Human capital, social capital and scientific production

Mathieu Goudard Michel Lubrano

June 2010

1 Introduction

What are the motivations for scientists to publish their results and which outside factors influence their publication performance? These are the main two questions addressed in the fields of sociology of science and economics of science. In her survey paper, Stephan (1996) (see also Diamond, 1996) suggests a diversity of explanations for the individual production of scientific papers. Above the simple satisfaction of scientific curiosity, one of the main motivations of scientists is the recognition awarded by the scientific community for being first to publish a main discovery. There is no prize for being second, remember the dispute around the discovery of the VIH. Scientific activity represents a very risky adventure. Scientific output is very unequally distributed.so that, for instance, Wages formation depends only for a fraction on scientific output. Lotka (1926) illustrated the concentration of publications among very few scientists with his famous statistical law. Merton (1968) pointed out the *Matthew effect* which shows that mature and recognised scientists are rewarded, both financially and by citations, above their real merit. The combination of these two characteristics (race for being the first and over recognition of already matured scientists) might explain the fact that we can observe scientists with an important and continuous production together with scientists who have a more cyclical production. It will thus be not very easy to find a single model to explain individual and yearly productivity.

A major ingredient for individual scientific production is human capital which is a combination of basic intelligence (initial conditions) and accumulation of efficient knowledge. The life cycle theory predicts that, due to the finiteness of life, investment declines over time. Combined with a depreciation of human capital, this explains the inverted U shape of scientific output. Several models were developed around this idea, notably by McDowell (1982), Diamond (1984, 1987) and Levin and Stephan (1991). These models recognise the importance of time in scientific discovery.

[0]We are grateful to Nicolas Carayol and Stephen Bazen for pointing out several references concerning the economic of science and the theory of human capital. We are grateful to Suzanne de Chevigné for arousing our interest in sociology. Of course errors that might be contained in this paper are solely ours.

However, these models, mainly based on time trends and cohorts effects lack explanatory power for explaining irregularities or cyclical behaviour in output. Using panel data, Levin and Stephan (1991) introduce individual fixed effect to take into account the differences in productivity which are not explained by a life-cycle effect.

These models are based strictly on individual behaviour, ignoring one fundamental aspect of human capital which can be increased by sharing. Thus the surrounding effects and the way they are exploited are determinant. Bourdieu (1980) developed the notion of social capital. After reviewing the traditional human capital model of scientific production, we shall see how it can be extended to environment and collective behaviour. We shall see that the hierarchical models are a good starting point.

2 A model of life cycle productivity for scientists

Most of the models which were built for modeling the research productivity of scientists are based on the model in continuous time of Ben-Porath (1967). This model describes the accumulation of human capital and explains the life-cycle profile of earnings. Individual invests in their human capital when they are young, anticipating future earnings. They continue to invest in their human capital, but at a lower rate, which becomes zero at the end of their career.

2.1 The theoretical initial model

Models for scientific human capital accumulation do not insist on the initial accumulation. There is no special need to measure the yield of an extra year of education. In a way, the main variable is a decision variable $s_t \in [0,1]$ which monitors yearly time allocation between using human capital K_t a in a proportion $(1 - s_t)$ for earning money and using human capital in a proportion s_t for augmenting the stock of human capital, over the career. When applied to academic scientists, $1 - s_t$ is the proportion of time devoted to routine academic occupation such as teaching, supervising PhD students, refereeing papers, participating to administrative tasks, while s_t is the proportion of time devoted to the writing of articles or books that will increase the prestige of the scientist, his number of citations, the recognition he has from his peers. The production function for supplementary human capital has the form of a Cobb-Douglas (see Diamond 1987)

$$Q_t = \beta(s_t K_t)^\alpha. \tag{1}$$

It can be augmented with other inputs such as in the original model of Ben-Porath. This function could receive several interpretations. It is a production function for supplementary human capital. So it represents an investment function, interpretation that is made clear by the equation representing the dynamic behaviour of human capital given below. But it is also a production function describing the production of scientific output. In Diamond (1987), human capital is seen as the prestige gained by the scientist and measured by the citations that other scientists make to his work. Due to the continuous progress of science, citations decrease over time and so human capital experiences an obsolescence so that the variation of K_t is given by

$$\dot{K_t} = Q_t - \delta K_t. \tag{2}$$

The objective function of the scientist is to maximise his discounted future income. Current income is provided by the exercise of his current activity (as described above, or renting his human capital for a unit wage w) and is given by

$$Y_t = w(1 - s_t)K_t. \tag{3}$$

In the initial period of formation, $Y_t = 0$ because $s_t = 1$. The objective function, similar to that of McDowell (1982), is

$$U = \int_i^T e^{-rt} Y_t \, dt = \int_i^T e^{-rt} w(1 - s_t)K_t \, dt, \tag{4}$$

the objective being to maximise the present value at age i of disposable income. A solution to this problem is found by writing the Hamiltonian

$$H = e^{-rt} Y_t + \lambda(Q_t - \delta K_t). \tag{5}$$

It expresses Q_t as a function of the parameters of the model and of the remaining time to retirement given by $T - t$. Production or investment in human capital Q_t is a non-linear decreasing function of time as:

$$\log Q_t = \frac{1}{1-\alpha} \log \beta + \frac{\alpha}{1-\alpha} \log \frac{\alpha}{\delta + r} + \frac{\alpha}{1-\alpha} \log(1 - \exp -(\delta + r)(T - t)). \tag{6}$$

It can be approximated using $\log(1 - x) \simeq -x$:

$$\log Q_t = \frac{1}{1-\alpha} \log \beta + \frac{\alpha}{1-\alpha} \log \frac{\alpha}{\delta + r} - \frac{\alpha}{1-\alpha} \exp(-(\delta + r)(T - t)), \tag{7}$$

which shows that the log of Q_t is exponentially decreasing with an acceleration at the end of the life cycle.

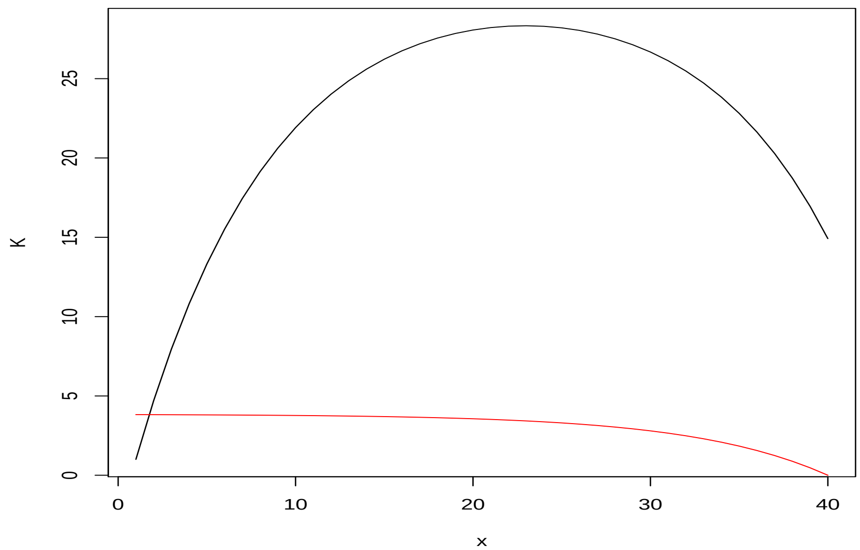
The original solution in Q_t is zero when $t = T$ (the age of retirement). Combining this equation with the capital variation equation, we can see that K_t has an inverted U shape life-cycle profile. McDowell (1982) uses this type of model to study the influence of the depreciation rate over the decision of investing in human capital, showing that children breeding interruptions causes a much smaller loss in human capital for female working in the social sciences that those working in physics.

