Individual Productivity Differences in Scientific Research :

An Econometric Study of the publications of French Physicists*

Laure Turner[a] and Jacques Mairesse[b]

July 2003

* Preliminary - Comments Welcome

Abstract - In the economics of science framework, many empirical studies have exhibited the persistence of the productivity hierarchies over the life cycle for a cohort of academic scientists. In this paper, we identify rewards in science which may contribute to this persistence and assess their impact on researcher's productivity relative to the influence of individual observed variables (age, gender and education). We have built an unique panel database that concerns French physicists over 1980-1997 and estimate an econometric model that controls for individual unobserved variables and allows time-stable variables. It is an issue for science policy, since the two effects – that of the rewards variables and that of the individual observed variables- are controlled with different tools The first effect depends on the incentive scheme at work in the scientific institution, and the second effect depends on the recruitment policy in academic research. We consider three measures of productivity: the number of articles and the quality of the publications assessed both by the number of citations and by the journals impact factor.

Key words: economics of science, researchers' productivity, count model on panel data, science policy
JEL Classification: C33, C51, O32

We express grateful thanks to Bronwyn Hall, Richard Blundell, Paul David, Emmanuel Duguet and Frank Windmeijer for their profitable comments and suggestions. We thank Emmanuel Duguet, Bronwyn Hall and Frank Windmeijer for allowing us to use their programs, respectively for the Zero-Inflated Poisson, the Poisson and log-linear estimation on panel data and the non-linear GMM estimation of exponential models. This paper also benefited from discussions at the 10th International Conference on Panel Data in Berlin 2002 and at the SPRU Conference "Rethinking Science Policy". Any remaining error is ours.

[a] EUREQua, Université Paris I Panthéon-Sorbonne, Maison des Sciences Economiques, 106-112 Bd de l'Hôpital, 75013 Paris, et CREST-LEI, 27 rue des Saint-Pères, 75007 Paris, e-mail: turner@univ-paris1.fr
[b] CREST-INSEE, 15 Bd Gabriel Péri, 92242 Malakoff, e-mail: mairesse@ensae.fr.

Introduction

In the economics of science framework, many studies have exhibited an empirical regularity: the persistence of the productivity hierarchies over the life cycle for a cohort of academic scientists. As in a variety of other domains, the probability of experiencing an event in the future appears to depend on the past experience - past publication matters for current and future publication.

Two explanations for this type of observed regularity have been proposed (Heckman, 1981). First, the experience of an event can modify the environment and the decision parameters of an individual in such a way that it has a lasting effect on his behaviour in the future. In our case, it could be that a scientist highly productive in the first years of his career gains an early access to the whole set of resources that facilitate research - a stimulating laboratory, funds, looser administrative constraints, etc. – so that he becomes even more productive. The second explanation is that individuals differ across time along unmeasurable criteria which determine their actions. Here, for instance, unobservable variables such as ability, intuition, motivation, positively influence the number of articles published at each date. When we fail to control for these unobservable variables in a regression in which past experience is an explanatory variable for current experience, the relation between past and present experience is overestimated, because past experience is a proxy of the individual unobserved variables.

Our research project is to isolate the explanation best suited to the issue of persisting inequalities in scientific productivity. In this paper, we address one aspect of the problem. We identify rewards in science that render future publication more likely, assuming that those rewards are obtained through past publication, and we assess to what extent they can capture the persistence between past and current publication. To do so, we compare the reward variables' impact on productivity with the influence of individual observed variables such as age, gender and education, in an econometric model that control for individual unobserved variables. We have built an unique longitudinal database for this study that concern French physicists and covers the recent period of 1980-1997.

It is an important issue for science policy, since the two effects on productivity – that of rewards and that of individual observed heterogeneity - are controlled with different tools. The first effect depends on the incentive scheme at work in the scientific institution, and the second effect depends on the recruitment policy in research.

According to American empirical studies, we expect to find a strong relation between the past and the current publication that can be explained by the incentives in research. Science offers non-monetary rewards to the researchers for their productivity, such as reputation, prizes, nominations in prestigious institutions, attributed under a “priority rule”. Consequently, the researchers aim at being the first to discover and to publish. This mechanism is efficient in two ways: it fosters the production of knowledge and supports the disclosure of the discoveries, since the utilisation of a result by the scientific community increases the reputation of the author.

Moreover, a process of “cumulative advantage” could be at work, which amplifies the impact of the non-monetary rewards on publication. In this process, the scientists who managed to get recognition from their peers at one point are more and more rewarded along their career. The interpretation is that successful researchers have access to grants, time, stimulating laboratories and teams, etc., which help them to increase or at least maintain their publishing activity and their reputation. On the contrary, a scientist who has experimented a bad start in his research activity might be obliged at one point to quit research because of the accumulated obstacles. In this context, the collective dimension of research is essential in two ways. A scientist can improve his reputation by working with well-known researchers, and multiple collaborations increase his chances of success in being the first to make a discovery.

The monetary rewards may also play a role in the link between the past and the current publication. Paul David (1994b) built a simulation model that characterizes the American academic research system as one which produces strong persistent inequalities among the researchers in terms of productivity and creates a path dependency from the career initial conditions to the research performance. These two characteristics are explained by a feature of the incentive scheme that grants the individuals’ research projects in a competitive selection process.

In our French case, our data allow us to identify two main categories of reward variables: the laboratory and promotion variables. We believe that membership of a dynamic laboratory that is central in research collaboration stimulates researchers' individual productivity and may be part of the process of cumulative advantages in which well-known scientists enhance their productivity and recognition by working in this type of laboratory. In the same way, promotion is expected to stimulate publication.

We should also find a strong influence of individual observable variables on research productivity. The relation between age and publication life cycle models has been often analysed. In the framework of the life cycle models, Diamond (1984) and Levin and Stephan (1991) conclude that there is a quadratic relation according to which the productivity is decreasing with age toward the end of the career. One issue at stake is to assess whether the aging of the scientific community, as it is occurring in the United States, will reduce the national scientific output. Many other studies have also focused on gender discrimination in science, which has always been an important topic to the scientific community (Stephan, 1998).

The novelty of this paper is to assess the relative impact of the different productivity determinants controlling for unobservable individual specific effects, using a new panel database. The paper is organised as follow. The next section describes the data, the model specification and the estimation methodology (I). Section II describes the results and section III concludes.

I. Specification and Estimation Methodology

Very few data bases exists that allow to study the publication behaviour in science. Levin and Stephan (1991) were the first to build a panel data base on publication counts per individual and per year to revisit some results on scientific productivity that have been obtained on cross-sectional data. To our knowledge, our database is the second of this type and is unique for Europe. It is original in several other ways. First, it size is important, since we are documenting almost 20 years of publication for approximately 500 scientists. In particular, we are able to study the effect of age and experience on publication. Secondly, the period of observation is recent (1980-1997). Levin and Stephan (1991) data concerned the period 1973-1979. It is important to use recent data, since the organisational schemes and incentives rules have changed over the past decades. In France for instance, the scientific policy shifted in the beginning of the 80’s when it has been decided that public research should become multi disciplinary and transfer the scientific knowledge produced to the industry. The rules governing the status of the researchers working in the public laboratories have been modified accordingly. Thirdly, the data concern not only publication counts but also two measures aimed at assessing the quality of the publications - citation counts and impact factors of the journals of publication. Finally, we have added several variables that describe the career of the scientists, and we know their laboratory address in 1997.

The next paragraph describes the configuration of the data, and paragraph B presents the methodology and the different variables that are used in the model.

A. *Configuration of the data base*

The source of our publications, citations and impact factors data is the *Science Citation Index (SCI)* which is produced by the *Institute for Scientific Information (ISI).* The *SCI* covers all the scientific domains and the articles of approximately 3200 most cited journals. The quality of the data is excellent, which make the *SCI* the international reference for bibliometric work.

