The Determinants of Research Productivity: A Case Study of Mexico

Claudia Gonzalez-Brambila
Advisors: Francisco Veloso and Granger Morgan

Abstract

This paper explores the determinants of research productivity using a data set of the most productive researchers in Mexico. In particular, an analysis of the dynamics of productivity over investigators' life cycle and an exploration of the differences among areas of knowledge is presented. Our findings confirm a quadratic relation between age of researchers and publications per year that other authors have found. However, according to our study, the decline of Mexican researchers' productivity begins when researchers are approximately 57 years old. This is 10 or 15 years later than what other studies have shown. This finding suggests that age is not very important in terms of research productivity, since researchers at 65 years old are as productive as those at 41. We also find that researchers are productive between 29 and 77 years old publishing a peak of 1.5 papers at 57. A comparison of research productivity life cycles among areas of knowledge shows important differences not only in the peak of publications, but also in the productive cycle. Researchers in Biology and Chemistry have the largest productive cycle, approximately 58 years, and peak the highest number of publications per year (2.4). Interpretations of other components of our model, such as gender, country of PhD, cohort effect, among others, are also presented. When adjusting for quality using the number of citations per publication, no important differences were found.

Introduction

The amount of funds dedicated to support science and technology (S&T) activities has increased enormously since Vannevar Bush's manifesto, "The Endless Frontier". In this seminal policy document, science is seen as the source of new knowledge that is published openly and drives the creation of new technologies that lead to economic growth. However, the last decades have shown that the social payoff for science and technology is neither direct nor clear. As a result, a necessity to asses the impact of S&T programs has grown. This has shaped a culture of evaluation and monitoring in research, where publications and citations are the most common ways to measure the importance of the contributions of a researcher or an organization.

Research output is evaluated and monitored at different levels and for different purposes. At the macro-level, it has been seen that governments have opted for increasing "project" funding of research, usually allocated on a competitive basis, in detriment of "institutional" funding[1] (OECD, 2002). To have access to this funding, researchers write and submit their research proposals to the national research funding entity. Then, these agencies request opinions for external referees and peer review committees about the most promising proposals. In this evaluation process there is evidence that shows that past publications have an important effect on the expected level of grant funding. (Arora, 1998-1 and 2).

At the micro-level, universities and research centers use publication and citation counts to monitor faculty performance and to give raises and promotions. Moreover, in ranking universities' departments, one of the most important measures is the aggregate number of publications and citations of their faculty. Publications are also important as a channel of communication with industry. Arundel and Geuna (2001) found that publications review is the method of obtaining the results of public research most frequently cited as important among firms in low, medium and high-tech sectors. Firms also use publications as a way of detecting expertise within universities, with subsequent hiring of faculty and graduates as consultants or employees.

Nevertheless, given the increasing importance of publications as a measure or research productivity, it is surprising how little it is known about the determinants of individual and collective research productivity.

Understanding the determinants of publications has important implications for administrators of universities and research laboratories, since they could forecast the expected productivity of researchers, taking into account individual characteristics and past history, and perhaps design policies to enhance productivity when a decline is predicted, or plan for a balance of ages in the groups to compensate for the potential existence of age effects. It can be particularly relevant for policy makers in countries

---

[1] Institutional funding refers to block funds that are allocated to universities or research centers on an annual basis and do not have strings.

where most of the research system is financed with public funds, since this information could help to design policies for the allocation of resources in order to enhance the productivity in the research establishment.

In the pursuit for more in-depth knowledge of the factors that condition individual research productivity, it is crucial to analyze a diverse set of countries because of the very different characteristics of the scientific community around the world, especially between the developed and less developed world (Nelson, 2003). This aspect has been overlooked by most existing research on the topic, which has analyzed only the developed world. For example, Mexico, as an advanced developing country, has a relatively small scientific community, with 0.6 researchers per 1,000 employees, while France has 6, and the U.S. 8.1 (Conacyt, 2002). The limited availability of research resources in developing countries is another issue. In Mexico S&T expenditure is 0.4% of the Gross Domestic Product (GDP), in the U.S it is 2.76% and in France 2.15% (Conacyt, 2002). Another reason is that developing countries have less access to some journals because of the journal selectivity, or policies of publication (Lawrence, 2003). Furthermore, the scientific community in many developing countries faces disadvantages due to differences in language and publication culture.