The first derivative of Q_t with respect to δ is negative for $t < T$ and not defined for $t = T$. It increases with t. It can be computed in the general case with Mapple and has a simpler expression for $\alpha = 0.5$. Levin and Stephan interpret this behaviour as $\partial Q/\partial V > 0$, meaning that obsolescence is lower for never vintages V.

The second derivative of Q_t with respect to both δ and t is positive, increases with age and then decreases with a peak depending on the parameters including α.

We have simulated this model, using calibration values inspired from McDowell.[1] We took $\alpha = 0.5$, $\beta = 1$, $\delta = 0.12$, and $r = 0.01$. The result is

Figure 1: Life cycle human capital



displayed in Figure 1 for $T = 65$, assuming that the career starts at 25. The red line represents investment and the black line the stock of human capital. There is just a slow declining effect in productivity. The graphs show that the age effect is negative and roughly linear. On the contrary, we can see a large life cycle effect on the stock of effective human capital after the age of 50 years with the chosen parameters. This would suggest a model having the total stock of publications as the dependant variable. Adopting this model solves the problem of scientists that have a zero production for one or more years.

2.2 A Mincer equation for scientists

If all the life cycle models of scientific productivity rely on simplifications or modification of the model of Ben-Porath (1967), none of them derive a convincing estimable form. McDowell analyses theoretically the influence of depreciation of the production of human capital. Diamond (1987) simply details the model. Levin and Stephan were the first to try to estimate this model, modeling cohort effects, but they simply propose a linear form for explaining Q_t. We find only in Mincer (1974) was the only one to derive a proper linear and estimable form for life-cycle earnings. We shall follow his derivation as explained in Heckman et al (2003).

Let us consider first a simplification of the production function of supple-

[1]McDowell estimated an annual rate of decay of 13.18% for three top US journal between 1950 and 1974. The estimated rate of decay includes both δ and the annual rate of growth of the number of published papers in the field. He assumed $T = 65$ and $r = 10\%$.

mentary human capital by supposing that $\alpha = 1$ so that

$$Q_t = \beta s_t K_t. \tag{8}$$

and second a discretisation of the accumulation of human capital equation

$$K_t = K_{t-1}(1-\delta) + \beta s_{t-1} K_{t-1}. \tag{9}$$

By successive substitutions, we get

$$K_t = \prod_{j=0}^{t-1} (1 + \beta s_{t-1} - \delta) K_0. \tag{10}$$

We distinguish between a period of formation where $s_t = 1$, devoting to the writing of the PhD dissertation and a period where $s_t < 1$ so that a proportion $(1 - s_t)$ is devoted to earning money with routine academic work and a proportion s_t is devoted to writing papers that will increase the stock of human capital. Using logs, we have

$$\log K_t = \log K_0 + \sum_{j=0}^{s-1} \log(1 + \beta - \delta) + \sum_{j=s}^{t-1} \log(1 + \beta s_j - \delta). \tag{11}$$

Using the approximation $\log(1 + x) \simeq x$, we get

$$\log K_t = \log K_0 + s(\beta - \delta) - (t - s - 1)\delta + \beta \sum_{j=s}^{t-1} s_j. \tag{12}$$

We now assume as Mincer does that the proportion of time spent to writing articles linearly decreases till retirement so that with an experience $x = t - s$ this proportion is given by:

$$s_{x+s} = \kappa(1 - \frac{x}{T}). \tag{13}$$

As $\sum_{j=0}^{x-1} 1 - j/T = x(1 + 1/2T) - x^2/2T$, we get

$$\log K_t = \log K_0 + s(\beta - \delta) - (x - 1)\delta + \beta(x(1 + 1/2T) - x^2/2T). \tag{14}$$

The log of the stock of articles of a scientists is a function of initial conditions, a term related to the initial formation (represented by $\log K_0 + s(\beta - \delta)$), a trend and a squared trend.

This log linear equation is a good start for estimation. However, it first describe the behaviour of a single scientist and second does not take into account the collective aspect of scientific research.

2.3 Individual and cohort effects

For the while, our equation looks like a Mincer equation, except that the dependant variable is the log of the stock of publications at the end of the period of observation. We have

$$\log K_i = \alpha_0 + \rho s_i + \beta_1 x_i - \beta_2 x_i^2,$$

where s_i could measure the number of years of schooling which could be the number of years needed to complete a PhD. It is more interesting to use this variable to introduce a vintage effect (date of PhD). Because of the secular progress of Science, a recent PhD could be thought of being more productive than a PhD coming from an older cohort. It could be completed by an indication on the place where the PhD was prepared. Vintage effects are usually difficult to identify because of the linear relation

$$\text{calendar date} = \text{experience} + \text{Vintage}.$$

So calendar, vintage and experience cannot be distinguished. A solution is to take vintage as a nonlinear function of calendar date by defining Vintage as a nonlinear function, most of the time a step function defining intervals of several years, for instance four years.

The term α_0 corresponds in fact to the initial stock of human capital and thus to an individual effect. Levin and Stephan (1991) introduced fixed individual effects. This was possible because their endogenous variable is the log of the yearly production. There are thus several observations of the same individual. Here we want to explain a stock at the end of a period of observation. So we have a single observation per individual. A fixed effect is thus not identified. A random individual effect is simply the residual term of the equation in our case. Thus α_0 will be simply related to the mean stock of human capital.

3 Individual and collective behaviour

The reasons for accumulating human capital were strictly determined on an individual basis. No mention was made to institutions, surroundings, collective behaviour. We know however that contrary to physical capital, human is expandable and self generating with use. It is transportable and shareable. Its efficiency depends on sociological factors, i.e. the way these intellectual capacities are organised and shared in an adequate surrounding. This aspect is particular well illustrated by the notion of social capital.

