

Efficiency and Effectiveness Evaluation in Local Public Services: A Survey of Methodology

KHEM K. DAHAL*

INTRODUCTION

The main aim of this paper is to make a concise survey on potentialities and problems of evaluation methods in local public services. After a brief description of the features of the local service the paper focuses on the specification of "efficiency and effectiveness" concepts within a local public service productive process. Emphasis will be placed on the problems of identification and measurement both from a qualitative and quantitative point of view.

Though methodologies of productivity analysis of public sector were developed during 1950s, the focus on efficiency and dissemination of the methodologies both at scientific and political level were carried out only from mid 1970s. By general agreement the development of micro-economic analysis of public productivity is seen as a consequence of the decreasing reliance on Keynesian economic policy measures based on cuts in budget expenses which has often proved unequal and ineffective. Even if this kind of economic policy is still prevalent in the experience of developing and developed countries, from mid 1970s awareness has been growing of the need for a more disaggregate and specific policy based on the single productive unit.

It is important to note, however, that if the economic reference of this turning point was the theory of "non-market failure", in most of the countries of Europe and Latin America the political and scientific debate had not emphasised the public - private alternative but, rather, the efficient public versus inefficient public one (Borchding et. al:1982).

The decreasing reliance on the level of public expenditure as an easily - controllable tool for economic policy has led to a progressive shift from an exogenous to an endogenous analysis of the public sector directed both to the decisional and the productive process. It is easy to understand how this shift made public sector analysis similar to that for the private sector applying similar tools. As a consequence, in the subsequent section the most relevant elements of a public service productive process will be indicated.

* Mr. Dahal is a lecturer at Central Department of Economics, Tribhuvan University, Kirtipur, Kathmandu, Nepal.

FEATURES OF LOCAL PUBLIC SERVICES

Henceforth, we will analyse primarily local public activities, but it is evident that most of the conclusion will be applicable on the whole bureaucratic activity of the central government. It can be assumed that local public activities are subdivided into two categories; the production of bureaucratic acts on one hand (certificates, licences, judgements, etc.), which seldom are intermediate products but often are external products; and the production of public services on the other hand. Such a classification is clear enough from a logical point of view but the distinction is more difficult in reality.