We draw from the *SCI* the publications, citations and impact factors data for 497 French physicists over the period 1980-1997[1]. The publications are obtained per year and per researcher and the citations are obtained per articles. We then aggregate the citations to get citations counts per year and per researcher as for the publications. The impact factor of a journal is calculated as the mean number of citations obtained by the articles of the journal within two years after publication. It gives information on the journal' s reputation and visibility. Each article is thus characterized by the impact factor of the journal where it was published. Again, we aggregated the impact factors per year and per researcher.

The scientists were working at the CNRS[2] in 1997, in the public funded research sector. They belong to the condensed matter section. The field of condensed matter was chosen for two reasons. First, its characteristics are suited to our study: its research is classified as pure basic science; journals with a sound reputation are clearly identifiable; the size of the field covered is clearly defined; and there is very little mobility among researchers from public research to teaching or to private research. Second, condensed matter is a fast-growing field, honoured by the Nobel Prize for Physics awarded to Pierre-Gilles de Gennes in 1991, and which currently accounts for close to half of all French academic physics[3]. The group of physicists studied here represents a major part of all CNRS researchers in this discipline (654 in 1996). The CNRS and the University (and, to a lesser extent, INRETS[*]) are the only public research institutions in this domain in France., and in 1996 there were 1650 condensed matter physicists working at the University (Barré, Crance, Sigogneau, 1999).

---

[1] For the citations only, the sample is not made of the 497 researchers but of a sub-sample of 352 researchers because of practical reasons that occurred when we collected the data.

[2] The Centre National de la Recherche Scientifique (CNRS) is a public organization for basic research, affiliated to the Ministry responsible for Research. With 25,000 employees (11,000 researchers and 14,000 engineers, technicians and administrative staff), a budget of 16 billion francs in 2001, and laboratories throughout the country, the CNRS covers all fields of knowledge. It relies on over 1,200 research and service units which employ the same number of University researchers as the CNRS' s researchers.

[3] Condensed matter includes all states of matter, on various scales (atom, molecules, colloids, particles or cells), between liquids and solids, in which molecules are relatively close. Its study is based on a heritage of traditions, both experimental (crystallography, diffusion of neutrons and electrons, magnetic resonance imagery, microscopy, etc.) and theoretical (static physics). It is also prompted to develop more and more relations with industry around materials used in electronics, granulars, plastics, food or cosmetic gels, etc.

The researchers entered the CNRS at different dates, between 1960 and 1997. Few researchers enter after 1990 (17), so they have been eliminated from the data. The remaining 480 scientists eventually worked in another research institution before their entry, so that we recorded the papers they published there. Yet, it is desirable to have a stable number of scientists over the period under study for the econometric estimations. Therefore, we have restricted our sample in three directions.

First, we render as homogeneous as possible the observation of the researchers. Among the remaining 480 scientists, 62% entered the CNRS before 1980, so we expect them to publish during the whole period of observation 1980-1997. 7.9% of them entered in or after 1980 and started to publish the year of their entry, so we keep them between their entry and 1997. An other 7.7% entered in or after 1980 but also started to publish after their entry, so we keep them between their entry and 1997, and assign a null publication each year from the year of entry to the year before the first year of publication. Finally, 22.5% of the 480 scientists entered after 1980 but started to publish before their entry, and we keep them between the year of their first recorded paper and 1997. Between the year of their first article and their entry at the CNRS, they are assigned the status that they reach at their entry.

But our sample remains unbalanced, the minimum period of activity lasting 9 years, and the maximum period 18 years. In order to get a balanced panel, we both restrict the period under study to the sub-period 1986-1997 and consider only the researchers that are active during this whole sub-period. Finally our final sample is balanced and concerns 465 scientists over 1986-1997[4].

B. Explaining Scientific Productivity

We study the determinants of researchers' productivity measured along three dimensions - in terms of the annual number of publications per scientist, in terms of the average impact factor of the journals of publication per scientist and per year, and in terms of the average number of citations per scientist and per year - which gives rise to three sets of regressions. As mentioned, we have been able to collect citations for 352 researchers only, so that the last regression concerns a sub-sample of scientists. Moreover, the impact factor and citation measures are strongly correlated. Therefore, our comments will concentrate on the two first regressions, and the citation equation is presented as an extension of the impact factor regression.

Table I.1 indicates the main statistics for these variables as well as for the explanatory variables used in the models. The physicists in our sample published approximately 15000 articles

[*] INRETS = Institut National de Recherche sur les Transports et leur Sécurité

over the period 1986-1997, which correspond to a mean number of 2.7 papers per researcher and per year, with a standard error of 3. The annual publication varies greatly among the scientists, between 0 and 62, the maximum over the period. The mean proportion of researchers with no publication in a year is 27%. The mean number of authors per article is 3.2, and the mean number of pages is 5.5. The scientists get published in journals whose articles receive 2.7 citations on average over two years. The quality of the journals of publications are different across the researchers, ranging between almost 0 and 21.5 citations in two years.

We study the relative impact of two series of determinants on scientific productivity: individual factors - age, gender, education, cohort -, and factors related to the incentive scheme of academic research - the career trajectory or experience, the size and activity of the laboratories in which the scientists work at the end of the period and a mobility indicator. We also take into account the unobserved individual specific effects.

A major problem arise when the aim is to measure the relative impact of variables such as age and experience, because those variables are strongly correlated. Moreover, in our specification, we need to take into account an observed fact, that is the exogenous increase of publication with time, and add time dummies in our model. The issue is considered in paragraph 3.

1. Individual observed and non observed variables

Age

As said previously, the age dispersion among the scientists under study is important. On average, they are 44.6 years old but the standard error amounts to 8.0. The economic studies concerning the age/publication relation use the framework of the life cycle models (Diamond, 1984, Levin and Stephan, 1991). These models enlighten the consequences of the end of the career on the individual productivity and on the allocation of research efforts over time. In Levin and Stephan (1991), the scientists allocate time in order to maximize their utility function over their career. At each date the function depends on the future financial rewards associated with the teaching and consulting activities and on the current research output seen as a proxy for the "puzzle -solving reward". One implication of the model is that research activity declines over the life cycle. Interestingly for our study, this proposition is tested on publication panel data drawn from the *SCI* and the *Survey of Doctorate Recipients*, that concern six sub-fields of physics and earth science including solid state and condensed matter physics over the period 1973-1979. To our knowledge, it was the first attempt to use

[4] For the citations only, the sample is reduced to 352 researchers.

longitudinal data to get evidence on the life cycle effect[5]. We are able to look at this effect as well and to compare our estimates with the ones of Levin and Stephan (1991). The age variable used in the model is described in paragraph 3.

Gender

Men represent 82% of our sample. A main concern in the issue of the gender influence on publication is whether the rewards in science are gender biased. Several studies have concluded that women scientists publish less than men and that they earn less as well (see Stephan 1998). Zuckerman, Cole and Bruer (1991) show that the process of cumulative advantages might be a cause of the persistent position of women in the "outer circle of science" because it amplifies an initial situation where the women published less than the men. But in the empirical studies, the relation between gender and outcome is often biased in the sense that the estimations rely on cross-sectional data that can not allow to account for unmeasurable individual effects reflecting personal motivation, talent or any omitted individual variable explaining productivity. Moreover, the samples used are non random but consists of the successful scientists, which introduce a selection bias in the results. Consequently, it is interesting to test the gender/publication relation on panel data, because we can control for the unobservable specific effects and take into account the evolution that led to the observed situation. In this framework, Stephan (1998) finds that gender is not a significant determinant of salary changes in US academe during the 1970's.

Education

Education is introduced as a dummy variable equal to one when the researchers studied in a "Grande Ecole" in addition to graduate from their PhD[6]. 16% of our sample did so. Among them, over 60% belonged to the Ecole Normale Supérieure, 6% to the Ecole Polytechnique, 10% to the Institut Supérieur d'Electronique du Nord, 6% to the Ecole Supérieure d'Electricité de Paris, etc.