3.1 Social capital

Social capital can receive two types of definitions, which can be opposed as discussed in Siisiäinen (2000). For Bourdieu (1980), Bourdieu (1986), the social capital is the value of an individual social network. This network is used as a resource in social struggle. On the contrary, for Putnam (1995), social capital is a collective good made of moral obligations and norms, social values and social networks. A society with a high level of social capital is an integrated society, functioning on trust and collaboration. Coleman (1988) demonstrated the importance of social capital in the process of human capital accumulation using the example of education. He pointed out the influence of a collaborative attitude of parents helping children for their homework to illustrate the influence of social capital at the family level. At the public level, he noted the very small dropout rate of students in catholic schools compared to public schools, explaining this by a common ideology of solidarity. Bourdieu on the contrary is very skeptical with respect to altruistic actions. They cannot be free of any specific interest of the actor. His position is thus totaly opposed to Putnam's and Coleman's romantic ideas of generalised trust.

If the theory of human capital has been widely used (and criticised) for explaining scientific production, that of social capital was rarely applied in the field of economics of science as underlined in Bozeman, Dietz, and Gaughan (2001) who propose to use it when evaluating research. See that paper for more details and the references it gives.

3.2 Scientific collaboration or scientific competition?

The two competing notions of social capital have an immediate transcription in the economics of science in term of collaboration or competition inside a scientific institution. Researchers are regrouped into departments and universities which characteristics have an impact on research output. Aghion, Dewatripont, Hoxby, Mas-Colell, and Sapir (2007) have tried to make up a list of different institutional factors such as governance variables illustrated for instance by autonomy in wage setting, or more direct variables like the number of students, the number of academic and the total budget. We are interested here in variables which could characterise the degree of collaboration (the Putnam's way) and of human capital sharing inside the institution. At the department level, this will be the number of leading scientists in the department, the size of the department and the proportion of papers scientific leaders write with members of the department.

In the struggle of publish or perish (the Bourdieu's way), individuals have several possible strategies. They can decide to publish alone, keeping their ideas for themselves and thinking that PhD students are a waste of time. But this strategy is rather rare, because of the risky nature of scientific activity. Having co-authors is a kind of risk-adverse behaviour. They can look for foreign co-authors who can be thought of being of a higher level or decide to engage in collaboration inside their department. Their final choice is the targeting of their publications: national, international or top level (a short list of top journals). This last choice can be interpreted a cultural norm, a kind of label which is an entrance ticket for social networks. For instance to have published a paper in Nature or Science.

3.3 Social capital in scientific institutions

This form of social capital is characterised by social values and social organisation that contribute to the value produced and constitute useful capital resources for individuals. Coleman (1988) details three forms of social capital: obligations and expectations, information channels, and social norms that we shall now try to investigate in the special case of universities and the production of scientific knowledge. We have also three components in Putnam's concept of social capital: social values such as trust, social networks (especially voluntary associations) and moral obligations and norms (see Siisiäinen (2000)).

1. *obligations and expectations, trust and cooperation*: you help somebody once and you expect that in the future he will help you in the same way if necessary. Mutual reciprocity and trust. At a department level, you discuss with colleagues concerning a problem where you need help. That

colleague gives you help and ideas. He then might expect either the same kind of service in the future or being the co-author of your work if the discussion goes far enough. Relations between professor and PhD students also enter this category. The PhD student expect good guidance, the professor expect good work and outside recognition.

2. *Information channels*: A community facilitates communication and general information. Collecting information is costly while being strategic, the group can provide it. In scientific communities, it is impossible to read all journals in order to maintain up to date information: the knowledge of the most recent research. Institutions organise seminars, conferences in order to diffuse this information. This lead to meet researchers outside the department and to write papers with them.

3. *social norms*: what is socially accepted and what is socially forbidden. What is imposed to the individual by the community in order to behave according to public interest and not according to personal interest. What is good scientific practice in relation with other scientists? Plagiary, scientific forgery have always existed, but always been condemned. A positive behaviour, always encouraged is scientific cooperation and PhD guidance. But the norm can go further such as imposing a certain type of publications. The importance of books is for instance declining in economics at the advantage of articles.

4 Measurement

Most, if not all, of the information we are looking for is contained in bibliographical data base provided they contain adequate affiliation data. In this paper, we shall use the ECONLIT data base, and thus restrain ourselves to the economic profession. Bibliographic data bases give information that goes far beyond counting publications, provided we exploit them on a relatively long period. But data have to be rearranged and processed in order to produce that information. A basic record is formed of one paper that has always the same attributes: one or more authors, their affiliations, a bibliographic source given by the name of a journal together with page numbers and finally a date of publication. This record has to be dispatched in as many records as there are co-authors for a paper, because we are now interested, not in the particular content of the paper, but in the strategy of publication of each author: the collaborations he had, the choice he made for publishing his paper and the score he can gain from that paper. We are also interested in institutions, to quantify the degree of collaboration within an institution, the presence of leaders and their impact on the other members of the institution.

Our data cover the period 1991-2007 which represents a maximum experience of 17 years. We could not consider a longer period, because affiliations are not reported before 1991. For more details on this data base, see Lubrano, Bauwens, Kirman, and Protopopescu (2003).

4.1 Measuring individual scores

The stock of human capital is difficult to define and to measure. McDowell (1982) or Levin and Stephan (1991) measured it as a stock of publications or a stock of citations. We have chosen here the stock of publication over a given period, using the van Damme (1996) formula:

$$K_i = \sum_{j=1}^{n_i} \frac{b(p_j)}{a(p_j)} v(p_j),$$

where $b(p)$ is a number related to the length of the publication, $a(p)$ is a number related to the number n_j of authors of the publication, $v(p)$ is related to the quality of the publication. n_i represents the total number of publications that author i is credited over the whole period. See Lubrano, Bauwens, Kirman, and Protopopescu (2003) for details. We have however, and contrary to Lubrano, Bauwens, Kirman, and Protopopescu (2003), taken $a(p) = 1$ because the number of authors can be used in the list of explanatory variables of our model. And it cannot appear on both sides of the regression.

4.2 Measuring experience

We have only a portion of the total trajectory of a researcher. Levin and Stephan (1991) have access to a list of PhD recipients together with the date of their PhD. They then look at those who hold an academic position in a top five US universities. This creates a selection bias for which they have to introduce a correction. Our data are not subject to sample bias because we consider the whole sample of person who have published at least one article. But, contrary to Levin and Stephan (1991), we do not know when a person started his academic career or his status (we have discarded individual who had no university affiliation). What we know is his first date of publication FPY_i and his last date of publication LPY_i. We shall thus measure experience e_i with:

$$e_i = LPY_i - FPY_i + 1,$$

and the starting year of carier as FPY_i. As we are considering a specific period of time (1991-2007), there is a censoring problem for authors who started their career before 1991. We do not have then a correct measure of their total experience or for the date of their PhD. However, the proportion of authors with an experience of 17 years represent 0.3% of the sample in the UK and 3.1% in the Netherlands. Only a fraction of these will have a longer and thus censored experience.

The distribution of the first publishing years for the UK was roughly uniform between 1991 and 2002. It sharply increased after 2002, indicating a large inflow of new young authors coupled certainly with a larger coverage of the data base. This feature is common to all European countries.

