results. The statistical evidence is amply buttressed by engineering information on this point. By adhering strictly to processes rather than complete plants, modifications in the formula can be made to account for individual capacity-cost relationships. For example, although a linear function of the logarithms seems to fit most of the data well, there is some process equipment for which a curvilinear function is required. For equipment such as cyclone separators, centrifugals, and towers, a function which is concave upwards seems to fit the data better. In most cases these curves indicate the existence of economies of scale up to a certain capacity (i.e., slope less than 1) and diseconomies beyond that point (i.e., slope greater than 1); hence an average cost curve for these items would turn up eventually but in general would tend to be flat-bottomed over a considerable range in capacity.

Let us outline a general simple procedure for analyzing the behavior of economies of scale using this process analysis. Suppose that plants in industry X can be divided into four main process areas and one "co-operating" area (e.g., the plant utilities system, piping, or transportation); further let us assume that application of the formula to each area has produced the following values for b_1:

Process area A	.25
Process area B	.60
Process area C	.80; 1.20
Process area D	1.00
Co-operating area E	1.10

From the above it is evident that there are economies of scale in areas A, B, and C, although in the last the economies exist only up to a certain point and then are replaced by diseconomies (e.g., the fractionating tower mentioned previously). Area E contains no possibilities for economies and area D provides constant returns to scale.

It would now be possible to investigate the behavior of economies of scale for different sizes of plant. Eventually the cost curve may turn up. It depends on the importance (from the standpoint of the per cent of total expenditure) of areas C and E. If 75 per cent of total capital expenditures normally occur in area C, or if the per cent of expenditures in that area increases for larger sizes of plants, diseconomies of scale may occur fairly rapidly. If, on the other hand, area A is the most important in the plant, then economies of scale may continue over the whole observable range.

In order to assess the problem we should also know whether the scale of effort in each area can be expanded in small increments or whether the capacities of equipment increase by discrete amounts. In the latter case, economies of scale are limited to specific congeries of equipment. The qualitative characteristics of the equipment must also be investigated since proportions may be affected thereby.

This would appear to be a relatively simple method of analyzing economies of scale in industry and one which is capable of use without an elaborate study of production functions. The engineers have compiled a good bit of information which can be used immediately and catalogues of equipment can provide more. This information is not in the form which can be used directly; usually it specifies the cost of an item which can perform a certain job such as grinding a certain number of tons a day, or conveying a certain charge per hour, etc. But these data can be utilized with only small changes; three steps are normally involved:

(1) The engineering data in technical journals and catalogues give cost relative to some engineering or physical magnitude (e.g., diameter of tank, square feet of heating surface, peripheral area, etc.).

(2) The physical or engineering magnitude can be related to capacity by an appropriate formula (e.g., the capacity of a tank can be related to the diameter). Chenery has suggested some ways this can be done for whole processes,[1] but what is suggested here is on a much simpler level; it may involve nothing more than an application of simple formulae of area and volume, for example. Of course, in the process some of the elements may be omitted but rough justice can usually be done to the relationship.

(3) From (1) and (2) it is then possible to express the relationship between cost and capacity and to analyze the behavior of economies of scale.

It would be interesting to apply this procedure, process by process, to plants in several industries, to go through, in short, a simplified version of design of a plant including an analysis of the changes to be made in equipment as size varies. It would not be necessary to consider the whole range of substitutions among capital inputs

1. Chenery, *op. cit.*

which are possible; sufficient indications of economies of scale could be obtained from perhaps three or four typical sizes, so that the amount of analysis necessary would be smaller than for a complete production function analysis. It is hoped that others may find in this method much to commend as a simple procedure for evaluating the evidences of economies of scale.

THE RAND CORPORATION
WASHINGTON, D. C.