Individual specific effects

Apart from the explanatory variables, we also introduce in the model random effects specific to the individuals, in order to take into account the unobserved individual heterogeneity. By doing so, we avoid the bias due to omission of them if the individual effects truly exist. And we can make two types of inferences: either inferences that are conditional on the individual effects, or inferences that are made on the ground of the whole set of effects - among which the specific effects of the model have

[5] This effect was not properly identified on cross-sectional studies, since they did not control for cohort effects.

[6] In the French educational system, after they graduate from high school, the students can either go to the University, which does not require any level nor grade achievement in high school, or they can apply to a preparatory class where they will be taught during two years the knowledge required to compete for the admission into a "Grande Ecole". Every student of the Gr andes Ecoles therefore succeeded in two selection processes: selection on the basis of their grades in high school, and exams to enter the Ecoles.

been randomly drawn. In the first approach, only the studied sample is of interest. We can not predict the publication behaviour of a scientist from outside the sample using our estimates. Practically, this approach amounts to treating the individual effects as individual dummies, that is to say as fixed effects. On the contrary, the second approach allows to draw conclusions for the whole population from which the sample comes. The individual effects are considered as random. In what follows, we adopt this second approach. The choice of our estimation method is determined by the existence of a correlation between the individual effects and the explanatory variables.

2. Incentives variables

Career

The researchers are distributed according to the evolution of their career, so that we can study the link between publication and promotion and more generally account for an incentive mechanism of the scientific institution. A researcher with a typical career profile enters the CNRS as "Chargé de Recherche" (CR), is then promoted research director of class 2, "Di recteur de Recherche de 2ème classe" (DR2), and finally research director of class 1, "Directeur de Recherche de 1 ère classe" (DR1). Yet, many researchers in our sample are never promoted and remain in the same status during the observed period (respectively for the status CR : 46.7% of the sample, DR2 : 10.4%, DR1 : 3%). Almost 30% of the sample get to be promoted DR2. The most difficult promotion to obtain is DR1: only 10% of the sample succeed to be promoted DR1.

The interactions between the career inside the scientific institution and publication behaviour are numerous. A descriptive study (Turner, 1999) on our sample established the existence of a positive link between the status and publication at each date, whereas the relation between the current position and the past publication is more complex. We would expect it to be positive as well, since the CNRS incentive scheme is such that promotion rewards publication.

On the contrary, does promotion give incentives to publish more, for instance by offering a better access to the resources needed for research? The descriptive study shows that on average the promoted researchers remain as productive as before their promotion. But the mean number of publications differs before and after the promotion according to the status reached, the age, and the period. For instance, the oldest scientists publish less after any promotion, especially after a DR1 promotion obtained during the last sub-period whereas younger researchers publish more after a DR2 promotion.

In our econometric model, we introduce a certain number of variables related to the career trajectory or promotion profile as explanatory variables. In a second version of the paper, we will take into account the fact that current position is influenced by past values of publication by considering those variables as endogenous.

The variables are chosen to capture observed changes in the publication behaviour as the number of years since promotion increases. We observe that the productivity of the scientists who are never promoted and who remain CR reach first increases with time and then decreases after a certain number of years spent in the status, as a sign of discouragement. We also observe that the DR2 productivity seems to remain on a constant trend, whatever the promotion perspectives or the total experience at the CNRS. Finally, it appears that the DR1 productivity is decreasing with tenure in the status. Consequently, we retain as variables the tenure in each status and dummies for the status DR2 and DR1[7]. The tenure variables are described more precisely in paragraph 3.

Laboratory effect

A study of the collaboration among our scientists presented in Mairesse and Turner (2002) has underlined the impact of some laboratory characteristics – the size, the productivity of its members, the quality of its publications, its international collaborations, etc – on the intensity of co-publication at the laboratory level. We want to assess the impact of the same laboratory variables on the individuals productivity. This "laboratory effect" is an issue when evaluating the recent policies that stimulate the creation of large research structures like technopoles. We have in mind that the membership of a dynamic laboratory stimulates researchers' individual productivity and may be part of a process of cumulative advantages in which well-known scientists enhance their productivity and recognition by working in this type of laboratory.

The variables are calculated for the laboratories where the researchers were working in 1997. They are time invariant and some are aggregates over the whole period of 12 years. More precisely, the laboratory variables are the following:

- ♣ Size: the size of the laboratory is its the total number of researchers in the laboratory in 1997 whether they are affiliated to the CNRS, to a University or to another research institution. The size variable is also centred and squared in order to measure quadratic effects.
- ♣ Productivity of the laboratory : it is calculated for every researcher individually over the whole period 1986-1997, by subtracting its personal contribution to the production of his colleagues who we have in our database and who are working in the same laboratory in 1997. Subsequently, it proxies the productivity of the researcher's environment.

[7] Several trials have later confirmed that any variable controlling for the total number of years the scientist has been working at the CNRS or the perspective of the career ending had no significant effect on productivity. The major effects are the effects of tenure per status.

- ♣ Quality at the laboratory level: it is also calculated for every researcher individually over 1986-1997 by subtracting its personal contribution to the average impact factor of his colleagues who we have in our database and who are working in the same laboratory in 1997. It reflects the quality of the researcher's environment.
- ♣ International openness of the laboratory: it is the proportion of articles at the laboratory level and over the whole period co-published with at least one foreign co-author (see Mairesse and Turner, 2002).
- ♣ Region dummies: a dummy for the Grenoble region and the Paris region are introduced to qualify the laboratories in 1997 because those two regions account for a major part of the total number of physicists and publications over the period. The regions are defined as Grenoble (resp. Paris) plus the set of towns geographically close to Grenoble (resp. Paris) - less than 100 km - in which CNRS laboratories are located in 1997.

Of course, the mobility of the researchers is of central interest for the evaluation of a "laboratory effect", and taking the 1997 laboratory of the researcher seems to give an incomplete information. This is not so because of two main reasons. First, the actual mobility is very low: 55% of the researchers in our initial sample never changed laboratories, 33% changed once, 11% changed twice, and 3% changed three times. We add a dummy when the number of laboratories frequented by the researchers is at most one, and a dummy when this number equals more than one to control for the frequency of changes. They are to capture an anticipated effect of mobility according to which the researchers who moved more than once are affiliated to more productive laboratories on average.

Second, the meaning of the laboratory changes is often weak, in the sense that it is not associated with either promotion or variability in the productivity and quality of the researchers environment. As a matter of fact, only nine persons changed laboratory as they were promoted, and only 20 persons who changed were promoted within two years. Finally, the "within" productivity and quality variability of the laboratories occupied by an individual who was mobile over his career represent on average less than 15% of the total variability.

The fact that the laboratory variables are aggregates over 12 years instead of being time-variant is also not a main issue because of three reasons. First, a long period is needed to assess the genuine or "steady-state" productivity and quality of the researchers who contributed during a certain period to the reputation of the laboratory[8]. Second, the mobility of the researchers is weak as mentioned. And third, by explicitly considering this mobility, we would get laboratory variables that

[8] We have done the same calculations over a six years period but it does not change the results.

would be over-dispersed across time, because for a significant number of laboratories, the number of researchers would fall or be unknown for some years.

Nevertheless, in a second version of the paper, we will consider our laboratories variables as endogenous, assuming that they are influenced by past publication.

3. Time, Age and Tenure effects

First statistics on our data suggest that publication increases toward the end of the period independently of the age of the researchers. We introduce years dummies in the regressions in order to capture this effect and account for any changes in the work environment or in the state of the art in condensed matter physics. But the estimation method that we use (section C) require to write the variables in deviation from their mean, which rise a major identification problem since the transformed age and time variables are collinear. Therefore, we decided to assess the age effect by four age cohorts instead of continuous age and age squared variables. It breaks the link between age and time so that it is possible to estimate both effects simultaneously[9].

