STATISTICA, anno LXVI, n. 3, 2006

MODELLING UNIVERSITY HUMAN CAPITAL FORMATION AND MEASURING ITS EFFICIENCY. EVIDENCE FROM FLORENCE UNIVERSITY

Guido Ferrari

1. INTRODUCTION

Human capital (HC) is a concept to which social sciences scholars and researchers are interested since at least a couple of centuries.

The debate that has developed around its definition and measurement has always remained alive, by feeding on new theories and interpretations, as well as on evidence, and still is topical. Indeed, it represents a question at which, besides the traditional, so to say, interest of economists and sociologists, points since long time the attention of statisticians and econometricians.[1]

HC formation, and the related activity, similarly to what occurs as regards the physical capital stock formation, is a question, if possible, even more topical, not to say that it represents one of the points on which it is necessary to think over more, due to its implications on the economic development and on the individual welfare.

In order to try to approach the matter in a general and, in the meantime, useful way to the developments that I wish to give to the paper in what follows, let's begin with the reflexes on the economic development, i.e., from the macro side, even though the title might drive at once to think of micro behaviours.

I realize that I'm approaching things a bit from afar; but I guess it is useful to clearly frame, although summarily at all, the rather technical treatment I intend to develop later on.

Let's start from globalization. Either we like it or not – and I admit that the ele-

[1] HC is commonly defined as the increasing in education and attitude to work due to education, formation directly received at work and improvement of health and psycho-physical conditions (see, for example, Vittadini *et al.*, 2003a; 2003b, also for a historical-systematic critical analysis of HC theory). HC theory views schooling as an investment in skills and hence as a way of augmenting worker productivity as well (Wolff, 2000).

In 2003, the European Council has pointed out that HC is a great lever for social cohesion and economic growth, by stressing that the achievement within 2010 of the objectives put in Lisbon in 2000 in the fields of education and formation is the essential condition to put the EU social and economic policy into effect.

ments that enliven the discussion about the theme, which in turn basically rests on social-political presuppositions, are difficult to be disentangled – we are however inside it and one cannot think of driving its path with the classic, consolidated and reassuring instruments to which the macroeconomic routine has accustomed us.

It is useless, it seems to me, to think of coming out and solving the problems it presents for the developed economies, first of all on the trade side and then on the whole economic financial one, by going back to tolls and tariffs. Indeed, someone proposes now the customs tariffs, but, after all, without actually trusting them and having he himself the perception that this is a blunt tool, which might also turn into a remedy worse than the evil.

Not even there is a need to deceive ourselves of being in a position to compete with the emerging countries with instruments that, while for the developed countries, with socio-economic structures still much welfare state oriented and basically old, are much blunt, for them are absolutely sharp.

What I mean is that for us, to make resort to the "easy", low cost and not unionized labour force, is not even thinkable, on the economic side more than on the ethical-social one, and indeed nobody not even speaks about it.

Thus, it is necessary to approach the globalization with new mentality and tools, which after all are not new, but rather old, and to go along ways so far unfortunately little explored, not to say, neglected.

In other words, it is necessary to put the emphasis on the intangible characteristics of western societies and to basically make resort to an instrument, they can have at their disposal: the high level HC formation and by firmely trusting it, that a society like ours, and I refer both to our country and, even more, to the European Union (EU), should be in a position to implement better than any other social context. This, if not because of historical background and substance itself of their development model, based on presuppositions of strong solidarity, wide human content and social equity.

Of personal HC, rather than of the national one, to conjugate, according to Dagum and Slottje (2000) "the endogenous economic development to the efficiency and the social equity, and therefore, to deal with the problems of socially unacceptable poverty and inequality".

In this way, we come back to micro, by referring to the individual aspects, or even better, to the personal ones, and namely, to the personal income distribution, which goes back to Pareto's work, who was the first to analyze a personal income distribution model.

Always according to Dagum and Slottje (2000), the theoretical production function implied by this personal income distribution concept includes labour among the arguments. But not a labour measured in man-hours worked, but a HC employed. Hence, the usual specification of the production function, $Y=f(L,K)$, with Y the output produced by labour L and capital K, transforms in $Y=f(H,K)$, with the HC, H, that substitutes the generic labour stock L, so that the output is produced on the basis of capital and of a set of factors such as experience, formation and education incorporated into the labour.

Hence, the question of production and of its formalization through an ade-

quate neo-classic-fashioned functional form, if one looks at it with eyes not veiled with too much technical schematism, takes just more interesting contours, if not from an ideal point of view, with the intervention, as an input in the production function, of the concept of HC and of a measurement of it.

From this use of the HC as a production factor to the further extension of the theoretical frame up to the production itself of HC, and hence, to the HC as the result of a given production process which uses certain inputs, the step is not short – and we'll see it later on – but it is almost spontaneous to develop the discussion in this direction.

In this respect, things are not simple at all. Indeed, to try to model HC production by means a production function with arguments represented by proper inputs is an anything but easy operation, first of all, at a conceptual level, and then, as a consequence, at the level of choice and definition of both independent and dependent variables. For instance, Dagum-Slottje (2000) again try to solve the problem by treating HC as a latent variable.

There are problems of a convincing adjustment of the theoretical frame to the neo-classic production theory context which is thought to describe and explain the production of concrete objects and not of an abstract concept such as HC. Solved this problems, others of choice, definition and measurement of the variables do appear.