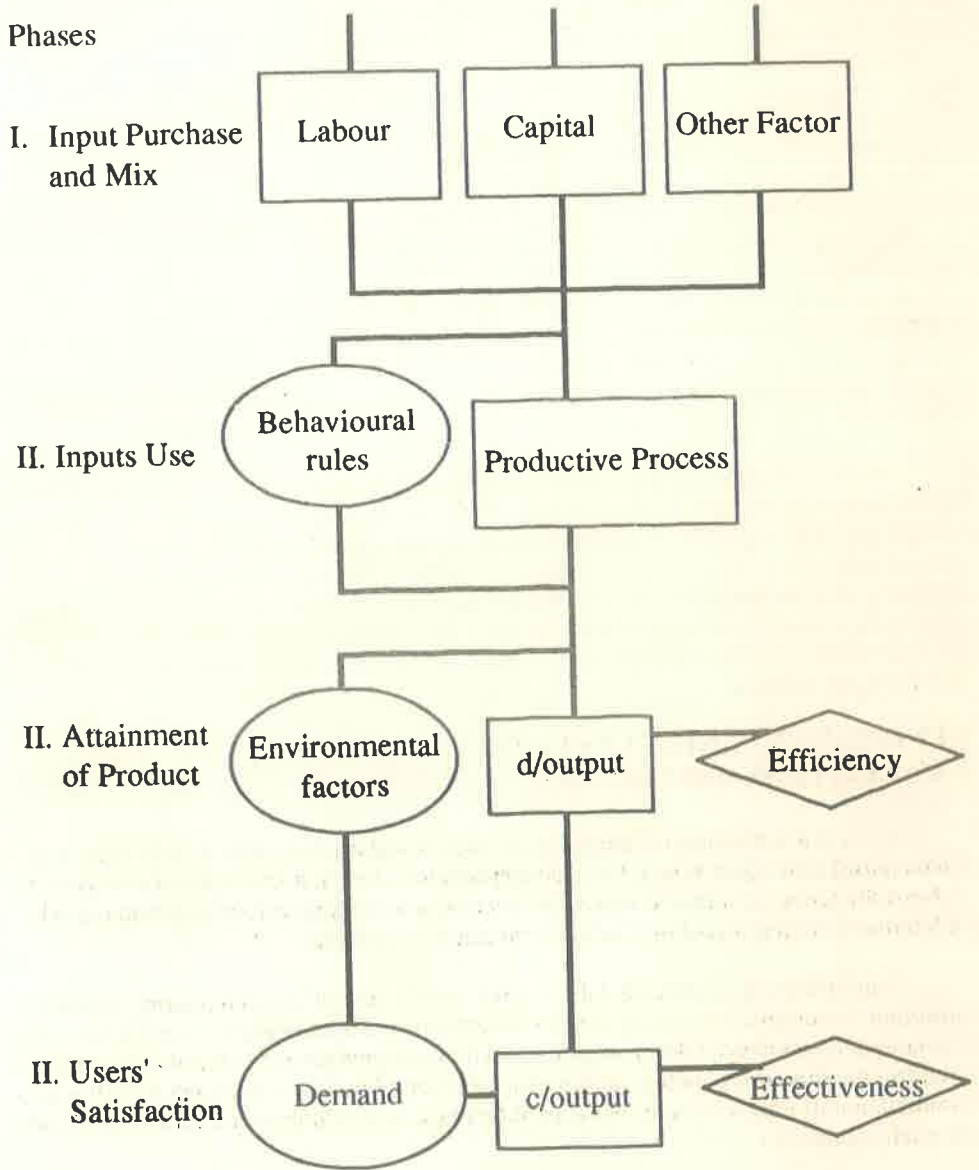
The productive process of bureaucratic activity is quite different from the public service provision and, therefore, the analytical tools should also be specific. In the former case, in fact, we have to do with a bureaucratic activity aimed at providing a pure public good (with the usual non-excludability and indivisibility characteristics), while in the latter, the local public services are usually impure public goods with congestion phenomena, i.e. a decreasing individual benefit corresponds to the increase of the users' number. So far as the bureaucratic activities are concerned, they often have a very low ratio between capital and labour with a very poor technological content. As a result, it is not easy to analyse them by means of the traditional economic tools (like the demand and supply functions), where as the science of organization and the information flows analysis seems more useful for this purpose. For this reason we will pay more attention to the public services to analyse the efficiency and effectiveness.

THE PRODUCTIVE PROCESS OF LOCAL PUBLIC SERVICES

For the analysis of efficiency of public services, it is useful to specify a scheme of their productive process which can make clear the crucial aspects to be examined.

Public Service Productive Process

Figure - 1



It is possible, in fact to distinguish some phases within the overall process; first we have the purchase of inputs in a given mix, the second phase is the use of the inputs. The result of these two phases is what is usually termed *d/output*, that is the technological and/or physical output of the process (e.g. tons of garbage collected, kms. travelled by city buses etc.). But it is clear that individuals do not use the public service directly in these terms and the *d/outputs* do not provide user's utility function. To this end, a further process of conversion from production to consumption is necessary which generates *c/output* consisting in the effects of the public service on the citizen's welfare (e.g. the degree of cleanliness of the streets) (for details see Bradford, Malt and Dates: 1959). The most important aspects determining the results of the process are the elements specified in Figure-1 on the left hand side, that is the behavioural rules of public managers and the environmental factors. Most local public services, in fact, are monopolies, so that it is possible to argue that decision-maker's choices are based often on different criteria from those of private managers. According to public choice theory, the politicians' objectives not always coincide with the voter's ones, since often the former are only trying to be re-elected by the latter. At the same time, the public managers could have different aims, either from politicians or voters (e.g. prestige, power, salary, etc.), which they pursue operating with a relevant degree of freedom; in fact the tools of control either from the top (the monitoring or politicians) or from the bottom (the voice of the users) are insufficient to avoid this kind of distortion.

In this context, we must consider that besides the allocative aim, the government has multiple targets, such as the distributive ones; the reduction of regional inequalities and the increase in employment. The environmental factors constitute in this framework the equivalent of the external economies (or diseconomies) in the private sector, affecting the productivity of municipalities but remaining beyond the reach of intervention. Such factors are sometimes of a physical nature and often depend on the socioeconomic structure. However, their influence affects not only the supply side but also the users' demand.

