Cost Analysis and Cost Forecasting in the Bank

This paper attempts to derive cost forecasting from an articulate method of cost analysis which is applied by some large banks. The discussion is of a mathematical nature. However, it will be carried out in a literary fashion as long as the mathematical framework will not be considered essential. In the second part of the paper the whole line of reasoning will be restated in mathematical terms and will be continued into some aspects of cost forecasting.

* * *

1. — Cost analysis has two basic features: (1) distribution of direct and overhead costs to the departments and divisions of the bank, *e.g.*, Loans, Individual, Accounts, Trust (corporate, coupon paying, custody, investments, etc.), Foreign (commercial credits, outward collections, etc.), Branches, Comptroller, Bookkeeping, Tabulating, Research, Personnel, Telephone, etc., and (2) computation of the unit cost of each function of the bank (commercial loans, consumer credit, checking accounts, letters of credit, documentary collections, remittances, etc.).

(1) Cost distribution involves two steps: (a) imputation of the direct expenses of each department, and (b) distribution of overhead costs and computation of the net, or final, costs of each department. (a) The first step involves a rather simple computational procedure: for instance, the distribution to the departments of the expenses incurred for their account by a central paying office. (b) The second step involves a quite laborious computation and for this reason few institutions go into a detailed analysis of cost distribution. Indeed, each department may debit expenses to other departments so that, in general, each department is also debited expenses by other departments. This is stated by an equation for each department, which reads: « total » costs are equal to direct costs plus the expenses debited by other departments. The net, or final, costs are simply that percentage of the total costs which the department does not debit to the other departments. The difficulty of the computation arises from the fact that total costs are found by solving a system of the (linear) equations described above. This computation may require perhaps 1000 man-hours — with the aid of desk computers — for a bank with 150 departments, depending on the number of charges among the departments. Of course, the number of calculations increases more than in proportion to the number of departments. Mechanical methods of computation have been devised to cut down the amount of labor involved: one employing punched-card machines, another by the use of a large electronic digital computer.

As an illustration of the computation of total and net costs, let us consider three departments, the initial costs of which are 100, 200, and 300. The first department debits 40% and 60% of its total costs respectively to the second and third departments; the second department debits 10% of its total costs to both the others; the third department debits 20% and 30% of its total costs to the first and second departments. Denoting the total costs of the first department by x_1, etc., we write the above equations,

Department I $\quad x_1 = 100 + 0.1\ x_2 + 0.2\ x_3$
Department II $\quad x_2 = 200 + 0.4\ x_1 + 0.3\ x_3$
Department III $\quad x_3 = 300 + 0.6\ x_1 + 0.1\ x_2$

The reader may easily verify that the solution for the total costs is

$$x_1 = 242.35$$
$$x_2 = 443.88$$
$$x_3 = 489.80$$

The net costs are given by
Department I $\quad (1 - 0.4 - 0.6)\ 242.35 = 0$
Department II $\quad (1 - 0.1 - 0.1)\ 443.88 = 355.10$
Department III $\quad (1 - 0.2 - 0.3)\ 489.80 = 244.90$
The first department, which debits all its expenses to the others, is a service department.

(2) The computation of the unit cost of each function is made on the basis of the net costs by breaking down the net cost of each department among the functions performed by that department, and dividing the total cost of each function, thus derived, by the level of activity of that function. Only operating departments — *e.g.* Commercial Banking, Loans, Individual Accounts, Bond, Trust, Foreign, Branches, etc. — have net costs because service departments — *e.g.* Comptroller, Bookkeeping, New Business Promotion, Research, Personnel, Purchasing, Mail, etc. — debit all their costs to the other departments. The number of operating departments — and therefore of net costs — is usually much smaller than the total number of departments — and of total costs. Furthermore, the number of functions is generally smaller than the number of operating departments. Therefore, while the estimation of the percentage of the net costs that are to be imputed to each function may require laborious time studies and other detailed analyses of the operations, the number of such ratios and the amount of work involved in the computation of unit costs are usually not large.