Once reasonable and consistent solutions to all these problems are found, one can proceed to the estimation of the parameters of the HC production function. Afterwards, one can estimate "how" HC is produced, that is, how much the production process is efficient in producing, and, of such an efficiency degree, to provide a measurement endowed with satisfactory properties.

The latter is the overall objective of this paper. Specific goal is to examine the HC formation at university level that is performed in Italy, by regarding it from the point of view of the technical efficiency, and by taking the case of Florence as an example.

This paper basically reproduces the discussion, the models and the outcomes that have occurred in the framework of the research project on the frontier efficiency concerning the production in textile and in services sectors and in the university education sector, the latter published in a number of papers (Ferrari *et al.*, 2001 and Ferrari and Laureti, 2004a, 2004b, 2005).

Thus, in paragraph 2, the theoretical foundations that are at the basis of the modelling that is convenient to implement in order to proceed in the analysis by making use of the tools provided by the neo-classic theory of production to describe and analyze the production process will be discussed. In paragraph 3, the above technical tools, represented by the production frontiers and by the efficiency measures, properly fitted to the specific case, will be illustrated and discussed. In paragraph 4, the dataset on which the empirical analysis is based will be described. In paragraph 5, a discussion of the main findings concerning individual graduates efficiency, the efficiency due to faculties and the factors that explain the average grade at exams will be carried out, whereas some conclusions will be drawn in paragraph 6.

2. THE HC PRODUCTION PROCESS IN THE ITALIAN UNIVERSITY

2.1 *The model*

The formation of a graduate is a very complex process, which involves cultural and intellectual elements, such as lectures, tutorials, discussions, seminars, study groups, as long as material and non material goods such as libraries, class rooms, computers, assistance services to students, secretaries and tutors. As such, this formation process can be regarded as a production process.

In fact, a student enters the university with a given background and specific knowledge, mostly deriving from his elementary, intermediate and secondary education and is transformed, after a certain number of years – during which he is submitted to a "cultural transformation" process – in a more educated person (that is, in a person with a higher cultural level).

Thus, one can affirm that, specifically, it is a HC production process in which university, through the education it delivers, transforms a "raw" material in a "refined" one, by means of the utilization of those cultural, intellectual material and non material production factors. In this context, the student who is submitted to the above cultural transformation process can be viewed as a production factor.

Nevertheless, it is evident that the education process is characterized by peculiarities that differentiate it from the "classic" production factors, that is, from those processes in which a set of material and non material factors is transformed in one or more material or non material, and make it a process much complex to conceptually bound.

Among these peculiarities, one should stress the *intangible and multiple nature* of the product, which is represented by a graduate endowed with a certain amount of knowledge, as well as the *heterogeneity of students*, who differ each other as regards many exogenous dimensions, such as gender, age, kind of secondary school diploma and, above all, as regards psychological, social, family characteristics. In other words, the process is characterized by a heterogeneity, both in production factors and in product, stronger than that one can see in goods and services production, which makes it a very peculiar case in neo-classic panorama.

A further peculiarity is represented by the fact that the student, besides being the object of the education process, takes part more or less actively in its delivering and influences it result as well.

Notwithstanding the above peculiarity, the analysis of this HC production process can actually be performed, as it has just been anticipated, within the neoclassic production theory.

Nevertheless, it is worth making some further remarks. First of all, in the neoclassic approach it is implicitly postulated that the production process in not affected by "external" conditions represented by the *characteristics of the area* in which it takes place. It is, of course, influenced by external infrastructures such as roads, ports, transportation means, etc., but the production model assumes that the product only depends on its own specific conditions.

On the contrary, the university education process is heavily influenced by fur-

ther conditions, besides the specific ones, such as the research activity carried out, the overall study and life atmosphere in the university, the socio-economic context of the area in which the university is situated. As a consequence, the condition that the process takes place in a isolated way is violated.

Additional difficulties are represented by the fact that some factors, such as the time devoted to study by the student, his interest, his inborn capacity, teachers' quality, etc., cannot be easily observed and quantified. Consequently, they cannot be expressed as production factors in the process, which therefore characterizes itself as one in which it is hard to *identify, quantify and measure* all production factors.

The same remarks can be made as far as the product is concerned. Indeed, it is often represented by qualitative variables difficult to identify and define in the quantitative framework.

In spite of all the above, it is possible to identify a reasonable set of production factors and products that allows to model the production process.

2.2 *The production unit*

As occurs in neo-classic production approach, there is here too the need of identifying the production unit which, by using a given set of factors, obtains a given set of products.

Let's set aside at once the intent, reasonable as based on the idea that it is capable of producing graduates by using teachers, class rooms, libraries, etc., of taking a faculty as the production unit. The empirical analysis will be bounded to Florence University, due both to data availability reasons, and to the need to eliminate as much as possible the influence of the above environmental factors. As the faculties are 11, the production space would be too restricted to provide significant results: given the number of production units and that of factors and products, the degrees of freedom would be too few.

Let's discard the idea to bring ourselves at a lower aggregation level, that of the degree courses, that are about 40, and to consider them as production units, since they present many problems of lack of data about teachers, class rooms, etc..

The question can be solved by approaching the reasoning from a representation given by Catalano and Silvestri (1992), of a production process where the university provides the student with goods and services to be utilized to carry out production activity, which allows to identify the production unit as the *student who produces himself* as a graduate.