EFFICIENCY AND EFFECTIVENESS: CONCEPTS AND EVALUATION METHODS

Even if the meaning of the terms efficiency and effectiveness is well known, it seems useful to classify how these concepts can be inserted in the scheme of figure-1 above. Efficiency (in terms of ratio between output and input) refers to *d/output* while effectiveness (in terms of user's satisfaction) refers to *c/output*.

Inefficiency is indicated by a higher cost function in comparison with the minimum. Internal efficiency can be split into a technological component and behavioural component. The former is caused by the imperfect use of productive factors (under-utilization of equipment, absenteeism of employees etc.), the latter by the non-minimization of cost (owing to a capital labour mix which does not take into account the relative prices).

With regard to effectiveness, it is common to distinguish between instrumental and structural effectiveness. According to the first concept the stress is on the qualitative characteristics of output, while in the second one the consequences for the

users are evaluated. Therefore, instrumental effectiveness can be estimated by means of the spread and quality of the public services as well as the ratio of demand to supply (like the frequency of garbage collection or the waiting time for an ambulance), and structural effectiveness can be assessed in terms of effects on users (degree of cleanliness of the streets or reduction in patients death rate).

The evaluation techniques are different and the difficulty of relating the results to the services provided has led to further development of the efficiency analysis methods. However, a satisfactory evaluation of the output is a common problem of both methods, either because most of the services are intangible or because there is more than one output. So, in the case of multiple outputs, we have to adopt method to weigh all products to homogenize them (in terms of labour is the most common way).

Finally, we can distinguish the statistical (analytical) approach, based on index numbers from econometric approach which uses estimates of cost and production function and of demand curve. The first is more useful for intra-unit and time series analyses while the latter is more suitable for inter-unit and cross-section analyses.

THE STATISTICAL APPROACH

The basis of this approach consists in obtaining a series of performance indicator as ratio between output and the factor(s) used in the productive process. Usually, the analyses of the indicator referring to a survey of production units. In many cases these indicators are based on single factors and therefore, they constitute partial productivity indices. The significance of this kind of tool is diminished by its failure to consider the quality of the factor (taking into account the quantity of labour, e.g., we do not evaluate the skill of the worker). Moreover, the increase of productivity may depend on the factor itself or on the mix with other factors or, finally, on technical progress.

In order to overcome this kind of defect attempts have been made to improve the method. On one hand more detailed information has made it possible to take into account different kinds of productive factors; on the other hand overall productivity indices have been developed which are able to measure the increase of total productivity as a rate of output and that of the total inputs. In order to distinguish the effects of technical progress from those of economies of scale, more sophisticated tools were used like the Divisia index (see Diewert : 1976).

However, it is true that the difficulty of attributing the productivity variations to each factor is higher when referring to the technological services; as far as the bureaucratic activities are concerned, the high importance of labour reduce the relevance of this weakness so that even a partial productivity analyses of labour may be useful. The analyses of the Activity Lines widespread in most of the developed countries correspond to this type of analysis. The aim of technique is to identify the causes of inefficiency by means of a very detailed disaggregation of all sectors of public administration into more and more specific elements. The process stops when every elementary sector is identified by producing one (or more) final activity (or activities) which might be considered as a product. Afterwards, by analysing a flow chart of each productive process, the actual use of labour is measured by means of comparison

between standard and real working times for each step of procedure. The final result is a very detailed system of unitary cost indicators.

On the analysis of effectiveness this kind of research is less developed. The only experiences which we can remember are time series or cross-section analysis of indicators based on diffusion of the services (e.g. percentage of served population), quality indicators and opinion surveys' (percentage of citizens who declare themselves satisfied with the services obtained.)

THE ECONOMETRIC APPROACH

In this stream of research, it is possible to evaluate the above maintained asymmetry. As for efficiency, we have several research efforts, both from a methodological and empirical point of view, while, on the contrary the effectiveness analysis is so far at an experimental level.