As an example, let us assume there are five operating departments — Loans, Consumer Credit, Individual Accounts, Branch I, Branch II — which have the following net costs for a given year: 250, 100, 150, 200, 50. We further assume that the bank has three functions: commercial loans, personal loans, checking accounts. The first, second and third departments impute all their expenses respectively to the first, second and third functions. The fourth department imputes respectively 33%, 27%, 40% to the functions; the fifth department imputes 30%, 20%, 50%. The total distribution of net costs among the functions will be given by

Department		I	II	III	IV	V
Function I	331 =	1.00 (250) +	0 (100) +	0 (150) +	.33 (200) +	.30 (50)
Function II	164 =	0 (250) +	1.00 (100) +	0 (150) +	.27 (200) +	.20 (50)
Function III	255 =	0 (250) +	0 (100) +	1.00 (150) +	.40 (200) +	.50 (50)

The total costs of the functions are 331, 164, 255. To obtain unit costs, the total costs of each function is divided by the number of operations of that function in the same period of time.

2. — The method of cost forecasting that is described here aims at deriving the relations between departmental net costs and levels of activity of the functions of the bank from the relations between the costs of these functions and departmental net costs. This is done on pp. 34, 35 in the mathematical section. In general such relationships are so complicated by scale economies, varying degree of utilization of large minimum units of some factors, limitational factors and complementarity of functions and factors, that we can only determine one point of the former relations with each observation. Many observation at different points would then give us an approximate description of these relations much in the same way as in the game of drawing a picture by connecting numbered points by lines. The counterpart to connecting points by lines in the game is found, in the analysis, in the statistical fitting of straight lines to the observed points. Unlike the analogy of the game, the statistical procedure enables us also to extrapolate the observed data, i.e. to describe the relations in regions outside the observed point. The statistical procedure is explained on p. 35.

More specifically, it is the purpose of this method of cost forecasting to derive the ratios of net costs to the level of activity of each function, and at least theoretically the whole technology associated with them, from the results of the computation of unit costs, *i.e.*

from the ratios of the total costs of the functions to net costs, at least for a certain range of levels of activity. In general the latter ratios vary together with all the levels of activity because of scale economies, limitational factors, etc. mentioned above. Consequently, also the former ratios vary, in general. The relations of these ratios to the levels of activity are approximated by fitting regression lines to the observed points by the standard statistical method of least squares.

The structure of the above relationships is given by technical conditions. Therefore, for a better understanding of the results, it may be useful to associate the ratios of net costs to levels of activity with certain techniques of production. The identification of the techniques involved may usually be made from a general knowledge of the operations of the bank. In some cases detailed operational research might be necessary. Needless to say, the larger is the technology, or set of possible techniques, and the more frequent are coexisting techniques and complementarity, the more information is needed from cost analysis for cost forecasting. Cost analysis yields more information when it is made at relatively short time intervals and when the levels of activity of the functions vary widely.

Some examples of the technical conditions determining the structure of the cost relationships may clarify this point. An example of scale economies may be found in the fact that centralized operations are usually cheaper than operations performed by the branches. Similarly, an increase in the activity of a department with respect to some functions might make possible a more efficient use of the existing facilities and consequently lower the costs of other functions performed by the department. Punched-card equipment is clearly a factor with a large minimum unit: when the volume of activity of a bank reaches the level at which it is convenient to install such equipment, this new technique will in general make possible lower unit costs. On the contrary, limitational factors — like office space, or the availability of skilled and experienced labor — may mark a region, or several regions, of activity levels with higher unit costs. Limitational factors illustrate also the possibility of more than one technique being applied to the same function at the same time. For instance, if two techniques are known, one of which is more economical under a certain set of salaries and prices but uses such large quantity of a limitational factor — of which a certain maximum amount is available — that a given level of activity cannot be reached, it may be expedient to use this technique in concurrence with the less economical one, which however uses a proportionately smaller amount of the limitational factor.