A last question remains. Including both the tenure variables and the time periods in the model rises another identification problem. Because a significant number of researchers are never promoted, time and tenure in status also show some collinearity, especially when considering the variables in deviation from the means. The solution proposed is to group the tenures in status into three cohorts equivalent in size. Again, it has the advantage to break the link between tenure and time and to allow the estimation of time-varying tenure effects.

C. Methodology

To estimate the first equation in which the dependant variable is the number of articles, a count variable, we need to estimate a Poisson model of productivity. In a first version of the model, we do the estimations under the main hypothesis that the explanatory variables are strictly exogenous.

The model is the following, with i={1,…, N}, N=465, and t={1,…,T}, T=12:

[9] To compare the effect of age on productivity with the one found by Levin and Stephan, we also run an annex regression in which the continuous variables age and age squared replaced the age cohorts. The time dummies were not correctly estimated, but we were simply interested in the age estimates. We verified that the age estimates were robust to the exclusion of the time dummies from the model.

$$E(\ y_{it} \big| X_{it}, Z_i\ ) = exp(\ \mu + Z_i \gamma + X_{it} \beta + \alpha_i\ ) \quad (1)$$
$$with \quad \mathrm{y}_{\mathrm{it}} \rightarrow Poisson$$

The variables in Z are stable across time but not across individuals and the variables in X vary in both dimensions. The random individual effects are α_i. We assume that the errors are not serially correlated. But according to the Hausman tests, we assume that the individual effects are correlated with the explanatory variables. In order to be able to estimate the coefficients of the time-invariant variables of the model, we assume that all the correlation with the individual effects is due to the time-varying variables in X, and that the time-invariant variables in Z are not correlated with the individual effects[10]. We have:

$EX_{it}'\alpha_i \neq 0$, $EZ_i'\alpha_i = 0$,
$EX_{it}'u_{it} = EZ_i'u_{it} = 0$,
$E\alpha_i = Eu_{it} = 0$,
$E\alpha_i\, u_{it} = 0$, (H1)
$E\alpha_i\, \alpha_j = \sigma_\alpha$ if i = j, and $E\alpha_i\, \alpha_j = 0$ otherwise,
$Eu_{it}u_{is} = \sigma_u$ if i = j and t=s, and $Eu_{it}u_{is} = 0$ otherwise.

We estimate β by the Conditional Maximum Likelihood Estimation[11] (CMLE) proposed by Hausman, Hall and Griliches (1984). In a second step, to estimate the coefficients of Z, we replace β by its CMLE estimate and estimate equation (2) using the non linear least squares method :

$$y_{it} / \exp(X_{ti}\hat{\beta}) = \exp(\mu + \gamma\, Z_i) + \varepsilon_{it} \quad (2)$$

We obtain consistent estimates of γ and μ. We run the Two Step estimation (TS) as well as the level estimation - the basic Poisson model (TOTAL) – in order to assess the size of the unobservables effect. Only the Two Step regressions take the individual effects into account. The results are in *table II.2.* In *table II.3*, we have calculated the marginal impact of the variables at the mean vector of the sample. This is the table to which we refer ourselves in the discussion of the results.

When the dependent variable is the average quality of the papers per researchers and per year – measured by the impact factor and as an extension by the number of citations, the Poisson model is replaced by the log linear model, since the dependant variables are continuous. The same Two Step

[10] By doing so, we do not estimate the raw effect of the time invariant variables. For instance, we do not separate the effect of being a woman from the correlated unobserved effect of the number of children. The woman variable embodies the fact of being a woman plus all the correlated facts. We believe that it does not reduce the interesting descriptive properties of the variables.

method is used to estimate the time-invariant variables. The results are in *table II.1, II.2* and for the citations in *table II.3*.

III. Results

A. *The determinants of publication*

Our comments concentrate on the TS regression. The TOTAL estimation reveals the same effects than the TS, the estimates of the time varying variables being higher meaning that the correlations between the individual effects and the variables are positive.

When looking at the age cohorts, the estimation suggests a quadratic relation between the age of the scientists and the number of their publications. According to the estimation, the researchers productivity increases between the first and the third age cohort, that is before 50, and than a decline occurs after 51 years. More precisely, the researchers aged 39 to 45 years publish on average 0.26 paper more per year than the youngest researchers aged 26 to 38, and the scientists aged 46 to 50 publish 0.36 paper more than the youngest cohort. The older researchers, aged 51 to 61, publish only 0.13 paper more than the youngest researchers.

To refine our idea on the effect of age on productivity, we have run an annex regression in which the continuous variables age and age squared replaced the age cohorts. As explained previously, the time dummies can no more be correctly estimated, but we are simply interested in the age estimates. We check that the age estimates do not vary to much when we exclude the time dummies from the equation. The TS estimates of age and age squared are respectively 0.022 and –0.0016 when the time dummies are included in the regression, and 0.025 and –0.0014 when the time dummies are excluded. Consequently, the quadratic relation between age and productivity is confirmed. According to the annex TS regression in which time dummies are included, the conditional effect of age is such that the researchers are more productive every year until they turn 52 but at a diminishing rate: they publish 0.9 paper per year on average at 30 years, 2.3 papers at 40, 2.9 papers at 52 and 2.5 papers at 60. The curves illustrating the age/publication relation according to the TS estimation and conditionally on the other variables is represented on *graph II.1*. It includes the mean point of 2.7 papers at 44.6 years old.

---

[11] Rappeler les raisons qui font que l'on peut estimer les effets aléatoires comme des effets fixes.

We compare our TS results to those of Levin and Stephan (1991) Tobit model with correlated fixed effects and time dummies (model B in their paper). The quadratic relation between age and publication is confirmed. Yet, the life cycle effect is much stronger in their model, in the sense that their results imply a relation more quadratic then ours. According to their model, the solid state and condensed matter physicists are productive between 33 and 57 years old, publishing 0.71 paper per year at 35 years old, a peak of 2 papers at 45, and 0.72 paper at 55. The mean number of papers is 3.8 over two years, which is slightly smaller than the average productivity in our sample but is still comparable (*graph II.2*). The differences in the findings can in part be explained by the fact that the specification of our model takes into account the count nature of the data and more importantly the great proportion of zeros.

The age effect is complemented by a time effect according to which the productivity increases with time for all the researchers. The scientists publish 0.9 paper more on average in 1991 than in 1986, and 1.6 papers more in 1996 than in 1986. This tends to suggest that the scientists progressively experiment a wider and faster access to publication related to the increasing numbers of existing journals.

We find a strong influence of the other individual variables – especially gender - on productivity. A woman publishes almost 0.9 paper less than a man on average per year. And the scientists who have been educated in a Grande Ecole publish 0.7 paper more than the others per year on average.

The TOTAL and TS estimations give a different picture of the career influence on the publication behaviour. In the TOTAL estimation, a DR1 scientist publishes on average 1.5 papers more per year than a CR scientist, and a DR2 scientist 0.5 paper more. Whereas in the TS estimation, the DR2 coefficient is not significant, and the DR1 coefficient is negative, the DR1 publishing on average 0.77 paper less than a CR. Reaching the status DR1 has an negative impact on productivity according to the TS regression. But this result has to be taken cautiously because the DR1 estimates are not very robust across different specifications of the time effect. As a matter of fact, more than 50% of the researchers reach the status DR1 after 1993, so that the variable DR1 is strongly correlated with the last time period 1993-1997. Running the regression without those years rises the DR1 estimate from –0.29 to –0.14, and the coefficient becomes non significant.