The efficiency analyses is based upon the estimate of production and cost functions, with particular reference to public services. The underlying hypotheses is that the public body is efficient when, given the technology, it acts on the minimum cost frontier. The efficiency indicators are derived by comprising the actual value of the unitary cost of a given local authority (or agency) and the value estimated by means of the cost function. The great merit of this method is to be able to take into account simultaneously several factors which affect the unitary cost (qualitative aspects, environmental and institutional factors), evaluating the individual contribution of each factor. However, it is not easy to evaluate the minimum cost frontier (or, the maximum production function) and in practice an estimate of the average cost frontier (or the average production function) is employed, selecting a sample of municipalities or of public agencies. According to the duality theorem (see FussMc Fadden: 1978), the effective cost function provides, under certain conditions, all the relevant information about a firms' technology such as the returns to scale, the substitution elasticity between factors and so on (see Petretto: 1987).

Therefore, we can start from an average cost function as follows:

$$C_i = f(Y_i, W_i, A_i, M_i) \quad i = 1, 2, 3, \dots, n$$

with

- C_i = Cost per unit
- Y_i = Output
- W_i = Input prices
- A_i = Environmental factors
- M_i = Quality level

We can assume the ratio $E_i = \frac{C_i}{C_i^x}$ that is the ratio of the observed value to the

estimated value of the cost per unit as an overall efficiency indicator.

Therefore, we can suppose that the public agency or the municipality i is efficient if $E_i < 1$, and inefficient if $E_i > 1$.

In order to separate the technological component from the behavioural one, an estimation of production function is necessary as follows:

$$Y_i = f(L_i, K_i, A_i, M_i) \quad i = 1, 2, 3, \dots, n$$

with

L_i = Labour input

K_i = Capital input

Comparing the observed (Y_i), and estimated output (Y_i^x) we can obtain an indicator

$$E_i^t = \frac{Y_i^x}{Y_i} \text{ of technological efficiency; it is evident that if } E_i^t < 1 \text{ the public}$$

company i is efficient,

while if $e_i^t > 1$ it is inefficient (in relative terms).

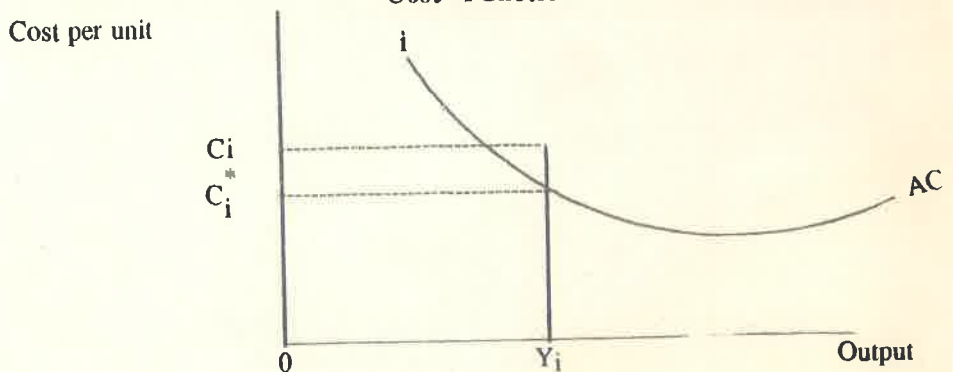
An evaluation of the behavioural efficiency can be made by means of a simultaneous use of the previous indicators. If we assume that the two components of overall efficiency are independent, we can obtain a third indicator $e_i^b = \frac{E_i}{E_i^t}$ of behavioural

efficiency, with

$$E_i = e_i^t \cdot e_i^b$$

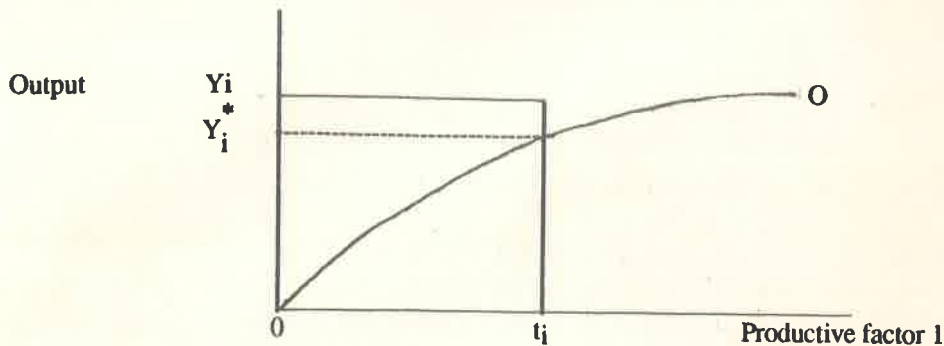
Graphically, this technique can be shown by means of figure 2 and 3. However, in practical application, this method has produced some problems, The most serious one is the lack of statistical significance of some estimated co-efficients, which often show a sing opposite to the expected one, due either to the relation between independent variables or to insufficient quality of data base.

