EVALUATING THE ACADEMIC SYSTEM[']

Rosa Giaimo, Filippa Bono and Domenica Matranga
University of Palermo
Italy

This paper follows two previous studies regarding an analysis of the efficacy and efficiency of the academic system. A bivariate, multilevel model has been proposed in order to measure the relative efficacy of each course by quantifying its contribution in obtaining a particular outcome, net of individual, environmental and course- specific factors. The concept of technical efficiency is also presented and two evaluating methodologies, which are based on a frontier function, are analysed. Both methods take into account differences in students' academic ability (which characterize the university system) and these are analysed on a geographical basis, thereby aiming at an investigation of differentials throughout all the regions in Italy. The results of this analysis will be presented during the conference.

INTRODUCTION

Preliminary studies regarding total quality concern the product/service provided by a particular structure; attention, therefore, has also been given to organisation of that structure. Nowadays, we can discuss the organisational quality of a factory, a bank, a hospital, a school, an entire city, a local office, or the system of a country. The topic of *evaluation* in this paper is dealt with within the sphere of university education, with particular reference to analysing the relative efficacy of courses and the efficiency of services provided by the university. This paper is a continuation and an in-depth study of methodologies which have already been proposed in order to measure two dimensions of quality in the university: efficacy and efficiency.

The evaluation of public services aimed at the individual and, in particular, education, takes on a didactic meaning, whether it be directed towards improving the quality of the service offered or not. The quality of services provided by the university answers a multiplicity of demands whose evaluating cycle in the 'educational system' can be represented by various phases in the evaluation; these phases concern resources, process and result (Cave et al., 1997).

Problems related to the evaluation of educational supply are dealt with in the first part of the work; that part of evaluation which regards the result of the process and which coincides with efficacy is also thoroughly analysed (Aitkin & Longford, 1986). In a previous paper two analyses were carried out in function of different ways of specifying the outcome of the educational process. In the first analysis, the degree class (which measures the outcome) was considered; in the second analysis, the indicator represented by the binary event of obtaining a degree after the duration of the course was chosen and a logit model was adapted. The choice of univariate responses in measuring relative efficacy answered the demand of defining an easy-to-understand model. In this paper we have proposed incorporating two output indicators within a multilevel bivariate model and evaluating if greater methodological complexity corresponds to further improvement in results.

The evaluation of efficiency is generically defined as the utilisation of productive resources at a given level of technology. Having estimated the efficiency rating of Italian universities, our analysis will focus on the specific case of Palermo University. The evaluation of the efficiency of services provided by the university was performed by using DEA methodology, which estimates a function of maximum efficiency by resolving a linear programming problem. The methodologies proposed in this paper will be applied to data published by the National Committee for Evaluating the University System. The results of this analysis will be presented during the conference.

[']This paper is the result of the collaboration displayed by all the authors in the field of MURST (PRIN 98) research, 'Multivariate Analysis and Total Quality: Methodologies, Computational Aspects and Applications'.

THE EVALUATION

The problem of evaluation has assumed an increasingly important and topical role in modern educational systems and these systems are based on the efficacious and efficient management of intervention, be it in the economic and/or social fields. The process of evaluation consists of examining information regarding single components of intervention, which have been gathered in a systematic way according to a measuring process (*assessment*), with the aim of improving the systems' efficacy and efficiency.

With specific reference to efficacy, it is possible to formulate different definitions according to whether the matter under investigation is a *search good* or an *experience good*). Regarding the former, the consumer is capable of identifying its qualitative features prior to the act of consumption, in that quality is relative to *output* and identical for all consumers. In these cases, the level of supply is tightly bound to the qualitative level of the good or service and evaluating efficacy can be explained basically by defining a standard to which efficacy can be compared, but also as a relationship between effective and expected results.

Regarding *experience goods*, their qualitative features cannot be singled out beforehand, nor can they be functionally correlated with the level of the intensity of supply. All types of intervention which are intended to modify behaviour or conditions (such as health or education) can be classified as *experience goods* and their efficacy can be evaluated according to observing the effects on the consumer of the provision of the service. The productive process, put into action by the agent, does not generate a service which is objectively measurable or an output. Instead, there is an *outcome, which is understood as the complex of the effects of supplying the service on the senders of that same service.* The outcome appears, therefore, as the measure of a change generated by the service and perceived in a different way by each consumer in function of his/her background. Consequently, this renders the procedure for comparing different agents as complex because a simple comparison of objective 'results' yields very little utility if the initial positions of consumers are not considered (Fahrmeir & Tutz, G., 1994).