4.3 Measuring individual strategies

Individual strategies of publication are relatively easy to describe when looking at data bases. We constructed a first variable called $P10_i$ which measures the number of articles that an author has published in a short list of top journals. The six top journals in economics are supposed to be *American Economic Review, Econometrica, Journal of Economic Theory, Journal of Political Economy, Quarterly Journal of Economics, Review of Economic Studies.* They are graded 10 on the scale used in Lubrano, Bauwens, Kirman, and Protopopescu (2003).[2] An author having achieved to get such publications is supposed to be a potential leader in his institution.

At the other extreme, we have identified authors who publish in national journals. We have computed Pnj_i as the proportion of articles that an author has published in national journals. We can interpret this variable as the characteristic of a person that favour national networks.[3]

Mobility which was identified by Putnam (1995) as a factor explaining the decrease in US social capital. Mobility of researchers. Variable to measure this.

4.4 Measuring collaborations

As research is a risky activity, there is a secular tendency to publish papers with a greater number of coauthors. But this choice is not uniform among the disciplines and the number of coauthors is still a decision variable reflecting a particular type of collaboration or absence of collaboration.

We have defined Pal_i as the proportion of papers that an author has written alone, reflecting thus the absence of collaboration, of participation to the social capital of his department. Conversely, Psi_i measures the proportion of papers that an author has written with at least one coauthor of the same institution. It is a measure of participation to the social capital of the institution. Pin_i measures the proportion of papers that an author has published with at least one foreign coauthor. This last variable measures international cooperation, but also the belonging to social networks, designed to improve one's personal situation.

4.5 Measuring social capital at the institutional level

Social capital is a collection of social relations inside a department that facilitate individual scientific production. These relations are not easy to measure. We have tried to build up variables at the department level that can have an influence on individual results.

- The presence of a leader can have a tremendous effect, both by attracting other top researchers and by supervising PhD students. We can identify a top researcher and perhaps a leader by noting if he belongs to the small circle of authors having published at least one paper in a top journal. The

[2] We could also have chosen the ten top journals in term of total citations in 2008, i.e. *American Economic Review, Econometrica, Journal of Political Economy, Quarterly Journal of Economics, Journal of Financial Economics, Journal of Econometrics, Review of Economics and Statistics, Review of Economic Studies, Economic Journal, Journal of Economic Theory.*

[3] A national journal can be easy to define, just by looking at the language it uses. But sometimes a national journal turned for English as *Economic Notes* in Italy or *Journal of Economics* in Germany or *Spanish Economic Review* in Spain. For the UK, the matter is more complex. We have to look deeper into the journal. A journal has a national coverage if it serves as a major means of diffusion for national authors. It is considered as mainly national if in addition it does not serve as a major means of diffusion for other countries. A more precise definition is given in Lubrano, Bauwens, Kirman, and Protopopescu (2003).

variable $N10_j$ will measure the total number of active top researchers in a department.

- Interactions between authors can be favoured by the size of a department. Below a minimal size, the possibilities of cooperation are nearly zero. But cooperation can be also impeded by the anonymity created by a too large number of colleagues. We call NZ_j the total number of publishing authors affiliated to department j.[4]

- The degree of cooperation can be measured by the proportion, at the department level of papers that are produced alone and by the proportion of papers that are written with at least two authors belonging to the same department.

- The turn over of the top members. When an author is a top researcher (as defined before), does he remains in the same institution or has he changed recently, by the way predicting thus a next change.

5 Hierarchical multilevel models

Hierarchical linear models were designed to model observations that are regrouped into clusters. The main domain of application in economics is school achievement where individuals are scholars endowed with a measure of schooling performance (see the US *High School and Beyond* data set from *The National Center for Education Statistics* and the survey article of Goldstein and Spiegelhalter (1996)). Hierarchical linear models will be the device that we shall use to mix together the individual research behaviour given by the life cycle model with the theory of social capital that consider the individual in his environment. Coleman (1988) says that "The conception of social capital as a resource for action is one way of introducing social structure into the rational action paradigm". He illustrates his demonstration using an example concerning high school dropouts. The same data are analysed, with similar conclusions in Raudenbush and Bryk (2002) using linear hierarchical models. We shall follow the same route, showing how individual decision variables coming from the human capital theory can be combined with proxies illustrating the social capital of an institution to explain how individual scientific production can be influenced by the characteristics of scientific surroundings.

5.1 Models with fixed and random effects

Let us consider log of the total score K_{ij} of an individual i belonging to institution j. Our regression model is

$$\log K_{ij} = \beta_0 + x'_{ij}\beta + v_{ij} \qquad v_{ij} \sim N(0, \sigma^2). \tag{15}$$

The log of the total score of an individual is explained by a constant term and a set of exogenous variables or predictors, all observed at the individual level. These variables are on one side the life cycle variables such as experience and

[4]In fact, NZ_j indicates the total number of active members of a department, those that have at least published one paper in the last four years and whose last affiliation is department j.

squared experience, the initial conditions represented here by cohort effects, and on the other side variables representing the individual network and publishing habits.

Authors are regrouped in departments (or clusters) where they can share unobserved common individual features, due for instance to a particular recruiting policy but also where they share a common social capital. The first way of introducing the possibility of a department effect is to consider a specific constant term per institution called β_{0j}. But a fixed effect, first cannot introduce correlation between individuals and second might cause inference problems (multi-collinearity) when there are many departments and few individuals per department. The correct solution consists in supposing that the β_{0j} are random and independent of the v_{ij}, so that we have a hierarchical linear model with

$$\begin{aligned} \log K_{ij} &= \beta_{0j} + x'_{ij}\beta + v_{ij} \quad && v_{ij} \sim N(0, \sigma^2) \\ \beta_{0j} &= \beta_0 + u_j && u_j \sim N(0, \omega^2), \quad u_j \perp v_{ij}. \end{aligned} \tag{16}$$

The correlation between individuals inside the same department is equal to

$$\rho = \frac{\omega^2}{\omega^2 + \sigma^2}. \tag{17}$$

The higher this correlation, the higher will be the unobserved sharing of a common social capital.

The interpretation of constant terms in this model depends heavily on the metric which is used for measuring the predictors. β_0 represents the average score of all the departments, when all predictors are set equal to zero. This particular value of zero might be meaningful for most variables, but certainly not for experience which has to be strictly positive. In order to recover a clear interpretation for the constant term, the predictors are usually centered around a common value, usually their sample mean. We have the choice between centering around the grand mean, which is here the country mean or around the local mean which is the institution or cluster mean.

- When the predictors are centered around the grand mean, β_0 represents the country average of the log scores for authors whose characteristics are equal to the country average. The reduced form of the model is

  $$\log K_{ij} = \beta_0 + (x'_{ij} - \bar{X}'_{..})\beta + v_{ij} + u_j. \tag{18}$$

  $\bar{X}_{..}$ represents the mean of x_{ij} taken with respect of both i and j. This interpretation is obtained because $\sum_{ij}(x'_{ij} - \bar{X}_{..})/n = 0$.