On the contrary, the tenure estimates are robust to the presence of time dummies. In the TS regression, the estimates of the third tenure cohort is negative when interacted with the DR1 dummy. A DR1 who belong to this third cohort tenure publish 1.3 papers less on average than a newly promoted DR1 researcher. We find this effect also in the TOTAL estimation. It means that the impact on productivity of reaching the status DR1 diminishes as the number of years spent in the status increases. Consequently, the incentives to publish appears to be lower with time for those researchers who have reached the higher status. The increase in their amount of administrative tasks must also be recalled.

Reaching the status DR2 has an positive impact on productivity according to the TOTAL regression but has no impact according to the TS regression. In the TS regression, the estimates of the tenure cohorts is negative when interacted with the DR2 dummy. A DR2 who belong to the second cohort publish 0.3 paper less on average than a newly promoted DR2, and a DR2 of the third tenure cohort publish 0.5 paper less on average than a newly promoted DR2. Again, it appears that the incentives to publish become lower as the time spent in the status increases. This effect is not significant in the TOTAL estimation. Finally, according to the TS estimation, tenure has no effect on the productivity of the CR researchers.

There is a important positive impact of the laboratory's productivity on the number of articles published by the individuals. We find that, in the case of a 10% rise in the productivity at the laboratory level, a scientist would on average publish 0.27 paper more than the number he would publish otherwise. But the quality of the laboratory has a negative impact on the publication of its members according to the TOTAL estimation. In case of a 10% rise in the quality at the laboratory level, a scientist would on average publish 0.25 paper less than the number he would publish otherwise A direct interpretation of this effect could be that working in a high quality laboratory requires efforts from the researchers to develop the quality of their papers, which is obtained at the detrimental of individual productivity. Statistically, a small substitution effect between individual quality and individual productivity is suggested by the negative correlation coefficient between the two variables (-0.036) conditionally on the other variables and a dummy for null publication. Yet, this effect does not hold in the TS estimation where the coefficient of the quality at the laboratory level is not significant.

Interestingly, the individual productivity is decreasing in the size of the laboratory. For a given laboratory, a 10% increase in size entails 0.09 paper less on average per scientist. This result is an element against the idea that grouping the researchers into large poles of research would foster their productivity. But collaboration with foreign laboratories has a strong positive impact on individual productivity. The last two laboratory effects are the following. First, it appears that the researchers working in the Grenoble region are more productive (almost 0.5 publication more per year on average), which can be explained by the experimental nature of the research done in this region where the greatest part of the technologies used in the field are located. Second, mobility does not play a significant role in explaining scientific productivity. In the TOTAL regression, the scientists who change laboratory twice or more publish on average 0.1 paper more per year than the other scientists, but no effect of mobility appears in the TS regression.

To conclude this section, the relative importance of the individual and incentive effects on productivity is as follow. Age plays an important role in the explanation of individual productivity. We find a quadratic relation between age and individual productivity according to which the productivity declines after 52 tears old. Tenure significantly contributes to explain the productivity decline with age after this threshold (graph II.3). It is related to the important negative effect of long tenure in the DR1 and DR2 status. The age and tenure effects are complemented by a time effect especially strong after 1991. Gender is a major individual determinant of productivity as well as, to a lesser extent, high education. Among the laboratory characteristics, only the international openness of the laboratory through collaboration and the productivity of the laboratory influence publication with the same order of magnitude than the previous variables. Yet, it is noticeable that the size of the laboratory has a small negative effect on individual productivity, and that the accessibility of the technologies for experiments - captured by the Grenoble region dummy - has a positive impact on productivity.

B. The determinants of the average quality of the journals of publication

Our comments focus on the TS estimation. The average quality of the journals of publication is negatively influenced by age and no quadratic relation emerges from the estimates of the age cohorts. The oldest researchers aged 51 to 61 publish in journals that receive on average 0.3 citation less than journals of the young cohort researchers. As previously, to refine our idea on the effect of age on impact factor, we have run an annex regression in which the continuous variables age and age squared replaced the age cohorts. According to the annex TS regression in which time dummies are included, their is a negligible increase in the impact factor between 26 and 41, from 2.50 to 2.67. After 41, the average impact factor declines slightly and is equal to 2.35 at 61. The curves illustrating the age/publication relation according to the TS estimation and conditionally on the other variables is represented on *graph II.4*.

The estimates of the time dummies are negative but suggest that if the impact factor decreased until 1988, it increased afterwards until 1997 where it approaches the level of 1986.

The gender and education variables are significant and influence similarly the quality of the papers published, but with a much smaller order of magnitude than the one found in the regression on the number of articles. The women publish in journals that receive on average 0.10 citations less over two years than the journals of men publication.

The dominant effect here is the ones of the quality of the laboratory. A 10% increase in the quality of the laboratory would increase of 0.58 the impact factor within two years. The other laboratory effects are comparatively weak. A 10% increase in the productivity of the laboratory would decrease the impact factor of 0.05, so that a substitution effect reciprocal of the one identified in the previous paragraph is suggested. A productive environment may stimulate individual productivity to the detriment of individual quality. Finally, the size of the laboratory has a small negative impact on the quality of the publications of its members, but at a decreasing rate, as shown by the positive size squared estimate.

The last thing to notice is that the status and promotion variables have no impact on the average impact factor of the journals.

To conclude this section, the main determinants of the average impact factor of the journals are quality of the laboratory publications, the gender and education variables, and to a lesser extent, the age.

C. *Extensions :* The determinants of the average number of citations per two years

The results of the estimation are shown in *table II.3.* We focus on the differences in the results a regard to the estimates of the impact factor model. It appears that age do not have an impact on the average number of citations per two years, according to the within estimates, whereas the impact was negative in the impact factor model and even in the total estimation on the citations. Being a woman has a negative effect on citations, almost four times higher than previously (–0.385). And the education effect is higher as well, almost five times higher than in the impact factor model (0.536). The last noticeable result is that the citations are strongly negatively influenced by the time spent in the status DR1. On average, a DR1 in cohort 2 of experience in the grade receives on average 0.6 citations less in two years per paper than a newly promoted DR1, and the figure amounts to 1.2 for a DR1 in cohort 3 of tenure in the grade.

III. Conclusion

This work explores the differences in productivity among scientists in public research, both in terms of the number of articles and of the quality of the publications. We use a unique longitudinal data base, concerning French condensed-matter physicists between 1980 and 1997. Two sets of factors have been considered as determinants of the researchers' productivity. First, individual variables – age, cohort, gender, education. And second, variables that may be related to the incentive devices at work in the scientific institution – the status and tenure variables, and the laboratory variables. Finally, a dynamic model also includes lags of the dependent variable in the regressors (to be completed) and treat the reward variables as predetermined.

We find a strong impact of the individual variables. In particular, we find a "life cycle effect" such that productivity first increases and then decreases with age, and a gender effect according to which women publish less than men and get less citations. We also find that a high education of the researchers is determinant to foster productivity. Some environmental factors seems to play a role as important. Individual productivity is stimulated in productive laboratories that are not too big and that are participating in international networks of co-publication. Finally, our results suggest that promotion plays as a research incentive. A long tenure in a status, and especially in the higher status, has a negative impact on productivity.

References

ADAMS J., GRILICHES Z., 1996, 'Research productivity in a system of universities", NBER.

ALLISON P.D., PRICE D.J. deSolla, GRIFFITH B.C., MORAVESIK M.J. and STEWART J.A., 1976, "« Lotka's Law : a problem in its interpretation and application », Social Studies of Science, 6.

ALLISON P.D., STEWART J.A., 1974, « Productivity differences among scientists : evidence for accumulative advantage », American Sociological Review, 39, august.

ALLISON P.D., LONG J.S., KRAUZE T.K., 1982, « Cumulative advantage and inequality in Science », American Sociological Review, 47, october.

AUERBACH F., 1910, « Geschischtafeln der Physik », Leipzig : J.A. Barth.

BARRE R., CRANCE M., SIGOGNEAU A., 1999, « La recherche scientifique française : situation démographique », Etudes et dossiers de l'OST, 1, avril.