Figure - 2
Cost Function



$$\text{Overall efficiency indicator } E_i = \frac{C_i}{C_i^*}$$

Figure - 3
Production Function



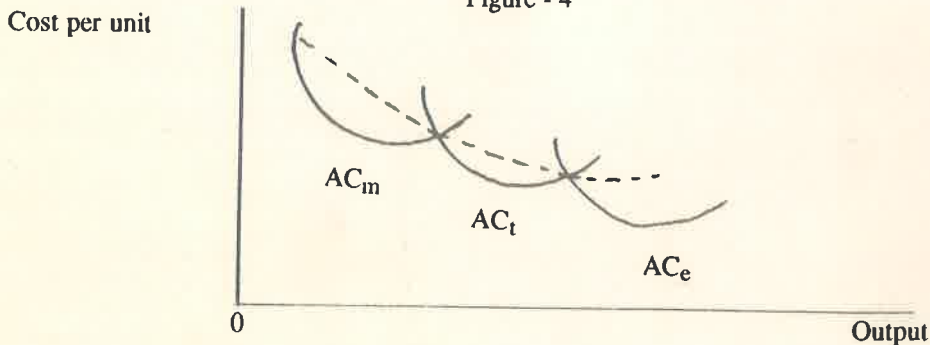
$$\text{Technological efficiency indicator } e_i^t = \frac{Y_i^*}{Y_i}$$

$$\text{Behavioural efficiency indicator } e_i^b = \frac{E_i}{e_i^t}$$

The econometric evaluation of efficiency is able to give useful results by means of the estimated co-efficient of the variable to show, for instance, the form of relation between unitary cost and the degree of inter-substitutability of factors as well as the scale of production.

The economies of scale in local public service are very important in order to suggest institutional reforms. As far as this point is concerned, it is important to analyse, in fact, a cross-section sample of productive units using different techniques at different productive levels may lead to confusion of the economies of scale with a downward shifting of different cost curves (see Figure 4).

Figure - 4



With reference to the effectiveness evaluation we can remark that only the analysis which have tried to make an overall estimate of the whole market of public services should be included in this group. In reality we have to face a *sui generis* market in the sense that the price does not balance the supply and demand often, in fact the market for local public services is a market in disequilibrium with a price fixed by law, lower than the equilibrium price, which causes an excess of demand (corresponding to the difference between potential demand and actual supply). In order to evaluate the potential demand (always unknown) it is possible to use surveys directed to the users to measure the relation between social and economic feature and the price on one hand and the quantity of demand on the other.

SELECTED REFERENCE

- Bradford, D. F., Malt, R.A., Oates, W.E., (1969), "The Rising Cost of Local Public Services: Some Evidence and Reflections", *National Tax Journal*, June.**
- Leibenstein, H. (1966), "Allocative Efficiency Vs. X - Efficiency", *American Economic Review*, 56.**
- Borcherding, T.E., Pommerehne, W. W., Schneider, F. (1982), *Comparing the Efficiency of Private and Public Production: The Evidence From Five Countries*, in Bos, D, Musgrave, R.A., Wiseman, J. (eds.) Public Production, Springer Verlag, Munchen.**
- Diewert, W. E. (1976), "Exact and Superlative Index Numbers", *Journal of Econometrics*, 4.**
- Fuss, M., Mc Faddan, D. (1978), *Production Economics: A Dual Approach to Theory and Application*, North Holland, Amsterdam.**
- Petretto, A. (1987), "Analisi della Produzione e Offerta dei Servizi Pubblici Locali: Aspetti Methodologici" in Maltinti, 6, Petretto, A. (eds.), Finanziamento ed Efficienzadella sepsa Pubblica Locale, Giappichelli, Ed. Torino.**