- If we now center the predictors around their local mean, $\beta_{0j} = \beta_0 + u_j$ represents the cluster mean of the log individual scores when the predictors are taken equal to their local mean. More precisely:

  $$\log K_{ij} = \beta_{0j} + (x'_{ij} - \bar{X}'_{.j})\beta + v_{ij} \tag{19}$$

  where $\bar{X}_{.j}$ is a vector of empirical means computed over i for a given j. This interpretation is obtained this time because $\sum_i(x'_{ij} - \bar{X}_{.j})/n = 0$.

How to center the predictors constitutes a large debate in the applied literature, see for instance Raudenbush and Bryk (2002) page 31.

The experience predictor e_{ij} has a particular status because it is an expression of time. So usually, this variable is not centered around its mean (local or global), but around a given value L. We have chosen to center it around $L = 1$, so the obtained mean score will be that of authors having one year of experience, which is the modal value of many countries of our sample.

When the proper method for centering variables is chosen, we note our model as

$$\begin{aligned} \log K_{ij} &= \beta_{0j} + \tilde{x}'_{ij}\beta + v_{ij} \\ \beta_{0j} &= \beta_0 + u_j \end{aligned} \tag{20}$$

where $\tilde{x}'_{ij}$ indicates the centered predictor. Due to the particular structure of the error term, these models have to be estimated using either iterated GLS or the EM algorithm.

5.2 Introducing variables at the department level

Cooperative social capital variables are introduced at the department level. They modify the potentialities of the individuals, either their mean score of their life cycle. We regroup in z_j these surrounding variables and center them around their country mean $\bar{Z}$ so that $\tilde{z}_j = z_j - \bar{Z}$. A first version of the enlarged model is

$$\begin{aligned} \log K_{ij} &= \beta_{0j} + \tilde{x}_{ij}\beta + v_{ij} \\ \beta_{0j} &= \beta_0 + \tilde{z}_j\gamma_0 + u_{0j}. \end{aligned} \tag{21}$$

This model says that the mean individual score varies around a country mean β_0 according to individual characteristics $\tilde{x}_{ij}$ and around the mean department score β_{0j} which is itself influenced by department characteristics.

We could further say that the yield of experience can vary across departments and be influenced by department variables. This amount to add a second random effect. Let us call $\tilde{e}_{ij}$ experience centered around L. The final model of this section is

$$\begin{aligned} \log K_{ij} &= \beta_{0j} + \tilde{e}_{ij}\beta_{1j} + \tilde{x}_{ij}\beta + v_{ij} \\ \beta_{0j} &= \beta_0 + \tilde{z}_j\gamma_0 + u_{0j} \\ \beta_{1j} &= \beta_1 + \tilde{z}_j\gamma_1 + u_{1j}. \end{aligned} \tag{22}$$

Both u_{0j} and u_{1j} are independent of v_{ij}. But u_{0j} and u_{1j} can be correlated. If u_j is the vector formed by the concatenation of u_{0j} and u_{1j}, we have

$$u_j \sim \mathrm{N}(0, \Omega). \tag{23}$$

u_{0j} indicates how much the mean log score of department j deviates from the grand mean β_0. u_{1j} indicates how much the average yield of a supplementary year in department j deviates from its country average β_1. A positive correlation would mean that the higher the average score of the department is, the higher the return to one year of experience in the same department.

Model (22) can be expressed in a reduced form which is convenient for estimation:

$$\log K_{ij} = \beta_0 + \tilde{z}_j\gamma_0 + \tilde{e}_{ij}\beta_1 + \tilde{e}_{ij}\tilde{z}_j\gamma_1 + \tilde{x}_{ij}\beta + u_{0j} + \tilde{e}_{ij}u_{1j} + v_{ij}.$$

Table 1: Sample characteristics

Country	Authors	Active academics	Stopped publishing	Unlisted	Median e_i	Percentage active $e_i = 17$
Belgium	1829	919	39%	10%	1	2.83
France	8115	3595	33%	23%	3	2.23
Germany	6181	2505	39%	20%	1	2.20
Italy	4908	2609	31%	15%	2	2.11
Netherds	3982	1999	38%	12%	2	3.05
Spain	5730	3536	30%	8%	1	0.62
Sweden	1935	952	39%	11%	2	1.90
Swiss	2156	819	38%	24%	1	1.50
UK	14361	7666	37%	10%	2	0.30

We note the very complex structure of this reduced form error term $u_{0j}+\tilde{e}_{ij}u_{1j}+v_{ij}$.

6 Empirical results

We have estimated our model using data collected from the JEL CD-ROM covering the period of 1991 to 2007 for nine European countries. We discarded authors who have ceased publishing, which means authors that have not published articles for the last four years. We also discarded authors who were unlisted or not affiliated to an academic institution. Characteristics of the sample are given in Table 1. The original sample contained five additional small countries Austria, Denmark, Finland, Greece, Ireland, which we discarded for ease of presentation.

6.1 Testing for the life-cycle model at the country level

We have estimated a regression model with a single random effect on the constant term, country by country, where experience was taken in deviation to 1 and the other predictors in deviation to their local mean. Consequently, the constant term measure the average country score (geometric mean because of the logs), so that countries can be ranked. We provide in Table 2 an estimation of the fixed effects.

The life cycle model seems to be validated for all the countries. The variables e and e^2 are very significant for all the countries. The coefficient of e is positive and has a value between 0.22 and 0.26 while the coefficient of e^2 is always negative between -0.01 and -0.005. *We have not yet introduced cohort effects.*

Differences in individuals strategies are very important to complete this model. Concerning publishing norms, choosing top journals has always a positive impact, compared to the very negative impact of publishing in national journals. This last effect is not significant for the Netherlands, Sweden and Switzerland, a reason being that national journals might not be very developed in these countries.

Concerning cooperation versus competition and personal networks, we notice that publishing alone is an inefficient strategy while publishing with a foreign

Table 2: Estimation of fixed effects

country	cste	P_{10}	Pnj	Pal	Pin	Psi	e	e^2
Netherld	1.21	0.38	-0.04	-0.27	0.06	-0.08	0.24	-0.005
	23.3	11.9	0.30	-4.35	1.13	-1.56	23.3	-6.98
UK	1.16	0.31	-0.21	-0.24	-0.04	-0.14	0.24	-0.005
	49.6	22.0	-3.77	-8.19	-1.61	-5.62	42.2	-12.1
Swiss	1.15	0.31	3.15	-0.22	0.20	-0.28	0.25	-0.008
	21.0	8.77	1.74	-2.13	2.45	-3.30	14.9	-6.11
Belgium	1.11	0.36	-0.56	-0.076	0.071	-0.18	0.23	-0.0048
	16.0	8.45	-6.24	-0.78	0.96	-2.31	14.9	-4.34
Sweden	1.08	0.42	-0.49	-0.14	0.05	-0.12	0.25	-0.007
	17.2	7.81	-1.69	-1.64	0.71	-1.62	17.2	-6.30
Germany	1.08	0.42	-0.37	-0.15	0.10	-0.12	0.26	-0.008
	35.4	13.7	-7.88	-2.83	2.20	-2.70	27.3	-11.8
France	1.00	0.23	-0.16	-0.25	0.04	-0.14	0.24	-0.007
	28.9	19.5	-5.85	-6.95	0.95	-4.24	37.1	-13.8
Italy	0.99	0.36	-0.28	-0.21	0.24	-0.18	0.22	-0.006
	40.6	15.6	-6.54	-4.69	5.73	-4.26	27.0	-10.4
Spain	0.96	0.36	-0.39	-0.04	0.17	-0.05	0.26	-0.01
	28.1	16.2	-14.3	-0.90	4.64	-1.46	38.3	-15.0

For each country, the second line represents t statistics. These results were obtained using the procedure MIXED of SAS. Countries are ranked by average total score represented by the constant term.

coauthor (personal network) can be fruitful in Italy, Spain and Switzerland.