The relationships of departmental net costs to levels of activity of the functions — reconstructed from cost analyses, or « observations », over a period of several years, by the method previously explained — may be used for several purposes, among which we may mention: (1) to forecast net costs of the departments and unit costs of the functions for different levels of activity, (2) to detect variations in the efficiency of departments at present as well as in the past, corresponding to changing levels of activity.

A complication in the analysis arises from variations in salaries, prices and rentals of goods and services used by the bank. In this connection a distinction must be made between the direct and indirect effects of these variations. The direct effect is simply that change in the expenditures of the departments, which is caused by variations in salaries and prices under the assumption that after these variations the departments use the same quantities of different kinds of labor and of goods as before. Of course this is a rather restrictive assumption. In fact, if more than one technique, using different proportions of kinds of labor, goods and services, may be applied to the same function, it may very well happen than an increase in salaries and prices of labor and goods (*e.g.* salaries of bookkeepers) largely used by the technique, or combination of techniques, presently applied (*e.g.* bookkeeping by manual machines) would make these techniques less economical than others (*e.g.* semi automatic accounting) which use large quantities of factors that have now become comparatively cheaper (*e.g.* punched-card equipment). As a consequence, to continue to operate on the most efficient keel the bank would have to switch to the more economical techniques. Variations in departmental expenses caused by these technical changes are the indirect effects of the variations in salaries and prices. Both effects of these variations have to be measured in order to bare the relationships of net costs to levels of activities on which our cost forecasting is based. The mathematical framework of these interactions and a method for the measurement of direct and indirect effects will be developed on p. 36.

* * *

3. — In the mathematical discussion we make use of matrix notation for simplicity in writing. A capital letter denotes a matrix, and a small letter represents a column vector (or one-column matrix). A small letter with subscripts indicates a term of a matrix or a component of a vector.

The gross total costs of the departments, v, are given by

$$u = E\,v,$$

where u are the direct or initial costs of the departments and E is the matrix of the percentages of gross total costs debited by each department to the others,

$$E = \begin{bmatrix} 1 & -e_{12} & -e_{1n} \\ -e_{21} & 1 & -e_{2n} \\ -e_{n1} & -e_{n2} & 1 \end{bmatrix}$$

The solution for the gross total costs, v, is of course

$$v = E^{-1}\,u.$$

E has the properties that $e_{ij} \geq 0$ and

$$1 - \sum_{i \neq j} e_{ij} \geq 0.$$

These properties ensure a solution for $v > 0$. The net costs are

$$y_j = \left(1 - \sum_{i \neq j} e_{ij}\right) v_j \geq 0.$$

The equality will hold for the service departments, the inequality for the operating departments.

The cost of each function of the bank (commercial loans, checking accounts, etc.) is derived from the net costs of the departments by a linear transformation

$$z = B\,y.$$

The dimension of z is generally smaller than the dimension of y; B is usually therefore, a rectangular matrix with more columns than rows.

4. — This is just a restatement of the framework of cost analysis as it is applied by some large banks. The method of cost forecasting that will be developed here aims at deriving the matrix of the (constant or variable) ratios of y to the levels of activity of the functions of the bank, from the B-matrix observed for different sets of levels of activity at different times. In general, the B-matrix varies with the levels of activity, *i.e.* $B = B(x)$, because each department will impute different percentages of its expenses to the various functions, depending on the volumes of activity of the functions. One observation of the B-matrix presumably defines an efficient point on the « production function », at certain salaries and prices. A number of observations at different levels of activities, defining as many distinct efficient points on the « production function », help to reconstruct such efficient production function by a more or less close approximation.

The method suggested here for such interpolation is devised to permit a certain elasticity of the model, which may yield a better fit to the real relationships, and to make the most out of an usually small number of observations which often, moreover, contain errors of estimates. In fact, the observations are usually few because they are made once a year and the series studies should not go back in time so far as to include important changes in technology, defined as the set of known banking techniques. Furthermore, the B-matrix is inferred by time-studies, or estimated on the basis of experience, knowledge of services, etc. with an inevitable margin of error.

