DYNAMIC MODELS FOR HIGHER EDUCATION IN VARIOUS SITES

ANNIBAL PARRACHO SANT'ANNA
Department of Production Engineering
Universidade Federal Fluminense
24210-240 Niterói-RJ Brazil

ABSTRACT *The complexity of academic activity, the eminence of the freedom to research and the reverence due to the authority of knowledge require full autonomy for the university departments. Above that, although dependent on central decisions on organization and resources availability, it is the action of Faculty and students within the departments that determines the results of the educational process. For this reason, it is at departmental level that must be driven the main effort to improve quality in higher education and the most important virtue of any academic evaluation system is its power to call attention of the departments to the importance of quality management and providing them information that may drive their efforts in this direction.*

This paper presents an approach to access the productivity of academic departments by combining model building and parameters estimation, in an iterative manner. Validation and outliers selection techniques can then be applied to feed automatic model revision procedures that will preserve continuity in the evaluation process. We start with a simple linear approximation to the ideal production function of each department with coefficients depending on quality of work explanatory variables. We follow the evolution of these coefficients estimates along time, as local observations become available. Departments with extreme estimated coefficients are taken as object of study aiming to detect useful innovation or any other cause of its different observed behavior.

An iterative algorithm for the estimation of the parameters of the dynamic hierarchical model thus generated is described here. An approach to combine the estimates of productivity with respect to each particular output in a global measurement is also proposed. Examples of different ways to model the academic activity are, finally, presented.

INTRODUCTION

Performance evaluation must primarily aim to identify factors of improvement and to arouse pursuit of such improvements. In order to reach such goals, it should pay attention to steady tendencies and structural differences rather than to final results. We present here a performance evaluation system based on modeling a production function with coefficients assumed to vary according to measurable factors that have to do with work quality, management and environment aspects. This allows not only to compute efficiencies on the basis of productivity estimates more reliable than crude ratios of observed cost and production, but also to identify systematic sources of variation and to detect the presence of unknown factors affecting efficiency.

By basing evaluation on fitting and verifying a statistical model for the production function, we provide elements to develop analysis on various levels of complexity. First, we collect data on inputs and outputs and register the academic relevant variables. After that, we use these data to produce aggregate measures of production and to compute estimates of productivity. These estimates are then used to compare efficiencies on the basis of the causes of performance variations. Finally, discrepancies between observed and predicted results are used to select for further analysis observation units with large adjustment residuals.

The initial choice of relevant variables, the development of data collecting reliable procedures and the model check involve all faculty in the quality effort. Measurement of local technical coefficients and of their inner relationships helps people in management position to develop their own efficiency improving and innovation fostering procedures. Selection of units with worse adjustment in order to there search for possible effects of factors left out of the model opens space for different sorts of benchmarking.

Usually, we can collect data on costs and production of each department, as well as on factors differently affecting productivity of its personnel. To such data we can adjust a hierarchical model where in the first level is a production function and, at a second level, the technical coefficients of such production function are themselves explained by a function whose coefficients can also be estimated. But, when there is not enough data about factors that might explain the production coefficients variation, or even when the number of academic departments is small as much as the number of available observations on each department, we may be led to give up identifying individualized local coefficients and then to compare performances on the basis of distances from observed productivity ratios to average or frontier such ratios. Doing so, we implicitly assume that all variation in costs and final result measurements is due to differences in efficiency. Since the larger observed deviations are seldom systematic, this approach results in fingering as the most efficient or most inefficient the units with less reliable data. If there are large errors in the measurements of the units taken as reference, the ratios obtained by comparison with such extreme points reflect no efficiency at all.

If we wish to admit that local parameters change with time, we have only one observation on each observational unit. But we may compare performances on the basis of estimated coefficients, even when dealing with such a small sample on each unit. We may separate in each local parameter a general component and employ data collected in all units in its estimation. And the variation left in the local component may have its evolution through time modeled in such a way as to satisfactorily reduce the number of parameters to estimate. We may, for instance, derive estimates for the parameters variance through time from the scale of the differences between observations on different units.

In synthesis, the approach developed here is based on the hypothesis of academic parameters systematically varying along the departments and through time. Not only a production function relating volumes of critical inputs and outputs but also the dependence of the coefficients of such function on management and environmental factors is, if possible, modeled. Periodically, a search for innovations and new sources of variation is realized in the units with worst adjustment and drives the model revision. With feedback expected from the beginning, initial modeling may be less elaborated, taking in due account that, usually, only after looking at the first results of model adjustment, even the most experienced administrators can give their best contributions to improve modeling.