Concerning the participation to the social capital of the institution, if it is measure by the variable Psi (all authors in the same institution) has a negative impact on individual scores.

6.2 Weak random effects

In Table 3, we have given an estimation of the variances of the two error terms, v_{ij} and u_j, computed the intra class correlation and indicated the P-Value P_Z corresponding to the test that there is no random effect $(\text{Var}(u_j) = 0)$. Clearly, we need a lot of data to identify a random effects. There is no random effect in Switzerland, but this also the country for which we have the smallest number of observations. The correlation inside departments is small on average, but in accordance with other data sets on secondary education. Without further indications, cooperation seems to be weak.

6.3 Testing for the social capital model

We are now going to introduce variables at the department level. The paper is still preliminary, so that some of the variables we wanted to include are not yet available. We have two variables: the number of top researchers who could act as leaders and thus improve the production of their colleagues and the size of the department. These variables are going to influence the mean individual score and the yield of experience. We have thus now two random effects. The model

Table 3: Random effects

country	Departments	Authors	σ_u^2	σ_v^2	ρ	P_Z
Belgium	8	904	0.0305	0.412	0.069	0.05
France	56	3586	0.0564	0.362	0.135	0.00
Germany	57	2424	0.0314	0.415	0.070	0.00
Italy	47	2512	0.0122	0.340	0.035	0.00
Netherld	12	1959	0.0263	0.435	0.057	0.02
Spain	49	3415	0.0442	0.294	0.131	0.00
Sweden	17	903	0.0443	0.383	0.104	0.01
Swiss	14	790	0.0246	0.420	0.055	0.08
UK	79	7260	0.0285	0.436	0.061	0.00

incorporate in this way social capital variables to explain individual scores and individual productivity as they ware initially described by the human capital model.

The key variable is undoubtedly the number of top researchers in a department which has always a positive effect on the individual log total score. Department predictors have a noticeable interaction with the random effects. They can reduce so much the variance of some of the u_j that the latter can be considered as zero for some countries. We shall present our estimation results by dividing the countries in two groups.

Table 4: Department effects (Constrained models)

country	γ_{01} N_{10}	γ_{02} NZ	γ_{11} $N_{10} * e$	γ_{12} $NZ * e$	σ_v^2	ω_0	ω_1	ω_{01}
France	0.024	-0.0010	0.0022	-0.00009	0.34	0.011	0.00029	0.0013
	6.55	-2.30	3.68	-1.45	41.8	2.76	2.61	2.77
Germany	0.019	-	-	-	0.40	0.0039	0.00057	0.0014
	5.01	-	-	-	34.6	1.06	2.46	2.18
Italy	0.015	-	0.0028	-	0.33	0.0023	0.00008	0.0
	3.69	-	3.42	-	35.6	1.27	1.78	-
Spain	0.017	-0.0007	-	-	0.28	0.0061	0.0012	0.0027
	10.0	-3.38	-	-	41.5	1.87	2.84	3.14
UK	0.007	-	-	0.00007	0.43	0.012	0.0004	0.0012
	4.46	-	-	2.54	59.8	2.87	3.18	2.85
Belgium	0.018	-	-	-	0.41			
	7.37	-	-	-	21.4			
Netherld	0.019	-0.0006	-	0.0002	0.43			
	8.00	-2.06	-	3.69	31.5			
Sweden	0.018	0.0020	-	0.0002	0.38			
	5.13	2.86	-	2.01	21.7			
Swiss	0.016	-	-	-	0.43			
	3.11	-	-	-	20.1			

γ_{01} and γ_{02} measure respectively the influence of $N10$ and of NZ on the local mean. γ_{11} and γ_{12} play the same role for the local yield of one extra year of experience.

- For the five large countries, we could identify two random effects which are positively correlated.

- For the small countries, the addition of variables at the department level killed the random effects whenever there was one. These variables have thus only a fixed effect.

- In large countries, the two random effects are positively correlated. The mean score of a department is highly correlated with the mean productivity of its members. The correlation is 0.73 for France, 0.98 for Germany, 0.99 for Spain, but only 0.55 for the UK. It is not significantly different from zero for Italy.

- One extra leader (top publisher) in a department increases individual scores by $\gamma_{01} = 2\%$ on average, even in small countries. But by less than 1% for the UK.

- The small effect of leaders on average score in the UK, coupled with the small correlation of 0.55 could indicate a comparative lack of sharing of social capital between members of UK departments.

- The influence of top researchers on the mean productivity of their department colleagues is positive for France and Italy ($\gamma_{11} = 0.2\%$ on average). It is not significant elsewhere.

- The effect of the size of the department is negative for France, Spain and the Netherlands where an extra member lowers the individual score by $\gamma_{02} = -0.1\%$ on average. In France, the department size influences negatively also productivity ($\gamma_{12} < 0$). In the UK, the Netherlands and Sweden, the size effect on productivity is positive.

Let us come back to the size effect. In Table 5, we give the list of countries were this effect was significant for explaining the mean department scores. Countries that have very large departments have a tendency to have lower mean

Table 5: Largest departments characteristics

country	Largest department	size	N_{10}	γ_{02}
France	Paris I	343	17	-0.0010
Spain	Complutence Madrid	254	1	-0.0007
	Valencia	253	2	
Netherlands	Erasmus	322	11	-0.0006
Sweden	Stockholm U	130	20	0.0020

department scores. There is a minimum size, under which no efficient research can be led. But a department which is too large becomes inefficient. This militates against the Napoleonic model of organisation for universities, model at work in France, Spain and Italy. This militates also against the recent movement in France for regrouping universities to constitute very large institutions. Stockholm U is the largest institution in Sweden. It has only 125 reported members, but 20 top researchers. This explains the positive sign of γ_{02} for this country.

7 A three level model for European countries

Despite the apparent large amount of individual observations, we experience a lack of significance for many variables in small countries. So efficiency could be gained by pooling the data in a larger model. Pooling data also allows to introduce new variables, this time at the country level. This is useful for analysing the influence of economic policy variables. For this exercise, we have kept the five small countries: Austria, Denmark, Finland, Greece and Ireland.