If the intervention, by which the agent's efficacy is evaluated, is extraordinary and specific, it is necessary to evaluate the *absolute or impact efficacy* as a comparison between i) the results which have been effectively obtained by consumers and ii) those which the consumers probably would have been able to obtain in the absence of intervention. Conversely, if the intervention reoccurs, it will be necessary to measure *relative efficacy* and this consists of comparing i) the results obtained by individuals (for whom the service is intended) with ii) those which the same individuals would have probably been able to obtain from an alternative service.

Services provided by the university can be identified as recurring *experience goods* to which the definition of relative efficacy can be applied. The senders of the evaluation are the students (in their capacity as *consumers* of the service) and individual degree courses (which are then organised into faculties) in their capacity as *agents* and *financers*, which are represented by the Ministry and private individuals.

Regarding the standardisation of supply and the aims of the evaluation, it is necessary to consider differences in terms of the average capacity of registered students. Bearing this in mind, the background of the first-year student is important. By 'background' we mean that complex of individual features which contribute in determining unequal conditions of 'competition' between individual degree courses. Individual features include: the class of diploma obtained, variables such as age, wealth, where the student’s family lives, as well as possible jobs carried out during the degree cycle.

Other relevant factors in the evaluation of educational supply are the so-called environmental or contextual factors which are measured by suitable indices such as, for example, size and importance of the university or the attraction of better students. Typical factors about the course can be added to these, the presence of which will make a course appear more efficacious than another because of its organisation, teachers and the accessibility of its teaching. Regarding this complex of factors, students/consumers constitute a hierarchical structure in that they are grouped according to degree courses and faculties.

ANALYSING EFFICACY

In order to analyse the relative efficacy of courses and faculties of a given university, in this paper we have proposed the use of multi-level analysis techniques (Goldstein, 1995)and hierarchical data; both allow us to obtain estimates by means of the combined evaluation of features on the micro-level (students/consumers) and the macro-level (degree courses and faculties). Multi-level methodologies allow us to measure the relative efficacy of every course by quantifying the contribution of factors (specific to the course) to the capacity of obtaining a result; the latter defines the outcome which is net of individual factors and the context (Shadish, W.,1998). Efficacy can be defined as the contribution of each course or faculty to the relationship between the variable outcome response and the explanatory variable, that is, the diploma score. The latter is net of: i) individual factors connected with student characteristics; and ii) a factor which is specific to each course and faculty, given by the teacher ratio per 100 students.

In the specific case of evaluating the relative efficacy of the degree courses of a given university, it is necessary to consider that the data relating to students/consumers is hierarchically arranged, primarily on the course level and then on the faculty level. Indeed, it is not appropriate to maintain that a given degree course is more efficacious than another on the basis of some previously-fixed criterion (continuous or discrete), if the features of the students on the two courses do not coincide exactly or if there exist different socio-environmental conditions (Willms & Raudenbush, 1995).

The idea underlying multi-level methodology considers the observed group as *a* random sample generated by a given population. The relationship, therefore, observed by the unities of the group is not unique although as many relationships may exist as there are higher hierarchical levels, in which case unities are nested. Such a plurality of relationships is merely a random sample generated by the real relationships which exist in the population. The results of a preceding paper (Giaimo & Matranga, 2001), regarding the relationship between the class of diplomas and degrees from Palermo University, lead us to a model with three variance components and eight fixed parameters. However, the faculty effect appears insignificant, notwithstanding the high variability of the teacher ratio per 100 students.

Passing on to the variance component logit model regarding the probability of obtaining a degree beyond its allotted time (a 'physiological' delay of two years in obtaining a degree), the event of *obtaining a degree beyond the legal duration of the course* has been defined as a binary one. This assumes 0 modality when the degree has been obtained with a delay of 0, 1 or 2 years; thereafter, the modality is 1. The preceding results for univariate models can be compared with those deriving from the bivariate model.