As a reasonable approximation it may be assumed that the levels of activity of the functions of the bank are independent variables (*). It is then clear that the expenses — reflecting the volumes of work — of the departments cannot vary freely. The variations in the department expenses are restricted to a subspace of at most the same dimension as the space of the functions — which generally are smaller in number. The net costs of the departments are related to the levels of activity of the functions by banking techniques. The techniques are constant vectors, because, if all factors producing a function by a certain technique are taken into account, a proportionate increase in the quantities of these factors should produce a proportionate increase in the level of activity of the function. Needless to say, varying combinations of such techniques make possible variations in the proportions among the quantities of factors. These combinations describe external economies, large minimum units of a factor, limitational factors, etc. already discussed.

(1) We now assume that the techniques used and also salaries and prices do not change. Denoting by x the levels of activity of the functions, as a first approximation we may consequently express the department net costs — which are proportionate to the volumes of work of the departments under constant salaries and prices — as linear functions of x, *i.e.*

$$y = A\,x.$$

The total costs of the functions, z, do not appear explicitly because they are simply given by the equations

$$z_i = \sum_j a_{ji}\,x_i.$$

The terms $\sum_j a_{ji}$ are, of course, the unit costs of x_i. If the A-matrices were constant, they would also be constant and this assumption would impose too strong a restriction on the analysis. Furthermore, inspection of the expressions below shows that if A is a constant such is also B. However, in most cases B is a variable. As a better approximation, therefore, A may be expressed as a linear function of x, because the combination of techniques used will depend on the levels of activity.

We then have,

$$z = B(x)y,\ y = A(x)x,\ z_i = \sum_j a_{ji}(x)x_i.$$

Therefore,

$$z = B\,A\,x = C\,x, \text{ and } B\,A = C,$$

where

$$C = \left[\begin{array}{cc|c} \sum_j a_{j1} & 0 & 0 \\ 0 & \sum_j a_{j2} & 0 \\ \hline 0 & 0 & \sum_j a_{jn} \end{array}\right] = \left[\frac{z_i}{x_i}\right]$$

We want to find A. It is possible to solve for A only if the expression $BA = C$ provides a number of terms in A. This is the case in general only if these matrices are non-singular, that is if costs of the departments are summed in as many groups as are the functions of the bank. The criterion followed in this grouping depends on the particular features and purposes of the analysis. For example, costs of departments which are to be more fully analyzed may not be added together, or a geographical criterion may be used in grouping, etc. This is, of course, a severe limitation on the analysis. However, it may be partially offset by an adequate criterion in grouping. Using the same notation with a bar to denote the above terms after this transformation, the aggregation of the costs of three departments — for instance — will be clearly given by

$$\bar{y}_l = y_r + y_s + y_t,$$

$$\bar{b}_l = \frac{b_r\,y_r + b_s\,y_s + b_t\,y_t}{y_r + y_s + y_t},$$

(*) Actually they partially depend on past expenses for building up the business (on advertising, public relations, etc.). In a more refined analysis, such expenses should be treated as independent variables; the relationships between past expenses and present levels of activity would be given by market conditions. The theoretical framework is rather simple. However, its applications represent a significant extension of the scope of the analysis, for they involve (1) identification of these expenses, and (2) market analysis.

$\bar{b}_i$ being the i^{th} column of the matrix $\bar{B}'$ which multiplies $\bar{y}_i$. Consequently,

$$z = By = \overline{B}\overline{y},\ \overline{y} = \overline{A}\,x, \qquad z_j = \sum_j \bar{a}_{ij}\,x_j,$$

and

$$\overline{A} = \overline{B}^{-1}\,C = \overline{B}^{-1}\left[\frac{z_i}{x_i}\right].$$

The computation of the inverse of $\bar{B}$ will be much simpler than the computation of the inverse of E — for the solution for the total costs — because the size of $\bar{B}$ is much smaller. $\bar{A}$ will have to be computed for each observation.