7.1 Variables at the national level

In most European countries, universities have a public status or get most of their budget from the state. The total number of universities in a country is a public decision, but also the budget per student or even the percentage of R%D in total GNP. From various sources, we have collected the following data given in Table 6. When reported to the population, there are countries which

Table 6: Miscellaneous data for main European countries

country	Pop (millions)	R and D intensity % GNP	Euros per student (thousands)	Econ departs
Austria	8.1	2.19	11.5	12
Belgium	10.3	2.33	12.2	16
Denmark	5.4	2.60	26.0	8
Finland	5.2	3.51	10.6	18
France	59.2	2.19	8.7	70
Germany	82.2	2.50	12.4	98
Greece	10.9	0.64	3.3	12
Ireland	3.8	1.12	11.3	8
Italy	57.8	1.16	7.0	72
Netherld	16.0	1.89	15.7	10
Spain	39.8	1.11	5.2	48
Sweden	8.9	4.27	20.7	21
Swiss	7.6	2.93	26.2	11
UK	60.0	1.87	11.5	96

Budget data result from computations made in Aghion, Dewatripont, Hoxby, Mas-Colell, and Sapir (2007). R&D intensity was collected from the European Commission, Research Division web site. Other data were previously collected in Lubrano, Bauwens, Kirman, and Protopopescu (2003).

have many economic departments. On average, there are 1.34 departments per million inhabitants. Some countries are below that figure such as France, Germany, Greece, Italy, Spain and mainly the Netherlands. In those countries, the size of the departments is large. Some other countries are much over that figure (1.34) such as Belgium, Finland, Ireland, Sweden and the UK. This would correspond to smaller departments.

On average, European countries are spending 12 thousands euros per student. Many countries including the UK are within that figure. All southern European countries are below or much below. Denmark, the Netherlands, Sweden and Switzerland are well over. We note that this variable results from calculations in Aghion, Dewatripont, Hoxby, Mas-Colell, and Sapir (2007) who adjusted published data for PPP. There can remains some suspicion on this variable. For instance we know that in France the spending per student is of 6 200 euros for universities. It can be higher for specialised schools (the grandes écoles). Table 6 indicates 8 700 euros. This is slightly too high for universities, but still reasonable after PPP adjustment. The figure given for Greece is 3 300 euros, which seems clearly too low.

7.2 Random effects, exchangeability and country level

Random effects imply a very particular assumptions on clusters. We have to suppose that the observed clusters come from a larger population and that they are observed at random. According to the expression coined in Lindley and Smith (1972), clusters are said to be exchangeable. It other words, the order in which they have been observed is not important, the same for any group of clusters. This assumption can be tenable in the long term for universities. For instance some universities were created quite recently such as Warwick (1965), Maastricht (1976) or Paris Dauphine (1968). Or some departments got an international fame only very recently such as Toulouse or Tilburg. This means that new top departments can appear randomly in the future.

The assumption of exchangeability is not tenable for countries which are fixed entities. This means that the γ coefficients cannot be made random across countries. The sole differences that can be introduced between countries are fixed effects by means of country predictors. Let us call w_k the variables that are specific to country k, but common to all universities and individuals belonging to that country. We extend model (22-23) with country effects so as to obtain:

$$\begin{aligned} \log s_{ijk} &= \beta_{0jk} + \tilde{e}_{ijk}\beta_{1jk} + \tilde{x}_{ijk}\beta + \tilde{w}_k\delta + v_{ijk} \\ \beta_{0jk} &= \beta_0 + \tilde{z}_{jk}\gamma_0 + u_{0jk} \\ \beta_{1jk} &= \beta_1 + \tilde{z}_{jk}\gamma_1 + u_{1jk}. \end{aligned} \tag{24}$$

The x_{ijk} continue to be centered at the department level, whatever the country, because we want to keep the interpretation of the β_{0jk} as being the mean score of a department, whatever the country. We could discuss at which level we should center the z_{jk}, the department variables. We could continue to center them at the country level. That could make sense in a full three level model. Here, as we have no random effect at the country level, centering these variables without taking care of the country structure make sense. The third level is represented only by the $w_k\delta$ variables. They are constant over the departments and the countries. In order to keep a clear interpretation of the β_{0jk}, we have to center them around an European mean.

This model is rather restrictive, because we restrict the individual fixed effects represented by the β to be the same across countries. We saw in the previous sections that there can be huge differences. As we cannot have random effects at the country level as explained above, the only solution is to increase the number of random effects at the department level. Systematic differences between countries are represented solely by the w_k.

7.3 Inference on fixed effects at the European level

By pooling countries, we obtain a total number of 26 676 individual observations grouped in 398 departments. We had to introduce a random effect on five of the individual variables. We now detail inference results, presenting the model in its structural form:

$$\begin{aligned}
\log s_{ijk} &= \beta_{0jk} + \beta_{1jk}Pal + \beta_{2jk}Pin + \beta_{3jk}Pnj - \underset{[-8.70]}{0.11}\,Psi \\
&\quad + \beta_{4jk}P_{10} + \beta_{5jk}e - \underset{[-36.5]}{0.0076}\,e^2 \\
&\quad + \underset{[7.59]}{0.012}\,EurosSt + \underset{[3.48]}{0.00085}\,NEco + v_{ijk} \\
\sigma_v^2 &= 0.377
\end{aligned}$$

$$\begin{array}{rcll}
\beta_{0jk} &=& \underset{[33.7]}{0.92} + \underset{[8.74]}{0.013}\,N_{10} - \underset{[-2.30]}{0.00033}\,N_Z + u_{0jk} & \omega_0 = 0.013 \\
\beta_{1jk}(Pal) &=& -\underset{[-11.0]}{0.17} - \underset{[-3.03]}{0.00034}\,N_Z & - \\
\beta_{2jk}(Pin) &=& \underset{[4.36]}{0.070} + u_{2jk} & \omega_2 = 0.016 \\
\beta_{3jk}(Pnj) &=& -\underset{[-13.3]}{0.27} - \underset{[3.21]}{0.0082}\,N_{10} + u_{3jk} & \omega_3 = 0.026 \\
\beta_{4jk}(P_{10}) &=& +\underset{[19.4]}{0.54} - \underset{[3.64]}{0.011}\,N_{10} + u_{4jk} & \omega_4 = 0.078 \\
\beta_{5jk}(e) &=& +\underset{[78.1]}{0.25} + \underset{[5.18]}{0.00011}\,N_Z + u_{5jk} & \omega_5 = 0.00058
\end{array}$$

These results confirms the results we found before, but provides a better precision. The individual strategy of publishing alone decreases your total score of 17% while publishing with an international co-author increases it of 7%. But choosing all your co-authors in the same institution decreases your score of 11%. So in term of collaboration, it is better to have co-authors and to share one's human capital, but it is better to share it with at least one outside member. The individual strategy of collaboration and participation to the increase of social capital is not evident.