BREUSCH T., MILZON G. and SCHMIDT P., 1989, « Efficient Estimation Using Panel Data », Econometrica, 57(3), May, 695-700.

COLE J. and COLE S., 1973, *Social stratification in science*, Chicago : U. of Chicago Press.

COLE S., 1979, "Age and scientific performance", American Journal of Sociology, 84(4), january.

DASGUPTA P., DAVID P., 1994, 'Toward a new economics of science", Research Policy, 23(5).

DAVID P., 1994a, "Positive feedbacks and research productivity in science: reopening another black box", in Ove Granstrand, Amsterdam: Elsevier Science B. V (Eds.), *The economics of technology*.

DAVID P., 1994b, " A science Simulation for studying the US and similar institutional setups (SCISIMUS)", miméo.

DIAMOND A., 1984, « An economic model of the life-cycle research productivity of scientists », Scientometrics, 6(3), may.

DUGUET E., MONJON S., 2001, « Creative destruction and the Innovative Core : is Innovation Persistant at the firm level ?" mimeo, Cahier de la MSE, 2001.

HALL B.H., JAFFE A., TRAJTENBERG M., 1999, « Market value and patent citations : a first look », mimeo, june.

HAUSMAN J., HALL B., GRILICHES Z., 1984, « Econometric Models for Count Data with an Application to the Patent R&D Relationship », Econometrica, 52(4), July, 909-937.

HAUSMAN J., TAYLOR, 1981, "Panel Data and Unobservable Individual Effects", Econometrica, 49(6)

HECKMAN, 1981, "Statistical models for discrete panel data", in Structural Analysis of Discrete Data with Economic Applications, eds C.Manski and D.McFadden, the MIT Press, Cambridge Massachusetts.

LEVIN S., STEPHAN P., 1991, "Research productivity over the life cycle: evidence for academic scientists", American Economic Review, 81(1), march.

LONG, J.S., 1992, "Measures of sex differences in scientific productivity", Social Forces, 71, 159-178.

LOTKA A., 1926, "The frequency distribution of scientific productivity", Journal of the Washington Academy of Sciences, 97, september.

MAIRESSE J. and L.TURNER, « Measurement and Explanation of the Intensity of Co-Publication in Scientific Research: An Analysis at the Laboratory Level ». First draft, in revision, 2001. (To appear in « New Frontiers in the Economics of Innovation and New Technology », Essays in honour of Paul David", eds C.Antonelli, D. Foray, B.Hall and E.Steinmueller, Oxford University Press.)

PRICE, D.deSolla, 1976, « A general theory of bibliometrics and other cumulative advantage processes » Journal of the American Society for Information Science, 27 (5/6).

STEPHAN P., 1996, 'The economics of science", Journal of Economic Literature, 31, september.

STEPHAN P., 1998, « Gender Differences in the Rewards to Publishing in Academe: Science in the 1970's », Sex Roles, Vol.38, 11/12.

TURNER L., «Explication des différences de productivité et de promotion pour un échantillon de physiciens du CNRS », Mémoire de DEA, Université de Paris 1, 1999.

ZUCKERMAN H., COLE J.R., and BRUER J.T. (eds.), 1991, *The outer circle: Women in the scientific Community*, New York: W.W. Norton.

Table I.1

	Mean	Standard Error	Median	1st Qrt	3rd Qrt
Dependent variables					
Number of articles per year ART	2,69	3,21	2	0	4
Average impact factor per researcher and per year NOT_I	2,66	2,30	2,54	0	3,8
Extensions					
Average number of authors per article (harmonic) NOTA	3.23	2.57	3.33	0	4.90
Average number of pages per article NOT_P	5,48	4,68	5,4	0	7,83
Individual variables + Time Dummies					
AGE	44.65	8.03	45	38	51
Age cohort 1 (AGE1) 26<=age<=38	0,25	0,44	0	0	1
Age cohort 2 (AGE2) 38<age<=45	0,25	0,43	0	0	0
Age cohort 3 (AGE3) 45<age<51	0,23	0,42	0	0	0
Age cohort 4 (AGE4) 51<=age<=61	0,27	0,44	0	0	1
Education in a 'Grande Ecole" (ECOLE)	0,17	0,38	0	0	0
Gender (WOMAN)	0,18	0,39	0	0	0
More than one mobility (DUMCH23)	0,13	0,34	0	0	0
Promotion variables					
Status (DR2_0)	0,08	0,28	0	0	0
Status (DR1_0)	0,33	0,47	0	0	1
Tenure in status CR cohort 1 (C1ANCCR)	0,21	0,41	0	0	0
Tenure in status CR cohort 2 (C2ANCCR)	0,19	0,39	0	0	0
Tenure in status CR cohort 3 (C3ANCCR)	0,18	0,39	0	0	0
Tenure in status DR2 cohort 1 (C1ANCDR2)	0,13	0,33	0	0	0
Tenure in status DR2 cohort 2 (C2ANCDR2)	0,10	0,3	0	0	0
Tenure in status DR2 cohort 3 (C3ANCDR2)	0,10	0,3	0	0	0
Tenure in status DR1 cohort 1 (C1ANCDR1)	0,03	0,18	0	0	0
Tenure in status DR1 cohort 2 (C2ANCDR1)	0,02	0,15	0	0	0
Tenure in status DR1 cohort 3 (C3ANCDR1)	0,03	0,16	0	0	0
Laboratory variables					
Size of the laboratory in logarithm LOGNBCH	3,23	1,39	3,64	3	4,09
Productivity of the laboratory in logarithm (LOGPROD)	0,72	0,42	0,79	0,44	0,99
Quality of the laboratory publications in logarithm (LOGQUAL)	1,09	0,45	1,27	1,09	1,33
Proportion of the laboratory articles with foreign co-authors (MPETR)	0,04	0,03	0,04	0,02	0,06
Dummy for the Grenoble region (ADUMGREN)	0,26	0,44	0	0	1
Dummy for the Paris region (ADUMPAR)	0,36	0,48	0	0	1
Dummy for laboratory with less than 3 researchers (DUMEF13)	0,14	0,34	0	0	0

♣ *Number of individuals = 465, Number of years = 12, Number of Observation = 5580*

Table II.1

	POISSON ON ART		MARGINAL IMPACTS	
Variables	TOTAL	Two Step	TOTAL	Two Step
AGE2	0.205*** (0.025)	0.098*** (0.029)	0.553	0.263
AGE3	0.233*** (0.025)	0.137*** (0.032)	0.628	0.368
AGE4	0.072*** (0.025)	0.086*** (0.028)	0.193	0.233
WOMAN	-0.273*** (0.024)	-0.33*** (0.046)	-0.736	-0.89
ECOLE	0.118*** (0.021)	0.26*** (0.043)	0.317	0.701
DUMCH23	0.041* (0.025)	-0.024 (0.051)	0.11	-0.064
DR1	0.543*** (0.041)	-0.287*** (0.096)	1.463	-0.772
DR2	0.174*** (0.031)	-0.045 (0.06)	0.47	-0.121
C2ANCCR	-0.094*** (0.029)	0.028 (0.054)	-0.252	0.077
C3ANCCR	-0.291*** (0.032)	0.054 (0.079)	-0.786	0.145
C2ANCD2	0.012 (0.031)	-0.129*** (0.042)	0.031	-0.347
C3ANCD2	-0.029 (0.034)	-0.2*** (0.062)	-0.078	-0.539
C2ANCD1	-0.05 (0.057)	0.023 (0.07)	-0.134	0.063
C3ANCD1	-0.287*** (0.058)	-0.494*** (0.091)	-0.774	-1.33
MPETR	2.942*** (0.432)	3.01*** (1.07)	7.928	8.112
LOGNBCH	-0.131*** (0.015)	-0.091** (0.036)	-0.131	-0.091
LOGNBCH2	0.016*** (0.006)	0.016 (0.011)	0.016	0.016
LOGPROD	0.233*** (0.032)	0.274*** (0.067)	0.233	0.274
LOGQUAL	-0.254*** (0.069)	0.018 (0.144)	-0.254	0.018
ADUMGREN	0.173*** (0.026)	0.179*** (0.053)	0.465	0.483
ADUMPAR	-0.055** (0.026)	0.009 (0.047)	-0.147	0.026
DUMEF13	-0.66*** (0.099)	-0.109 (0.202)	-1.78	-0.293