Indeed, in the university education, where the student is in the meantime production factor, protagonist and recipient of the production activity, it seems reasonable to keep him as the production unit which, by utilizing didactic resources and equipment provided by the faculty he attends, produces himself as a graduate.

Thus, the production process is modelled as a factor-products space where the observed points are represented by the students.

2.3 *Factors and products*

The identification, quantification and measurement of the production factors and of the products associated to the above production process is somewhat problematic.

The first difficulties are met on the production factors side. Indeed, the production process is carried out by referring to a student of a given faculty. Therefore, the factor set is formed by the material and non material goods provided by the faculty and by the student himself. The former are represented by professors, researchers, sits in classrooms, tutorials, textbooks, reviews and papers and by all the assistance services, both bureaucratic and tutorial. The latter are represented by the education level reached by the student before approaching university, that is, by the degree of "accumulated knowledge" and by individual characteristics. As a consequence, at least identification and measurement problems there emerge.

If as for the production factors there are uncertainties, the problems related to the identification, quantification and measurement of proper products are even more complicated.

There are, first of all, noteworthy difficulties in the interpretation of the "exit" of the process, as the graduate is characterized by a set of variables, mainly qualitative, that reflect the level of education he received.

This is, as anticipated in the Introduction, a complex and not easy to be solved question, since actually there is a need of measuring the HC represented by a graduate, even though in an indirect way and without the intent – as this is not an objective of this paper - to deeply enter the many and debated problems that are linked to it.

It is evident that here the problem is approached and solved in a very operational way, by making resort to a reasoning scheme that is finalized to the goal, not to identify a HC production function, but, in a much more restricted way, to identify a production function in which the exits of the process i.e., the products – we will see in a moment that there will be one exit only – are a reasonable and acceptable quantitative representation of the specific production activity that has been modelled.

One can imagine various quantitative representations. A possible, reasonable and available measure of the education level embodied in the graduate is constituted by the average grade at exams. And indeed, this will be the representation of the product that will be used in this paper.

3. THE MEASUREMENT OF THE EFFICIENCY IN THE UNIVERSITY EDUCATION PROCESS: THE FRONTIER EFFICIENCY

3.1 *The non-parametric approach*

Two are the approaches that one can follow in the construction of production frontiers: the non-parametric and the parametric ones. While the former is essen-

tially of a deterministic type, much utilized in engineering framework, the latter specifies further in deterministic and stochastic approaches. In this paper, the non-parametric approach will be used.

In the non-parametric approach, the (relative) production efficiency, that is, the frontier efficiency, is elaborated from a set Y_0 of observed data, which is a production set formed by the vectors of production factors and products $(\mathbf{x},\mathbf{y})$, that is, the set of the production processes (also named net output vector) formed by the factors and products vectors.

The above frontier efficiency is measured with reference to a "hypothetical" production set Y, which, as is usual to do, is assumed to be generated by the set Y_0.

If one intends to measure the efficiency as regards the production factors, the input-efficiency, of the set Y_0, the problems is to establish which is the frontier of the production set Y, which is usually unknown, as regards to which to measure the distance of each observed point.

In this context, Farrell (1957), following Debreu (1951) and Koopmans (1951), in defining and applying the method that looks at the frontier as the *convex envelopment of the factors free disposal* called Data Envelopment Analysis (DEA) (Charnes and Cooper, 1995) has proposed the following technical input-efficiency measure of the production factors vector in the production of $\mathbf{y}$: $F(\mathbf{x},\mathbf{y}) \equiv min\{\lambda|\lambda\mathbf{x}\in L(\mathbf{y})\}$[2].

The production factors vector belongs to the sub-set of all production factors vectors $\mathbf{x}$, named *input requirement set*, $L(\mathbf{y})$, capable of producing at least the products vector $\mathbf{y}$.

The above is an input-efficiency measure defined as the inverse of a *(radial) distance function* or *input distance*, $d(\mathbf{x},\mathbf{y})$, represented by a positive scalar $\lambda = \frac{\|\mathbf{x}\|}{\|\mathbf{x}^*\|}$, where $\|\cdot\|$ denotes a vector module and $\mathbf{x}^*$ denotes the input-efficient vector, that is, the vector that on the frontier of $L(\mathbf{y})$ is identified by the line from the origin to $\mathbf{x}$. Hence, this distance is an inefficiency measure as well.

It is evident that, if $(\mathbf{x},\mathbf{y})\in Y$ is input-efficient, $\mathbf{x}=\mathbf{x}^*$ and the above measure is equal to 1. Consequently, $\forall (\mathbf{x},\mathbf{y})\in Y$ input-inefficient, i twill be $F(\mathbf{x},\mathbf{y})<1$ (in fact, $\mathbf{x}=\lambda\mathbf{x}^*>\mathbf{x}^*$ and therefore $\mathbf{x}\in L(\mathbf{y})$). Viceversa, if $F(\mathbf{x},\mathbf{y})>1$, $\mathbf{x}=\lambda\mathbf{x}^*<\mathbf{x}^*$ and therefore $\mathbf{x}\notin L(\mathbf{y})$ and $(\mathbf{x},\mathbf{y})$ is not a feasible process. Hence, $0 \le \lambda \le 1$.