Individual publication strategies keep the same profile. Publishing in national journals is equivalent to loose roughly one year of productivity. On the contrary, one more top paper increases your score of 54%. These results have a clear interpretation. The cost of the habit of publishing in national journals is terribly high in term of scientific diffusion. This should be put in parallel with the findings of Bauwens, Mion, and Thisse (2007) concerning the use of English as a scientific vehicle. But English is not the sole concern, international cooperation is its natural corollary. If a department is not a good place for choosing one's coauthors, it is a good place for attracting foreign visitors and sharing human capital, and perhaps also by this way increasing the social capital of the institution. These results should be remembered when negotiating funds for inviting visitors.

There is definitively a social capital effect at the department level because first of correlation between random individual effects and second because of the presence of collective variables. Top researchers have a strong influence on their surroundings. Their presence and their number increase the mean score of their department by 1.3%. They have a positive effect on the members who use national journals in a proportion below that of their department, but the reverse for the others. They have a positive effect on the other top publishers, provided these have published a number of top papers below the department mean. Otherwise, their effect is negative (the abuse of competition?).

A bigger size for a department has a negative impact on its mean score and on the mean score of authors publishing alone. The department size has a positive effect on experience. The department size manages to have a global positive effect only for those with an experience greater than 4 years and who publish alone in a proportion lower than that of their department.

At the national level, it seems better to have more economic departments than fewer but bigger ones. The number of euros spent per student is determinant. The yield of a one euro increase in the budget per student is 1.1 per cent in term of national research output in the long term.

Let us now examine the correlation matrix coming from Ω, which is the variance-covariance matrix between the random effects explaining the random differences of individual coefficients between departments.

We still have a positive correlation between the mean score and experience, with 0.69 at the department level. The correlation between the mean score effect and the National Journals effect is negative and rather strong with -0.41. So it is bad for a department to promote the social norm of publications in national journals. There is a negative correlation of -0.35 between mean experience and the average number of top papers in a department. When members of a department get older or more experienced, there is a tendency for that department to publish less top papers. Finally, there is a negative correlation of -0.34 between the average proportion of international co-authors and the average number of top papers in a department. This mean that the individual strategy of having foreign coauthors is efficient to increase the total score, but not for increasing top publications. All other correlations are not significantly different from zero.

Table 7: Variance covariance matrix of random effects

	β_0	e	Pnj	Pin	P_{10}
β_0	0.013				
	[0.00]				
e	0.69	0.00058			
	[0.00]	[0.00]			
Pnj	-0.41	0.0	0.026		
	[0.01]	[0.54]	[0.00]		
Pin	0.00	0.00	0.00	0.016	
	[0.16]	[0.82]	[0.67]	[0.00]	
P_{10}	0.00	-0.35	0.00	-0.31	0.078
	[0.99]	[0.00]	[0.16]	[0.04]	[0.00]

Figures in squared brackets indicate P values. Diagonal elements are variances while off-diagonal elements represent correlations.

8 Conclusion: Human and social capital

In this paper, we have shown how to derive a theoretical model of scientific production based on the human capital model. We empirically verified the life cycle assumption which appears to be an important factor explaining scientific production. However, this simple individual decision model is not enough to explain the diversity of scientific productivity. We completed this model by situating the individual action in a social context, namely the department. Researchers are regrouped in departments, universities, themselves regrouped in countries. An institution works nicely because it has accumulated social capital that becomes an essential ingredient for individual human capital accumulation. We have tried to measure this influence, using a hierarchical linear model. Institutional social capital is based on cooperation, trust, altruism. Individual researchers pursuit on the contrary personal interest and tend to develop personal networks outside the institution. We have introduced individual variables that could act as proxies for these individualistic behaviours. Some of them are successful (international collaboration), some others are not (publishing alone). Our model has still to be improved for adding new collective variables in order to have a better grasp of social capital.

References

Aghion, P., M. Dewatripont, C. Hoxby, A. Mas-Colell, and A. Sapir (2007): "Why Reform Europe's Universities?," Bruegel Policy Brief 2007/04, Bruegel, Brussels.

Bauwens, L., G. Mion, and J. Thisse (2007): "The resistible decline of European science," CORE-DP 2007-92, CORE, Université catholique de Louvain.

Bourdieu, P. (1980): "Le capital social: Notes provisoires," *Actes de la Recherche en Sciences Sociales*, 3(2-3).

——— (1986): "The forms of social capital," in *Handbook of theory and research for the sociology of education*, ed. by J. G. Richardson. Greenwood, New York.

Bozeman, B., J. S. Dietz, and M. Gaughan (2001): "Scientific and technical human capital: an alternative model for research evaluation," *International Journal of Technology Management*, 22(7-8), 716–740.

Coleman, J. (1988): "Social capital in the creation of human capital," *American Journal of Sociology*, 94, S95–S120.

Goldstein, H., and D. Spiegelhalter (1996): "League Tables and their Limitations: Statistical Issues in Comparisons of Institutional Performance," *Journal of the Royal Statistical Society, Series A*, 159(3), 385–443.

Levin, S. G., and P. E. Stephan (1991): "Research Productivity Over the Life Cycle: Evidence for Academic Scientists," *The American Economic Review*, 81(1), 114–132.

Lindley, D. V., and A. F. Smith (1972): "Bayes estimates for the linear model," *Journal of the Royal Statistical Society, Series B*, 34(1), 1–41.

Lotka, A. (1926): "The frequency distribution of scientific productivity," *Journal of the Washington Academy of Sciences*, 16, 317–323.

Lubrano, M., L. Bauwens, A. Kirman, and C. Protopopescu (2003): "Ranking European economics departments: a statistical approach," *Journal of the European Economic Association*, 1(6), 1367–1401.

McDowell, J. M. (1982): "Obsolescence of Knowledge and Career Publication Profiles: Some Evidence of Differences among Fields in Costs of Interrupted Careers," *The American Economic Review*, 72(4), 752–768.

Merton, R. K. (1968): "The Matthew Effect in Science," *Science*, 159(3810), 56–63.

Putnam, R. D. (1995): "Bowling alone: America's declining social capital," *Journal of Democracy*, 6(1), 64–78.

Raudenbush, S. W., and A. S. Bryk (2002): *Hierarchical Linear Models: Applications and Data Analysis Methods*. Sage Publications Ltd, London, 2nd edn.

Siisiäinen, M. (2000): "Two Concepts of Social Capital: Bourdieu vs. Putnam," Discussion paper, University of Jyväskylä, Department of Social Sciences and Philosophy, Paper presented at ISTR Fourth International Conference: The Third Sector: For What and for Whom? Trinity College, Dublin, Ireland, July 5-8, 2000.

Stephan, P. (1996): "The Economics of Science," *Journal of Economic Literature*, 34(3), 1199–1235.

van Damme, E. (1996): "Measuring Quality of academic journals and scientific productivity of researchers," Unpublished mimeo, CenTer, Tilburg University.