In the next Section, procedures based on combining local and external information are developed. Then, efficiency criteria developed to compare units on the basis of their productivity estimates are discussed. Finally, models with different complexity are compared.

ESTIMATION OF PRODUCTION COEFFICIENTS

The model structure used here is centered on a linear production function relating the volume of output of a few different products to the global cost involved in producing them. The coefficients of this production function are allowed to vary through time and from unit to unit according to environmental or managerial variables, also observable. A third set of equations explaining these second level coefficients may be added and so on.

Formally, this structure may be given, by:

Observation equation:

$$C_{it} = P_{it}\beta_{it} + \varepsilon_{it},$$

C_{it} denoting observed aggregate cost,

P_{it} denoting the vector of observed values of the outputs that the model considers,

β_{it} denoting the vector of inverse productivities with respect to this set of outputs,

ε_{it} denoting zero mean random disturbance in cost measurement in unit i at time t.

Structural equations:

$$Y_{kit} = F_{kit}\gamma_{kit} + \eta_{kit},\ Y_{lkit} = F_{lkit}\delta_{lkit} + \nu_{lkit}, \text{ and so on.}$$

Y_{kit} denotes an estimator of β_{kit}, the inverse productivity in unit i with respect to output k at time t, F_{kit} denotes a vector of second level observable explanatory variables, Y_{lkit} denotes an estimator of the l-th coordinate γ_{lkit} of the vector γ_{kit}, the vector of second level technical coefficients relative to output k, η_{kit} denotes a zero mean stochastic disturbance for a second level equation, F_{lkit} denotes a vector of third level observable explanatory variables, δ_{lkit} denotes one of the vectors of third level technical coefficients and ν_{lkit} denotes one of the third level zero mean disturbances.

Mean evolution equation:

$$a_{it} = G_{it}.m_{it-1},$$

for G_{it} known, a_{it} and m_{it} denoting the mean of the prior and posterior distribution of the vector of coefficients of last level equations. In many situations, we may simplify estimation by assuming all G_{it} identity and even that the final level coefficients do not depend on the unit i.

Variance evolution equations:

$$\sigma^2_{it} = k^2_{it}.\sigma^2_{it-1},\ V_{it} = J_{it}.W_{it-1}.J_{it}',$$

for k^2_{it} and J_{it} known. σ^2_{it} is the variance of the distribution of φ_{it}, V_{it} is a positive definite matrix such that $\sigma^2_{it}.V_{it} = \mathrm{Var}(\theta_{it} | D_{t-1}, \sigma^2_{it})$, the variance of the prior distribution of θ_{it}, and W_{it} a positive definite matrix such that $\sigma^2_{it}.W_{it} = \mathrm{Var}(\theta_{it} | D_t, \sigma^2_{it})$, the variance of the posterior distribution of θ_{it}, for θ_{it} vector of coefficients of one of the structural equations or of the observational equation and φ_{it} the disturbance of the same equation.

If we cannot find observable explanatory variables, we are forced to take the structural equations out of the model. But we must not drop the index i from the observation equation coefficients unless we are willing to admit identical production equations in all units, case in which differences on units results are only accidental and irrelevant to the goal of deriving quality improvement actions. In this case, of absence of structural equations, the algorithm has no smoothing stage and all we have to do is to initialize, possibly with the estimates of the constant coefficients regression, and, after that, update each unit productivity estimates separately, anytime a vector of new observations is available.

By modeling estimates instead of parameters in the structure equations, we are able to set an independence structure in the dynamic hierarchical model that permits to isolate each level conjugation computations and to infer about the ratios between the disturbances variances. Assuming independence between all different equations disturbances and between coefficients and disturbances, and assuming a normal and inverted gamma distribution for the coefficients and the variance of the disturbance of each equation, after starting with known prior mean and variance for the distribution of θ_1 given σ^2_1 and known prior mean and number of degrees of freedom for σ^2_1, we can easily relate parameters of prior and posterior distributions at time t for the whole set of equations.

This can be done through successive steps of forecasting and updating. Forecasting follows from simple multiplication by the matrices G and successive F and use of the constants k^2 to change variance scales.