The Farrell efficiency measure presents a not negligible inconvenience. Since it has the isoquant as reference set – being defined as radial reduction (which implies a proportionate reduction of all inputs) – and although the isoquants include the efficient sets, the reciprocal is not necessarily true. Thus, it may occur that the

[2] In non-parametric case, among the possible methods, DEA is one of the most widely used, as it presents several attracting features. By and large, it is a deterministic method – and as such it is used in this paper – as it assumes all differences from the frontier to be the results of inefficiency. Nevertheless, there is a study stream that extends the analysis to the stochastic ground, although with many complications.

efficient vector is *one* of the efficient vectors and not the unique. In this case, the condition that a production process is technically efficient if and only if it is impossible to produce more of some product without using more of some production factor, or without producing less of some other product is violated, with the consequence of producing with bigger quantities of some factor, or of obtaining smaller quantities of some product than it would be possible.

This is an anything but theoretical possibility, that can occur in a wide range of specifications of the production technology much utilized.

This drawback can be overcome by making resort to non-radial technical efficiency measures.

The Russell measure, a non-radial generalization of Farrell measure, that has been proposed by Färe and Lovell (1978), and that is not the case to report here, is flexible enough to allow to select a vector from the set of all the efficient vectors as regards to which to calculate the technical efficiency.

In this paper, since some factors are not under the direct control of the student, the Farrell efficiency measure that is used is an output-efficiency, $\phi = \frac{\|\mathbf{y}\|}{\|\mathbf{y}^*\|}$.

This is a measure defined as a *(radial) distance function*, or *output distance*, where $\mathbf{y}^*$ denotes the output-efficient vector, that is, the vector that on the frontier of the sub-set P(**y**), defined as the set of all vectors of products affordable from the production factors vector **x**, is identified by the line from the origin to **y**.

Looking at the production activity in a wide meaning, one should note that, while in *private sector* the objective of the production unit is unique and consists of maximizing profit, for the *public production unit* the objectives are manifold, difficult to quantify and partly irreconcilable among them. According to Pestieau and Tulkens (1969), they can be summarized in: (*i*) efficiency (both technical and economic), (*ii*) equity, (*iii*) financial equilibrium, and, (*iv*) macroeconomic improvements (as regards unemployment, inflation, etc.). Now, it is clear that they represent as many constraints in the decisions that must be taken by the public entrepreneur, and, furthermore, contradictory constraints, since choices that would be positive from the economic point of view, can be questionable on the equity field, and vice-versa. In the evaluation of the global *performance* of a public production unit, it is necessary to take account of this circumstance, so that one should be able to establish in which proportion each of the objectives has been achieved, which is decidedly difficult.

A way to overcome the complexity of the problem consists of restraining oneself to evaluate the behaviour of the public production unit from the point of view of the *technical efficiency*, the only objective whose achievement does not prejudice the achievement of the others, since, as pointed out by Pestieau and Tulkens (1990), to produce too little or to use too many production factors as regards how much is technically possible is surely the consequence of a wrong behaviour that cannot be justified by the pursuing of none of the other above listed objectives.

Since the students who produce themselves as graduates make it in universities

that are for more than 90% public, it seems opportune to conduct the analysis in the framework of the concept of technical frontier efficiency.

The production set Y which satisfies the hypothesis of variable returns to scale, the assumption that better is justified as it is very general and better fits the data base, can be expressed in the following form:

$$P = \left\{ (\mathbf{x},\mathbf{y}) \;\middle|\; x_i \;\geq \sum_{j=1}^{n} \lambda_j x_{ij} \;\; (i = 1,..,s) \;\; y_r \;\leq \sum_{j=1}^{n} \lambda_j y_{rj} \, (r = 1,..,m), \lambda_j \geq 0 \;\; (j = 1,...,n) \right\}$$

where n is the number of units, s the number of factors and m the number of output and λ_j is the weight of the j^{th} unit when the reference point on the frontier is defined. The variable returns to scale imply that the set Y forms a polyhedron.

The Farrell output efficiency measure, specified as $F_O(\mathbf{x},\mathbf{y}) = \max\{\phi | \phi\mathbf{y} \in P(\mathbf{x})\}$, expresses the radial expansion of the output vector **y** determined as ratio between observed vector and vector on the frontier.

The measure can be obtained by solving the following linear programming system:

$$\max \phi_j$$

subject to:

$$\sum_{h=1}^{n} \lambda_h y_{rh} \geq \phi_j y_{rj} \qquad r = 1,...,m; \quad j = 1,...,n$$

$$\sum_{h=1}^{n} \lambda_h x_{ih} \leq x_{ij} \qquad i = 1,...,s$$

$$\lambda_h \geq 0;$$

$$\sum_{h=1}^{n} \lambda_h = 1$$

3.2 *The evaluation of the faculty role*

The efficiency exhibited by each student in producing himself as a graduate can partially be attributed to the efficiency of the faculty he attends.

In fact, besides the accumulated knowledge and the individual characteristics, individual production factors provided by each student and different case by case, the faculty provides the students with the other factors; as above said, class rooms, furniture, libraries, teachers, just made available to all students in the same quantity and quality

Now, it is evident that if one would aim to evaluate the efficiency of the students of a faculty, there would be no problem, as the measure would not be influenced by differential factors other those represented by the individual characteristics.