To find a statistical approximation to $\bar{A}\,(x)$ — based on the assumption of random deviations, which seems to be reasonable — a regression line of each term of $\bar{A}$ on x is computed by the method of least squares. These calculations may be facilited by choosing the x_i on which each term of $\bar{A}$ may reasonably depend, not only to save labor, but also to obtain a better fit, because the reliability of the results depends on the number of degrees of freedom, equal to the number of observations minus the number of regression coefficients. In case of uncertainty as to the x_i to which a particular term of A would be related, it might be expedient to compute a regression equation of this term and different sets of x_i, and to compare the sets of regression coefficients thus obtained. The result for any period of constant techniques will be of the form

$$\overline{A}\,(x) = \overline{\overline{A}}_0 + \overline{\overline{A}}_1\,x_1 + \overline{\overline{A}}_2\,x_2 + \ldots + \overline{\overline{A}}_n\,x_n,$$

where $\overline{\overline{A}}$ are matrices of rank equal to or smaller than $\bar{A}$. Standard tests may be applied, determining the level of probability at which the coefficients of $\bar{A}\,(x)$ are significant.

$\bar{A}\,(x)$ gives a forecast of the costs of the various offices and groups of offices and of the costs of the functions for any level of activity of the bank, at least within a certain range. However, it does not give a forecast of the cost of a single department if this cost was summed to others for the computation. If this forecast is required, it may be roughly obtained by supplementing the information of $\bar{A}\,(x)$ with a statistical regression line of the costs of the departments in the group under consideration, y, on the functions of the bank, x. A more rigorous procedure, on the other hand, would be to repeat the computation for $\bar{A}$, regrouping the departments in order to isolate the ones for each there is a specific interest. Obviously, if the computations were repeated once or twice, the whole original matrix $A\,(x)$ could be reconstructed.

(2) The expenses of a department may vary because the techniques used and also the salaries and prices change, beside levels of activities. If at any time the techniques — or methods of work — of the bank undergo important changes, for example because the technology changes, the statistical applications should be made, separately, for the periods before and after these changes. A comparison of the results for the two periods may help in determining the effects of technical changes on costs.

Changes in salaries and prices, if they are large and not proportionate, may be expected to make some unused or little used techniques become more economical than others, bringing about a shift in the techniques applied. Therefore, these changes, beside affecting the costs directly, because salaries and prices form the costs, may have an indirect effect on costs by modifying the proportions of techniques applied. As we pointed out, for instance, if salaries of clerical workers should increase by a large percentage, it might become economical to use new methods of work, substituting mechanical equipment with semi-skilled operators for skilled bookkeepers, tellers, etc. in some departments.

For cost analysis it is essential, of course, to distinguish between variations in net costs caused by changes in levels of activity and by changes in salaries and prices. If salaries and prices changes proportionately, net costs vary in the same proportion, *coeteris paribus*. However, generally in the computation of net costs changes in initial costs caused by variations in salaries and prices are to be kept separate. We write the relationship of initial costs to total costs as follows:

$$u = u' + u'' = E\ v = E(v' + v'')$$
$$u' = E\ v',\ u'' = E\ v'',$$

where the double-primed letters represent the direct effects of changes in salaries and prices on costs with respect to a point of time taken as a basis, $t = O$, and the primed letters are the difference between costs and the direct effects. The solution for total costs is

$$v = E^{-1}\ u,\ v' = E^{-1}\ u',\ v'' = E^{-1}\ u''.$$

Once E^{-1} has been calculated, not much labor is involved in finding v' and v'', beside v. Net costs are then given by the simple expressions

$$y_j = \left(1 - \sum_{i \neq j} e_{ij}\right) v_j \geq O$$

$$y'_j = \left(1 - \sum_{i \neq j} e_{ij}\right) v'_j \geq O$$

$$y''_j = \left(1 - \sum_{i \neq j} e_{ij}\right) v''_j \geq O.$$

The statistical relationships $A\ (x)$ are found between y' and x, *i.e.* between net costs after the direct effects of variations in salaries and prices have been eliminated, and levels of activity.

If large indirect effects of variations in salaries and prices — *i.e.* technical changes — are present any time, they should be treated like technical changes from any other cause, and the statistical applications will be made for the periods before and after the changes, as it was previously explained.

Pietro Castiglioni