A87	0.117*** (0.046)	0.114** (0.046)	0.315	0.307
A88	0.203*** (0.045)	0.219*** (0.045)	0.548	0.589
A89	0.336*** (0.043)	0.362*** (0.044)	0.905	0.974
A90	0.191*** (0.045)	0.254*** (0.046)	0.514	0.684
A91	0.227*** (0.044)	0.329*** (0.046)	0.612	0.885
A92	0.262*** (0.044)	0.393*** (0.046)	0.707	1.06
A93	0.457*** (0.042)	0.599*** (0.045)	1.232	1.615
A94	0.388*** (0.043)	0.55*** (0.046)	1.046	1.481
A95	0.297*** (0.044)	0.482*** (0.048)	0.8	1.298
A96	0.393*** (0.043)	0.583*** (0.047)	1.059	1.572
A97	0.373*** (0.043)	0.575*** (0.048)	1.005	1.55
C	0.95*** (0.105)	0.372** (0.19)	2.561	1.003
Log-vraisemblance R^2-Adj for Step 2	-14009.8	-8978.58 0.03		

- ♣ *Number of individuals = 465, Number of years = 12, Number of Observation = 5580*
- ♣ *Standard Errors computed from analytic second derivatives (Newton) for the TOTAL and First Step and from quadratic form of analytic first derivatives (Gauss) for the Second Step.*

Table II.2

	LOGLINEAR ON NOT_I		MARGINAL IMPACTS	
Variables	Total	Within	Total	Within
AGE2	-0.041* (0.021)	-0.002 (0.029)	-0.109	-0.004
AGE3	-0.104*** (0.026)	-0.06 (0.042)	-0.276	-0.16
AGE4	-0.144*** (0.03)	-0.111** (0.055)	-0.383	-0.295
WOMAN	-0.022 (0.016)	-0.04** (0.016)	-0.058	-0.106
ECOLE	0.032* (0.017)	0.042** (0.016)	0.085	0.111
DUMCH23	-0.012 (0.019)	-0.01 (0.019)	-0.031	-0.026
DR1	0.135*** (0.042)	-0.054 (0.074)	0.36	-0.143
DR2	0.058** (0.026)	-0.041 (0.046)	0.155	-0.109
C2ANCCR	-0.004 (0.02)	-0.02 (0.04)	-0.01	-0.054
C3ANCCR	0.015 (0.028)	-0.011 (0.056)	0.041	-0.029
C2ANCD2	0.041 (0.026)	-0.021 (0.031)	0.11	-0.055
C3ANCD2	0.063** (0.029)	0.039 (0.046)	0.166	0.105
C2ANCD1	0.033 (0.052)	0.025 (0.059)	0.089	0.066
C3ANCD1	0.006 (0.05)	0.069 (0.077)	0.016	0.184
MPETR	0.031 (0.326)	0.121 (0.327)	0.082	0.322
LOGNBCH	-0.033*** (0.012)	-0.025** (0.012)	-0.033	-0.025
LOGNBCH2	0.021*** (0.004)	0.024*** (0.004)	0.021	0.024
LOGPROD	-0.046* (0.024)	-0.049** (0.024)	-0.046	-0.049
LOGQUAL	0.57*** (0.051)	0.579*** (0.05)	0.57	0.579
ADUMGREN	0.029 (0.019)	0.044** (0.019)	0.077	0.118
ADUMPAR	0.009 (0.018)	0.025 (0.018)	0.023	0.067
DUMEF13	0.501*** (0.074)	0.551*** (0.073)	1.335	1.467

A87	-0.051* (0.03)	-0.051* (0.028)	-0.135	-0.135
A88	-0.123*** (0.03)	-0.122*** (0.028)	-0.327	-0.325
A89	-0.109*** (0.03)	-0.103*** (0.029)	-0.29	-0.274
A90	-0.1*** (0.03)	-0.085*** (0.03)	-0.265	-0.227
A91	-0.123*** (0.03)	-0.104*** (0.031)	-0.328	-0.277
A92	-0.076** (0.03)	-0.054* (0.032)	-0.203	-0.143
A93	-0.191*** (0.031)	-0.167*** (0.033)	-0.509	-0.444
A94	-0.087*** (0.031)	-0.065* (0.034)	-0.232	-0.172
A95	-0.057* (0.031)	-0.03 (0.035)	-0.151	-0.08
A96	-0.036 (0.032)	-0.007 (0.037)	-0.095	-0.019
A97	-0.062** (0.032)	-0.035 (0.038)	-0.166	-0.093
DUMMY (ART=0)	-1.119*** (0.014)	-1.093*** (0.016)	-2.979	-2.91
C	0.566*** (0.077)	0.516*** (0.071)	1.506	1.373
Log likelihood R^2-Adj	-3482.77 0.557	-2941.03 0.602		

♣ *Number of individuals = 465, Number of years = 12, Number of Observation = 5580*