Since, however, one aims at evaluating the production *performance* of students belonging to different faculties, it is necessary to keep into consideration also the faculty, in order to measure the efficiency independently on the weight represented by the faculty itself.

In other words, it is necessary to decompose the efficiency in the part ascribable to the student and in the part ascribable to the faculty.

This can be done through a procedure called *Programme Evaluation* (PE), that is, evaluation of the programme or faculty, which allows to decompose the overall efficiency in a *within* component and in a *between* component.

In order to better understand this decomposition, it may be useful to refer to the figure below, by imagining, for the sake of simplicity, a production situation with one factor and one product only, relative to students from two faculties, *A* and *B*.

The frontier of Faculty *A*, based on DEA with the hypothesis of variable return to scale, is represented by the line from x axis and passing through A_1, A_2, and A_3; the frontier of Faculty *B* is represented by the line from x axis and passing through B_1, B_2, and B_3, both of them are *within* frontiers (the students' frontier internal to the faculties). The overall frontier, that envelops all graduates, is represented by the line from x axis and passing through A_1, A_2, B_2'', B_4'', and B_3.

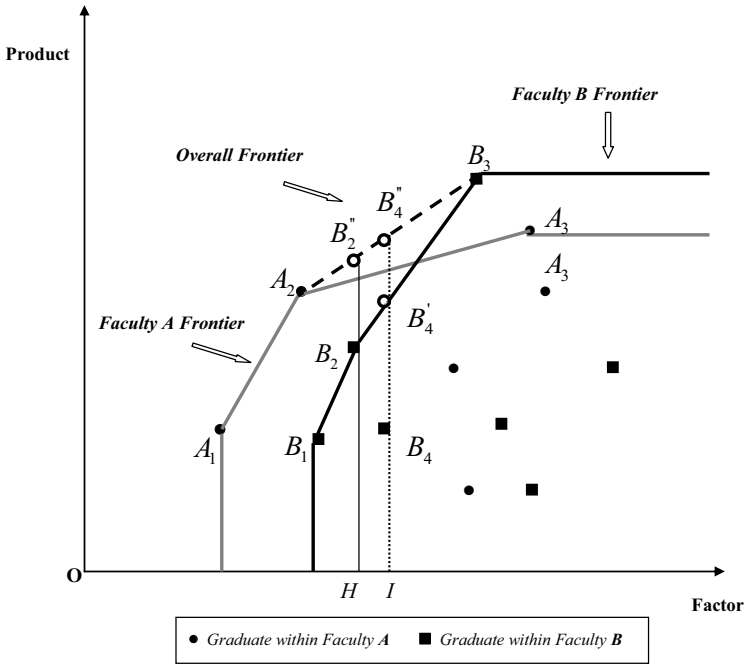

Fig. 1 – Overall and within frontiers under the hypothesis of variable returns to scale.

Consequently, in the output efficiency framework, the student B_4 of the Faculty B, has an overall efficiency equal to $IB_4 / IB_4^{''}$.

This measure can be decomposed in: (*i*) a *within* efficiency, defined as the ratio $IB_4 / IB_4^{'}$, which represents the portion of efficiency of the student B_4 as regards the best results obtained by the students of the Faculty B only ascribable to his individual characteristics; (*ii*) a *between* efficiency, represented by the ratio $IB_4^{'} / IB_4^{''}$, which provides a measure of the impact of the Faculty B on student's *performance*.

Conversely, the student B_2, always of Faculty B, exhibits a *within* efficiency equal to 1, and a *between* efficiency equal to $HB_2 / HB_2^{''}$, due to the inefficiency of his faculty.

Finally, the student A_1, situated on the global frontier, exhibits both *within* and global efficiency equal to 1.

Actually, a way to estimate the faculty weight consists of calculating the output efficiency for graduates separately for each of 11 faculties and of bringing the inefficient graduates up to their faculty frontier, so to obtain a new envelop where the inefficiency is eliminated. Then, the significance of the differences among faculties is tested by means of a non-parametric test, usually the Kruskal and Wallis one.

4. DATA USED

The data used in the analysis come from different sources, although all in the framework of the survey conducted at the Department of Statistics of the University of Florence concerning the labour outfall of about 2,500 graduates in 1998 in the University of Florence.

Additional data on didactic production factors of the faculties during the period 1990-98 have been collected at the administrative archives and the Central Library of the University of Florence.

After eliminating 41 outliers as regards the variable "length of study" by using the so-called three-sigma rule, by putting at less than 5% the probability that such a variable is three times greater than the standard deviation, the final group was formed by 2,236 graduates. The following production factors have been selected:

Human resources:

1. average number of full and associate professors by student;
2. average number of researchers by student.

Capital:

3. average number of sits in class room by student;
4. average number of class rooms by student;
5. average number of books in library by student;
6. average number of reviews in library by student;

7. average number of furniture provided by University by student;
8. average number of instruments provided by University by student.

Individual factors:

9. secondary school diploma grade;
10. actual length of degree course.

As regards the product, the average grade at the exams has been chosen to represent it.

5. EVIDENCE FROM THE UNIVERSITY OF FLORENCE

5.1 *Overall graduates efficiency*

The output efficiency measures ϕ_i, obtained by means of the software DEAP 2.1 (Coelli, 1996), have been grouped by classes, as reported in Table 1.