ANALYSING EFFICIENCY

The quality of services provided by the university can also be evaluated from the viewpoint of efficiency and this is generically defined as the optimum utilisation of productive resources at a given level of technology. The aim of analysing the productivity or efficiency is generally the quantification of differentials: one efficiency indicator is by definition the result of a comparison between an observed production process and a process which is taken as a reference standard. The information provided by this process has features which are intrinsically relative. On the other hand, a productivity indicator does not provide useful information if considered as absolute, whilst it becomes more interesting if it is compared with analogous indicators referring to production processes, which have been appropriately chosen. In a previous paper (Giaimo & Bono, 2001), an evaluation of the efficiency of the Italian university system was performed on the basis of estimating a mean production function (Johnes, & Taylor, (1990) and estimating a borderline function of maximum production.

Estimating the mean production function has allowed us to compare the performance of universities according to an approach based on the use of indicators. A comparison of performance was carried out on the basis of result, resource, process and contextual indicators, as defined by the Committee for Evaluating the University System. Bearing in mind the results of previous analyses, we have applied a DEA analysis to the efficiency calculation of Palermo University.

DEA

Unlike the usual measurements of relative efficiency which estimate a common set of weights (Farrell & Fieldhouse, 1962) for the totality of the unities considered in analyses, DEA analysis estimates one set of weights for each unit which enables to attain optimum performance, as compared with other unities. The set of efficient units is the so-called comparison group and it defines the borderline of efficiency. The models applied to the DEA methodology can be *input-oriented* or *output-oriented* models, according if inputs must be minimised, given a certain output, or the output must be maximised, maintaining the inputs at a constant level. DEA analysis allows us to identify the potential change which each unit must undergo in order to reach efficiency (Charnes, Cooper, & Rhodes, 1978). Competitive benchmarking is undertaken in relation with the so-called comparison group, which is made up of unities operating in conditions of efficiency (input or output efficiency) for the set of estimated weights (Cooper, Seiford &Tone, 2000). Given two non-specific universities A and B, it is possible to identify the potential change which university A must undergo in order to increase its own efficiency as compared to university B.

Given two non-specific universities A and B, it is possible to identify the potential change which university A must undergo in order to increase its own efficiency as compared to university B (Fig. 1). Thus, for example, considering the equivalent student variable to teachers, it is possible to note that, on the basis of the variables listed in Figure 1, university B performs better by 75% than university A. Such a result is obtained by employing a proportion of lower input (16% less students coming from secondary schools, 44% less financial resources, 43% lower for the teacher ratio per 100 students and 52% lower for technical and administrative variables per teacher). It follows that university A would have to undergo changes (ie, minimizing the resources employed) in an input-oriented model.

REFERENCES

Aitkin M., & Longford, N. (1986), Statistical modelling in school effectiveness studies. *Journal of the Royal Statistical Society, Series A, 149,* 1-43.

Cave, M., et al. (1997), *The use of performance indicators: the challenge of the quality movement.* UK: Higher Education Policy, Kingsley Press.

Charnes A., Cooper W. W., Rhodes E. (1978), Measuring the Efficiency of Decision Makink Units, *European Journal of Operational Research,*2, 429-444.

Cooper W.W., Seiford L.M.& Tone K., (2000), *Data Envelopment Analysis. A Comprehensive Test with Models, Applications,* References and DEA-S Solver Software, KAP, Boston.

Fahrmeir, L., & Tutz, G. (1994). *Multivariate statistical modeling based on generalized linear models.* Springer Verlag.

Farrell, M.J., & Fieldhouse, M. (1962). Estimating Efficient Production functions under increasing Returns to Scale. *Journal of the Royal Statistical Society, 120,* 252-267.

Giaimo, B. & Bono F. (2001), Misura dell'Efficienza produttiva: Un'applicazione al Sistema Universitario, Il Mulino.

Giaimo, M., & Matranga D., Valutazione della Qualità dell'Istruzione Universitaria nell'ottica dell'Efficacia, Il Mulino.

Goldstein, H. (1995). *Multilevel statistical models*. NY: Oxford University Press.

Johnes, J., & Taylor, J. (1990). *Performances indicators in higher education*. The Society for Research into Higher Education & Open University Press.

Shadish, W. (1998), Some evaluation questions, *Practical Assessment, Research &Evaluation*, *6*(3).

Willms, J.D., & Raudenbush, S.W. (1995). The estimation of school effects. *Journal of Educational and Behavioural Statistics*, *20,* 307-335.