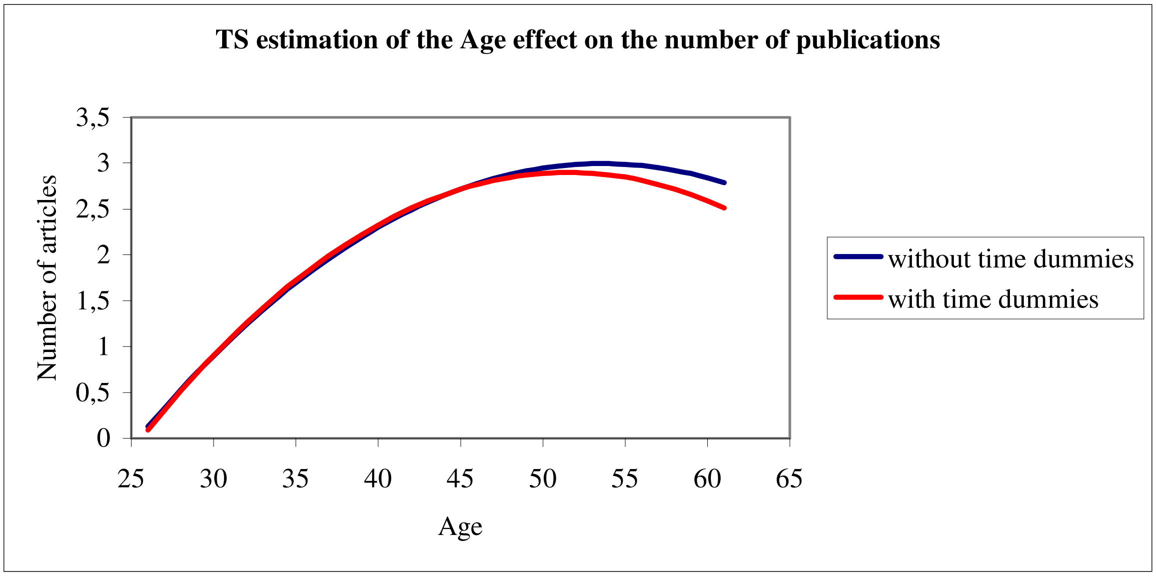
Table II.3

Variables	LOGLINEAR ON MCIT_2 TOTAL	TWO STEP	MARGINAL IMPACTS TOTAL	TWO STEP
AGE2	-0.022 (0.036)	0.006 (0.05)	-0.077	0.023
AGE3	-0.103** (0.043)	-0.028 (0.073)	-0.357	-0.097
AGE4	-0.172*** (0.049)	-0.019 (0.096)	-0.597	-0.066
WOMAN	-0.049* (0.029)	-0.111*** (0.03)	-0.172	-0.385
ECOLE	0.062* (0.033)	0.154*** (0.033)	0.215	0.536
DUMCH23	0.078** (0.033)	0.083** (0.034)	0.27	0.29
DR1	0.306*** (0.084)	-0.164 (0.134)	1.065	-0.571
DR2	0.101** (0.045)	-0.057 (0.067)	0.352	-0.197
C2ANCCR	-0.054 (0.038)	0.015 (0.052)	-0.187	0.052
C3ANCCR	-0.011 (0.05)	0.016 (0.078)	-0.038	0.054
C2ANCD2	-0.022 (0.044)	-0.13** (0.053)	-0.077	-0.451
C3ANCD2	0.023 (0.048)	-0.072 (0.083)	0.081	-0.251
C2ANCD1	-0.051 (0.096)	-0.183** (0.093)	-0.177	-0.638
C3ANCD1	-0.01 (0.097)	-0.355** (0.148)	-0.034	-1.234
MPETR	1.756*** (0.597)	1.944*** (0.62)	6.109	6.766
LOGNBCH	0.011 (0.021)	0.053** (0.022)	0.011	0.053
LOGNBCH2	0.015** (0.007)	0.018** (0.007)	0.015	0.018
LOGPROD	-0.104** (0.043)	-0.056 (0.044)	-0.104	-0.056
LOGQUAL	0.28*** (0.09)	0.301*** (0.093)	0.28	0.301
ADUMGREN	0.128*** (0.035)	0.141*** (0.037)	0.446	0.489
ADUMPAR	0.031 (0.033)	0.082** (0.035)	0.107	0.284
DUMEF13	0.29** (0.13)	0.544*** (0.133)	1.009	1.894

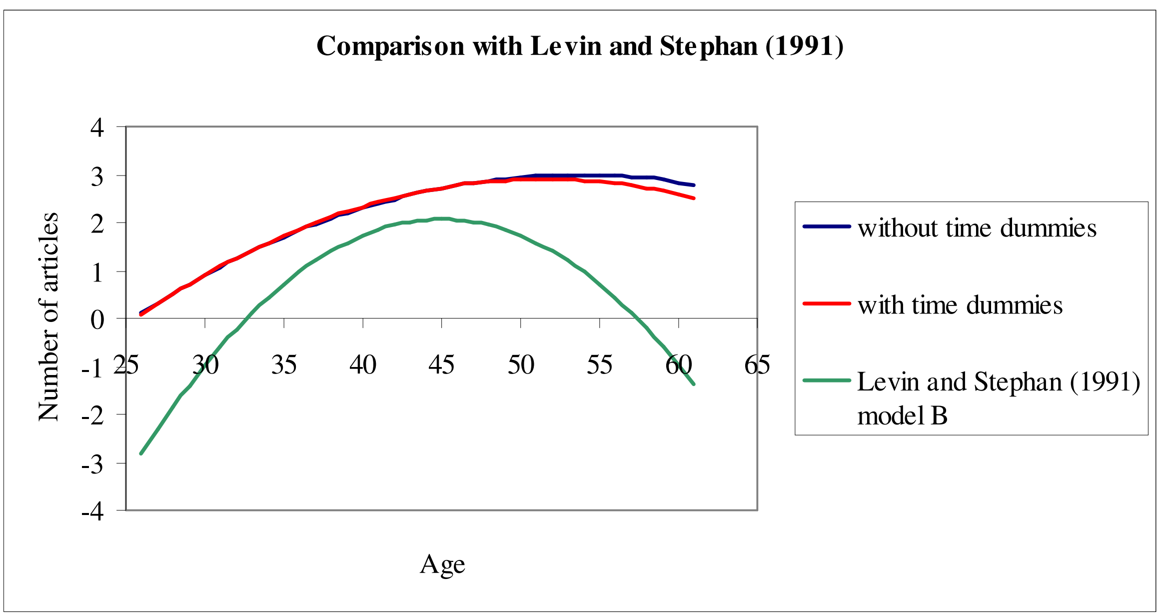
A87	0.042 (0.046)	0.046 (0.043)	0.147	0.16
A88	0.0001 (0.046)	0.008 (0.044)	0.0002	0.028
A89	0.01 (0.046)	0.022 (0.046)	0.036	0.077
A90	0.006 (0.047)	0.027 (0.05)	0.019	0.095
A91	0.0004 (0.047)	0.032 (0.053)	-0.001	0.11
A92	0.01 (0.048)	0.056 (0.057)	0.034	0.195
A93	-0.349*** (0.048)	-0.309*** (0.061)	-1.215	-1.073
A94	-0.209*** (0.049)	-0.17*** (0.065)	-0.729	-0.593
DUMMY (MCIT_2=0)	-1.233*** (0.024)	-1.095*** (0.027)	-4.29	-3.808
C	0.835*** (0.135)	0.549*** (0.133)	2.904	1.909
Log likelihood R^2-Adj	-3223.11 0.50	-2728.45 0.58 (step1) 0.46 (step2)		

♣ *Number of individuals = 352, Number of years = 9, Number of Observation = 3168*

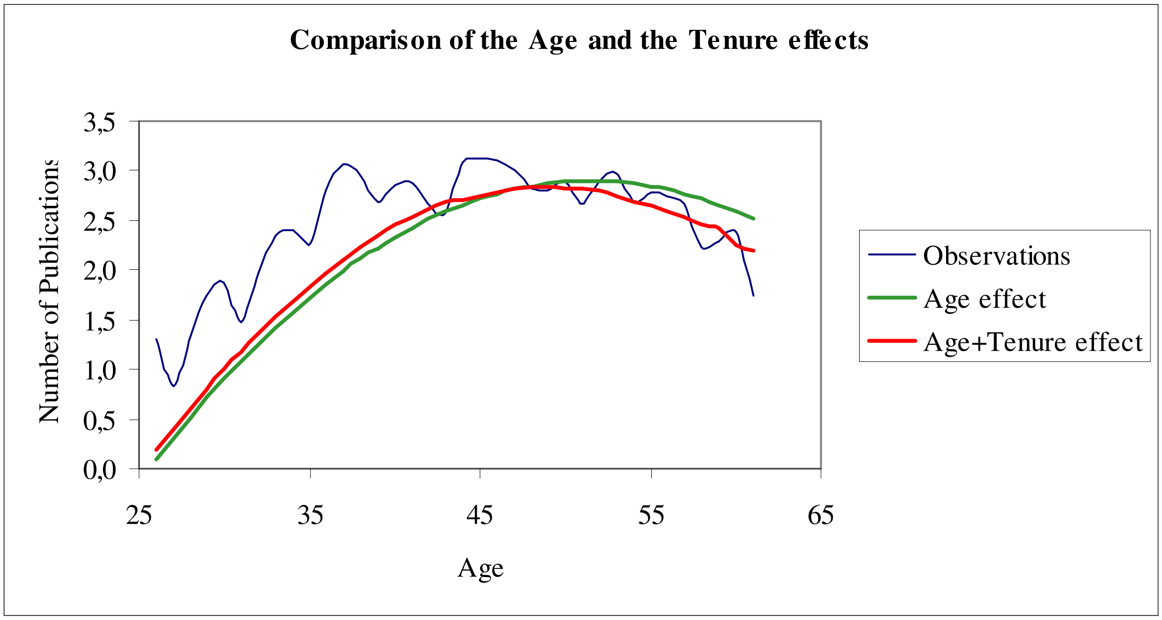
Graph II.1



Graph II.3



Graph II.3



Graph II.4