A brief comment of them is in order (a detailed discussion of the case of constant return to scale can be found in Ferrari and Laureti, 2005. Here, the aim is to outline the main findings under the assumption of variable returns to scale useful to a better understanding of the development given in the following pages).

TABELLA 1

Frequency distribution of the overall technical efficiency measures

Efficiency classes	*Frequencies* n_i	*% Frequencies* p_i	*Cumulated % Frequencies* P_i	*Reverse Cumulated % Frequencies* P_i^*
0.70-0.80	32	1.4	1.4	100.0
0.80-0.85	181	8.1	9.5	98.6
0.85-0.90	419	18.7	28.2	90.5
0.90-0.95	753	33.7	61.9	71.8
0.95-1.00	851	38.1	100.0	38.1
Total	**2,236**	**100.00**		0.0

First of all, it is worth noting that no graduate has an efficiency lower than 70%: if 50% efficiency is taken as a turning point for efficiency from low to acceptable level, one can affirm that by far there is no HC formed in an inefficient way, all the more so because more than 90% of graduates exhibits an efficiency greater than 85%.

Full efficiency, that is, $\phi = 1$, occurs in 161 cases, 7.2% out of the total.

Anyway, very satisfactory efficiency, that is, efficiency $\geq$ 90%, is reached by 71.8% of graduates. But also, an efficiency of 80% to 90% is not a despicable level at all and it is reached by 26.8% of graduates.

As the overall average efficiency is 0.925, the unexploited production capacity is 7.5%, that is, the inefficient graduates do not have exploited 7.5% of the production factors. In other words, they could have increased the average grade at exams by 7.5% without augmenting the human resources, capital and individual characteristics factors. This average margin is rather low and testifies that the HC formation in the University of Florence takes place in a satisfactoriles efficient

way. The factors are, of course, not fully efficiently employed, but not at a dramatic level at all.

5.2 *Efficiency by faculty*

Technical efficiency measures by faculty are reported in Tab. 2.

TABELLA 2

Technical efficiency measures by faculty

Faculty		*Number of graduates*	*Overall average (geometric) efficiency*	*Within efficiency*	*Between efficiency*
Fac1	Agriculture	70	0.907	0.950	0.955
Fac2	Architecture	396	0.946	0.946	1.000
Fac3	Economics	388	0.889	0.932	0.955
Fac4	Pharmacy	53	0.949	0.949	1.000
Fac5	Law	275	0.927	0.927	1.000
Fac6	Engineering	214	0.894	0.928	0.964
Fac7	Literature & Philosophy	230	0.959	0.966	0.993
Fac8	Medicine & Surgery	80	0.922	0.964	0.956
Fac9	Sciences, Mathematics, Physics	170	0.894	0.935	0.956
Fac10	Political Sciences	164	0.942	0.944	0.998
Fac11	Formation Science	196	0.966	0.969	0.997
	Overall	**2,236**	**0.925**	**0.943**	**0.981**

A look at column 2 shows that graduates in Formation Sciences are the most efficient, with their overall efficiency average level of 0,966, a little bit higher than that of the graduates in Literature & Philosophy (0.959), Pharmacy (0.949), Architecture (0.946), Political Sciences (0.942), and Medicine & Surgery (0.922), against the graduates in Economics, who are the less efficient, with their overall average level of 0.889, which, however, is not dramatically far away from the above levels.

After checking through the Kruskal and Wallis test that the association between faculty and degree of efficiency is significant, the application of the PE procedure has provided the results shown at columns 3 and 4 of Tab. 2.

As has been stressed above, the weight of the faculty in students' performance is relevant. Also, it varies across faculties, as was easy to expect and as column 4 of Tab. 2 clearly show.

Moreover, there is evidence that the contribution of faculties to the overall efficiency is greater than that ascribable to students' individual characteristics. In fact, the *between* efficiency, except that in the case of Medicine & Surgery, is always higher than the *within* one.

The faculties that more contribute to students' efficiency are Architecture, Pharmacy and Law, with between efficiency accounting for 1. This also shows that the three faculties are fully efficient, that is, that they do not need to increase the average number of full and associate professors, the average number of researchers for student, the average number of sits in class rooms, the average number of class rooms, and the average number of books in library; a somewhat unexpected result that need to be further investigated in subsequent analyses.

Again, Political Sciences, with a between efficiency of 0.998 and Formation Science, with a between efficiency of 0.997, provide a relevant contribution, vir-

tually equal to 1. All the other faculties show lower contributions, even though in no case lower than 95%.

On average, the faculties contribution to efficiency accounts for 0.981, which means that the students could have increased the average grade at exams by only 1.9% without the faculty should increase the above mentioned human resources and capital factors.

The graduates individual characteristics contribution to efficiency, besides being lower than that provided by faculties, except in the case of Medicine & Surgery, as remarked above, is somewhat lower on average, 0.943 - the students could have increased the average grade at exams by 6.7% without the need of augmenting the individual characteristics factors - ranging from the highest *within* efficiency exhibited by Formation Science, 0.969, to the lowest efficiency exhibited by Law (0.927). In the latter case, the students' individual inefficiency is particularly heavy as they do not fully take benefit from the excellent organization of their faculty (*between* efficiency equal 1).

As a consequence, the Faculty of Law does not need to increase its efficiency, but should take account of the fact that its students have on average more reduced individual capacity and therefore eventually establish preliminary programmes of integration of the education its students have received in secondary school.