For the presentation of the updating relations, a little more notation is needed. Let m_{it} denote the mean of the posterior distribution of θ_{it}, that is $m_{it} = E(\theta_{it} | D_t, \sigma^2_{it})$, and let $a_{it} = E(\theta_{it} | D_{t-1}, \sigma^2_{it})$, the mean of the prior distribution of θ_{it}. Analogously, let n_{it} and l_{it} be the numbers of degrees of freedom of the distributions of σ^2_{it} conditional on the information available at times t and t-1, respectively, denoted by D_t and D_{t-1}. Let $s^2_{it} = E(\sigma^2_{it} | D_t).n_{it}$ and $r^2_{it} = E(\sigma^2_{it} | D_{t-1}).l_{it}$. Finally let e_{it} denote the prediction error resulting from estimation of θ_{it} by a_{it}, that means, $e_{it} = C_{it} - P_{it}.a_{it}$ if θ_{it} is the vector of inverse productivities or $e_{it} = Y_{it} - F_{it}.a_{it}$ if θ_{it} is the vector of coefficients of the structural equation whose left hand is Y_{it}. The updating relations are then given by:

$$W_{it} = V_{it} - (1+F_{it}V_{it}F_{it}')^{-1}V_{it}F_{it}'F_{it}V_{it}$$
$$m_{it} = a_{it} + (1+F_{it}V_{it}F_{it}')^{-1}V_{it}F_{it}'e_{it}$$
$$s_{it} = r_{it} + (1+F_{it}V_{it}F_{it}')^{-1} e_{it}'e_{it}$$
$$n_{it} = l_{it} + 1.$$

While forecasting goes from the last level to the first, updating starts from the first level. The coordinates of the vector of means of the posterior distribution of β_{it} can be used as the Y_{kit} for the second level equations, the second level m_{it} used to compute the prediction errors for the third level updating and so on. Then, each iteration of the estimation procedure starts with updating the first level estimates with the means of the posterior distribution of βit and proceeds with updating until we reach the last level, where we start using the forecasting relations to smoothen the vectors of coefficients estimates.

To complete the description of the estimation algorithm, it rests to show how to initialize it. The means for the initialization priors may be obtained from specialists. These means need not be realistic. Wishful values may be used to impact initial predictions and drive efforts towards desired goals. The distortions thus induced will be corrected by the estimation algorithm in a few runs.

When starting analysis, it is seldom available any information on prior distributions for the variances. Uninformative priors, characterized by large variances and small number of degrees of freedom, can be used in the first run. Instead of that, we may prefer to initialize with the assumption of coefficient estimators uncorrelated and with the same variance of the disturbances. After a few runs, when the set of explanatory variables to be used will probably have been changed according to the new information then received, and the distortion due to the unaccounted correlation between the outputs in the first level equation evaluated, we may reinitialize with improved correlation assumptions.

While using the uncorrelated equal variances hypothesis, in order to avoid the effects of differences in measured variances due to scales of measurement, it may be advisable to standardize all explanatory variables by equalizing scales. Standardization of all dependent variables may also help developing patterns for comparison between different levels residuals.

THE CONCEPT OF AGGREGATE EFFICIENCY

Once obtained final individual estimates allowing for local quality standards, it remains the problem of global comparison of departments with different teaching, research and service structure. A linear combination of the estimated technical coefficients will be the best uni-dimensional productivity measurement, whenever the weights of such linear combination correctly reflect the importance of each output. We choose as weights the volumes of output in the unit whose efficiency is being accessed.

Besides measuring productivity, we must also choose the standard behavior with respect to which to compare the measurements. A general formulation for an efficiency measure is given by the ratio of two linear combinations with the same weights, given by the volumes of outputs produced by the evaluated unit, the first of optimal and the other of local technical coefficients. By this way, local aggregate productivity will be compared with the most efficient observable way of producing the same output attained in the unit under evaluation.

The optimal coefficients may be obtained as the best linear combination of the estimated coefficients of different units. The weights in this linear combination must be those that minimize the cost of producing the outputs P_{jio} of the unit i_o under evaluation, admitted eventual excesses in the production of particular outputs.

Thus, we have the efficiency measure formally given, for unit i_o, by:

$$\text{Efficiency}(i_o) = \inf \Sigma(\lambda_i D_i/D_{io}),$$

this sum evaluated for i varying along the whole set of observed units, and the infimum computed among all possible vectors of contributions λ satisfying $\Sigma(\lambda_i P_{ki}) \geq P_{kio}$, for all k. Here, as before, P_{ki} denotes the volume of output of the k-th product at the i-th university department.

On the other side, D_i is introduced here to denote the corrected cost for unit i, obtained through the prediction generated according to the chosen model, and not necessarily the observed costs.