More generally, any faculty with a relatively higher *between* component should behave similarly.

Vice-versa, any faculty with a relatively higher *within* component should try to improve its own efficiency, by equipping itself with better human resources and capital, such as better teachers, administrative personnel, structures, etc.

5.3 *Factors that explain the average grade at exams*

Selecting the graduates with *within* efficiency equal to 1 or nearly 1, whose number total 513, through a regression analysis it is possible to analyse their characteristics.

It is convenient to carry out distinct estimation for each faculty; it is thus possible to analyze the graduates under the same organization conditions and resource endowment.

The dependent variable is represented by the exams average grade. The explanatory variables, which express individual characteristics, are gender and residency, whereas the school background is expressed by type of secondary school diploma (scientific or classic lyceum, or other type of secondary school), elementary, intermediate and secondary school attendance regularity, i.e., whether the graduate has been a regular school path or he had to repeat some year, and whether the graduate has been working during schooling or not.

Results concerning the two most interesting faculties, of course to this respect, are presented at Tab. 3. As for the other faculties, the analysis of results is still under control and discussion and will be globally presented and discussed in a subsequent paper.

TABELLA 3

Regression analysis for Law and Literature & Philosophy

Faculty		**Law**			**Literature & Philosophy**	
Variable	*Coefficient*	*Std. Err.*	*p-value*	*Coefficient*	*Std. Err.*	*p-value*
Gender						
(Reference category = Males)	1.165	0.516	0.029	0.354	0.173	0.044
Residency						
Florence province	-0.061	0.523	0.908	0.013	0.169	0.937
Other province	-1.828	1.331	0.177	0.190	0.314	0.546
Previous education						
Diploma						
(Reference category = No Lyceum)	0.423	0.592	0.479	0.426	0.194	0.031
School regularity						
(Reference category =Regular)	-1.608	0.669	0.020	-0.189	0.233	0.421
Working during schooling						
(Reference category = No)	-1.138	0.501	0.028	0.236	0.168	0.164
Constant	27.213	0.726	0.000	28.648	0.233	0.000
N.observations	53			82		
Adj.R-square	0.2125			0.1215		

If one takes the efficient graduates in the Faculty of Law, which exhibits the lowest level of *within* efficiency, among the individual characteristics that influence the achievement of high average grades he finds the primary and secondary school regularity, the gender and working activity during the study.

Among the most efficient graduates of the Faculty of Literature & Philosophy, the gender and the kind of diploma show a particular relevance in the achievement of high average grades.

By and large, it should be stressed that the school regularity, that is not significant for the graduates in Literature & Philosophy, takes a special importance in nearly all the other full efficient graduates of Florence University, particularly for graduates in Agriculture, Architecture and Economics Faculties.

6. CONCLUSIONS

In the measurement of the technical efficiency of 2,236 graduates in 1998 in the University of Florence, that is to say, in the formation of HC at university level, the starting point has been represented by the modelling of the production process as one in which the student produces himself as a graduate.

This is an original way of looking at the production process which generates HC that allows to operate in a satisfactory technical way in the framework of the production theory, with interesting and promising results, that can provide further useful developments.

The tool that have been used is the DEA methodology, one of the most widely used in the framework of the non-parametric approach to the frontier technical efficiency, chosen for undertaking this analysis, that seemed to be the most adequate to describe and interpret such a production process. Likewise, it appeared most adequate to conduct the analysis under the variable returns to scale assumption.

The factors utilized are a set of human resources and capital provided by faculties, joined to two individual factors represented by the secondary school diploma grade and by the actual length of degree course.

The analysis has been conducted both for all graduates and at faculty level, in order to focus the contribution provided by the latter to efficiency, by disjoining it from the efficiency due to graduates individual characteristics. Moreover, inside the group of 513 graduates with full or nearly full efficiency, a regression analysis has been conducted, with the aim of identifying the factors correlated to the average grade.

Globally, there was evidence that the 2,236 students who graduated in 1998, did it with an average efficiency of 0.925, and therefore, an inefficiency of 7.5%, a not dramatic unexploited production capacity at all. Moreover, the efficiencies are shifted towards 100%, with 71.8% of graduates that shows an efficiency equal or greater than 90%.

At faculty level, the faculty with the most efficiently graduated students was Formation Science; on the contrary, the Faculty of Economics is the one with the less efficient graduates.

The contribution provided to technical efficiency by faculties' organization (*between* efficiency) is greater than that provided by the students individual characteristics (*within* efficiency).

The faculties that most contribute to graduates efficiency are Architecture, Pharmacy, Law, Political Sciences and Formation Science

There is evidence that among the factors that more explain the efficiency due to individual characteristics, there are primary and secondary school regularity, gender, working activity during university study, and kind of diploma.

Of course, these are initial, somewhat compact results, which deserve further deepening that might allow to reach more detailed and robust conclusions. Nevertheless, they constitute an already very explanatory information material, which allows to get some conclusions and offers indications for further research extension.

Dipartimento di Statistica "G. Parenti" GUIDO FERRARI
Università di Firenze

REFERENCES

G. CATALANO and P. SILVESTRI (1992), *Il finanziamento del sistema universitario italiano*, "Ministero del Tesoro, Commissione Tecnica per la Spesa Pubblica".

A. CHARNES and W.W. COOPER (1985), *Preface to topics in data envelopment analysis*, "Annals of Operations Research", Vol. 2, pp. 59-94.

T.J. COELLI (1996), *A guide to DEAP version 2.1: a data envelopment analysis (computer) program*, "CEPA Working Paper" 96/8, Department of Econometrics, University of New England, Armidale NSW Australia.

C. DAGUM and D.J. SLOTTJE (2000), *A new method to estimate the level and distribution of household human capital with application*, "Structural Change and Economic Dynamics", 11, pp. 67-94.

G. DEBREU (1951), *The cefficient of rsource utilization*, "Econometrica", Vol. 19, N. 3, pp. 273-292.

R. FÄRE and C.A.K. LOVELL (1978), *Measuring the technical efficiency of production*, "Journal of Economic Theory", 19 (October), pp. 150-162.

M.J. FARRELL (1957), *The measurement of productive efficiency of production*, "Journal of the Royal statistical society". Series A, General, Vol. 120, pp. 253-281.

G. FERRARI, T. LAURETI and M. MALTAGLIATI (2001), *Evaluating the efficiency of the human capital formation in the university of Florence through the DEA*, "Atti della Società Italiana di Statistica, Convegno Intermedio", Roma, 4-6 giugno, pp. 25-28.

G. FERRARI and T. LAURETI (2004a), *Multi-factor efficiency analysis of the Florence University graduates*, "Atti della XLII Riunione Scientifica della SIS", Bari, 9-11 giugno, pp. 503-506.

G. FERRARI and T. LAURETI (2004b), *Evaluation of University graduates technical efficiency using data envelopment analysis combined with multiple correspondence analysis: the case of the University of Florence*, "Quantitative Methods in Economics", Bratislava.

G. FERRARI and T. LAURETI (2005), *Evaluating the technical efficiency of human capital formation in the Italian University: evidence from Florence*, "Statistical Methods and Applications (SMA)", pp. 14-2.

T.C. KOOPMANS (1951), *Analysis of production as an efficient combinations of activities*, in Koopmans T.C. ed., "Activity Analysis of Production and Allocation", Cowles Commission, for Research in Economics, 13.

P. PESTIEAU and H. TULKENS (1990), *Assessing the performance of public sector activities: some recent evidence from the productive efficiency viewpoint*, "CORE Discussion Paper" n. 9060, Center for Operation Research and Econometrics, Université Catholique de Louvain, Louvain-La-Neuve.

G. VITTADINI, M. COSTA, C. DAGUM and G. LOVAGLIO (2003a), *A microeconomic recursive model of human capital, income and wealth determination: specification and estimation*, "Meeting of the American Statistical Association, San Francisco".

G. VITTADINI, M. COSTA, C. DAGUM and G. LOVAGLIO (2003a) (2003b), *A method for the estimation of the distribution of human capital from sample surveys on income and wealth*, "Meeting of the American Statistical Association", San Francisco.

E.N. WOLFF (2000), *Human capital investment and economic growth: exploring the cross-country evidence*, "Structural Change and Economic Dynamics", 11, 4, 433-472, December.

RIASSUNTO

La modellazione della formazione del capitale umano nell'università e la misura della sua efficienza. Il caso dell'Università di Firenze

In questo lavoro viene effettuata un'analisi dell'efficienza tecnica di 2.236 laureati nel 1998 nell'Università di Firenze, vale a dire, della formazione di capitale umano a livello universitario, attraverso la modellizzazione del processo di produzione come di un processo nel quale lo studente produce se stesso come laureato.

Lo strumento utilizzato è la metodologia DEA, sotto l'ipotesi di rendimenti di scala variabili.

I fattori utilizzati sono rappresentati da un insieme di risorse umane e di capitali forniti dalle facoltà, insieme a fattori individuali rappresentati dal voto di diploma di scuola secondaria e dalla lunghezza degli studi universitari.

L'analisi viene condotta sia per tutti i laureati globalmente, sia a livello di facoltà, allo scopo di evidenziare il contributo fornito da queste ultime all'efficienza.

Si evidenzia che gli studenti si sono laureati con una efficienza media superiore al 90% e, quindi, con una capacità produttiva inutilizzata inferiore al 10%.

A livello di facoltà, quella più efficiente risulta essere Scienza della Formazione; quella più inefficiente è Economia.

In generale, il contributo all'efficienza fornito dalle facoltà è superiore a quello portato dalle caratteristiche individuali degli studenti.

SUMMARY

Modelling university human capital formation and measuring its efficiency. Evidence from Florence University

In this paper, an analysis of the technical efficiency in the formation of 2,236 graduates in 1998 in the University of Florence, that is, in the university human capital formation, is performed, by modelling the production process as one in which the student produces himself as a graduate.

The tool utilized is the DEA methodology, under the hypothesis of variable returns to scale.

The production factors are represented by a set of human and capital resources provided by the faculties, along with individual factors represented by secondary school diploma score and by the length of university study.

The analysis is conducted both for the overall graduates, and at a faculty level, in order to emphasize the contribution provided by the latter to efficiency.

There is evidence that the students graduated with an average efficiency greater than 90% and therefore with an unexploited productive capacity lower than 10%.

At a faculty level, Formation Science appears to be the most efficient, whereas Economics is the less efficient one.

By and large, the contribution to efficiency provided by faculties is greater than that brought by students individual characteristics.