The distances from predicted to observed costs should also be examined, in order to look for efficiency differences due to factors still not taken into account within the model. But, while these factor are not suitably measured, we should not intend to derive any efficiency measure from their effects, specially when mixed with purely random disturbances. If such factors seem important, we ought to first identify them through the *in loco* analysis of the units where the known factors are less able to explain the measurements obtained.

The rule for the choice of units for *in loco* analysis should be simple and easy to explain. For instance, choose the two units with higher predicted to observed ratios among those not selected in the two last years. The analysis of the units then selected should take into account the coefficients estimates obtained in the successive model levels for that unit and also for the whole set. For instance, explanatory variables previously considered relevant but excluded from the model adjusted due to low estimates should there have their measurements carefully checked. But the main search should be for elements not yet considered: quality improvements on teaching or development of new research procedures may explain positive residuals in the adjustment, as much as simple waste due to lavishness in the application of resources; on the other side, negative residuals may be due to loss of quality as well as to technological advances enhancing productivity. Substantive improvement in the quality management will come from measuring these possible changes and taking them into account in the production function.

COMPARING DEPARTMENTS OF VARIOUS SIZES

In [1], we fit a model to academic production, with only two first level explanatory variables: Teaching, measured by the total number of hours of students enrollment in courses, and Advising, measured by the number of final dissertations advised. The input variable is the value spent by the University in paying Faculty. As second level variables, enter two uncorrelated quality enhancing factors for each of the first level productivity coefficients.

The teaching coefficient is supposed to increase with a measure of students satisfaction and a measure of reliability of the evaluation procedures used by teachers. The advising coefficient is allowed to grow with a measure of the Department Faculty presence in community projects and with the number of research papers published. Positive signs in the coefficients of these second level variables reflect the contention that the rates of resource utilization increase with the quality of the education offered to students. The model presents a third level of explanation, where these coefficients are modeled as been reduced by the presence of management resources, that would make easier to the Faculty to offer better service. Other explanatory variables might appear at this level to take into account specific aspects of the production process of departments belonging to different areas.

Another philosophical approach would separate the results of Teaching and Research. We might also add a third explanatory variable measuring the amount of Extension services. The interaction between the activities leading to final products measured under each of these three headings may difficult the estimation. On the other side, it may be useful to fit such a model to evaluate the effect on productivity of concentrating on each of these fields.

The production variables for this second model will not be so straightforward as those in the first model. In order to choose relevant and reliable observable variables and weights to combine them into measurements of Teaching, Research and Extension we must involve specialists of the different areas of knowledge. Tools for quantifying preferences should be developed to deal with this problem in such a way as to serve also the goal of calling attention for quality management aspects. Frequent revision of weights should be exploited with this same purpose.

The adjustment of a structural model with the parameters of the first level equation modeled to explore the possibility of being specialization the reason of technological or managerial improvements that reflect on productivity can be compared with that of an unstructured model. In order to take into account the operational scale in terms of each output, that means specialization, we can choose as second level explanatory variables indicators like the ratios input/output for each output. These explanatory variables may be used to explain the respective production function coefficients or may be combined in a new factor used to explain any of the first level coefficients.

FINAL COMMENTS

Only after modeling academic production as varying systematically from one university department to another and then explicitly modeling efficiency, we are able to compare performances in an instructive way. If it is verified that a structured model fits the data, their technical coefficients not only set a goal towards which the other departments should move, but also hint on ways to reach such goal.

Assuming productivity coefficients varying through observational units we can estimate local parameters even when the observations are subject to small disturbances. A hierarchical model built by allowing the production coefficients to vary according to second level observable factors gives, if not the best efficiency estimates, the best way to analyze the differences in efficiency.

On the other side, just detecting unexplained components, constant coefficients regression residuals hint on units that should be objet of further analysis but do not provide a suitable basis to access efficiency. Observed to average or regression predicted values ratios can offer sensible efficiency measurements in the case of two groups of production units similarly sprayed on the production space, one of them less efficient than the other. But, if efficiency extremes are in a very small set of units, the predictions obtained by fitting a global model should be expected to be distorted. In that case, even from the point of view of selecting observation units to deeper analysis, fitting a different equation for each observation unit is more informative.

The first estimates of varying coefficients will be based on a relatively small data set and shall not be very reliable. But, in an iterative modeling and estimation approach, we are always improving on them. And deriving model improvement from local analysis of units sampled by bad fit has the advantage of attracting specialists in the departments to think about quality and to act on it.

REFERENCES

[1] Sant’Anna, A. P. **-** “Modelagem da Produtividade Iterativa com a Avaliação do Desempenho”, *Pesquisa Operacional*, 17, 1, 1997, 71-86.