Another dimension that needs careful consideration in studies aiming at comparing the performance of researchers is the area of knowledge. A number of existing studies have documented how the pattern of publications varies significantly across areas (Cole, 1979; Xie and Shauman 1998). Nevertheless, as with the contrast between countries, previous work has not made a systematic effort to explore how the scientific area may condition the pattern of publication.

The purpose of this paper is to explore the determinants of research productivity using a data base of the most productive researchers in Mexico. Our main interest is to analyze the dynamics or productivity over the life cycle, and to explore the differences among areas of knowledge. To our best knowledge, this is the most comprehensive study of its kind, not only for the number of researchers in the sample, but also because it considers all areas of knowledge. It is also the first to look at those issues outside the developed world.

This paper is organized in five sections. The following section presents a literature review of some studies that address the determinants of productivity. Section 2 describes the data base of Mexican researchers; section 3 presents a descriptive analysis of the data. In section 4 we formally present the model that is used to estimate the effects of the explanatory variables in international publications. We present our results in Section 5, and our conclusions and suggestions for future research in Section 6. The results of the regressions are in the appendix.

1. Literature Review

The dynamics of research productivity over the life cycle is one of the aspects of individual productivity that has received more attention. Gary Becker (1962) and Theodore Schultz (1963) are the pioneers exploring how life cycle may condition productivity in occupations where human capital plays an important role. Their models suggest that human capital investment declines over time due to the finiteness of life, which would lead individual productivity to follow an inverted U shape pattern (Stephan, 1996). Subsequently, a number of researchers have tried to test for such effects in a number of professional contexts, in particular trying to understand if they would apply also to the academic profession.

Using a sample of American chemical engineering departments, Bernier et al. (1975) find that publications and citations peak for those in the 40-44 age group. The study also makes a first attempt to correlate measures of quality such as number of citations, number of PhD's graduated, or funds with each other, as well as with number of publications and with the peer evaluations of researcher quality. Overall, they find that quantity and quality are highly positively correlated.

Analyzing a cross-sectional data set involving American researchers drawn from six different fields, Cole (1979) finds that age has a slight curvilinear relationship with both quality and quantity of scientific productivity. He suggests that the small increase in productivity through the thirties and a corresponding limited decrease over the fifties, is explained by the scientific reward system, i.e. the reward system reduces the number of scientists who are actively publishing, so those who continue publishing are only the best members of their cohort.

One limitation of these earlier studies is their use of cross-sectional data, which make it possible that aging effects can be confounded with cohort effects (Stephan, 1996). For example, one type of cohort effect is associated with change in the knowledge base. Since there is a general presumption in science that the latest educated are the best educated (Levin and Stephan, 1991) and, as a result, better prepared to publish, regressions based in cross sections would attribute potential cohort effects to the younger age of researchers, biasing their productivity upwards. Another type of cohort effect is associated with the cumulative advantage, where past publications is an important factor to have access to research resources, so the more experienced researchers could be expected to be the more productive.

To counter these problems, more recent studies use longitudinal data on individual researchers. For example, studying the publishing activity of Berkeley mathematicians, Diamond (1986) finds a slight decline with age. Using a pooled model to estimate the growth rate of publications for 1000 Israeli scientists, Weiss and Lillard (1982), find a quadratic relation between the number of publications and the phase of the academic career.

One of the most important studies is Levin and Stephan (1991). They develop a model of scientific productivity that considers that scientists engage in research not only because of the future financial rewards associated with it, but also for the satisfaction of "solving the puzzle". Using longitudinal data of American scientists, they find that life cycles effects are present in five of the six areas of physics and earth sciences studied. This means that publishing activity initially increases and then declines somewhere in mid-career. In their model B, which considers straight publication counts, the solid state and condensed matter physicists reach a publication peak of 2 papers per year at age 45. However, the atomic and molecular physics reach a peak at age 39 and geophysicists at 59. The impact when they adjust for co-authorship, for journal quality or both is generally for the age peak to be reduced in 1 to 5 years. It is important to mention that they do not find strong evidence that the latest educated are the most productive.

In a very recent study, Tuner and Mairesse (2003) analyze the impact of research productivity relative to age, gender and education of French condensed matter physicists. They find that there is a quadratic relation between the age of the scientists and the number of publications. That is researchers' productivity increases before 50 and then declines after 51. The results using the average number of citations per two years are not significantly different from those of publications. Their results also suggest that there has been a wider and faster access to publication because of the increasing number of available journals. Finally, the results suggest that graduates from the French Grande Ecoles publish more and that a woman publishes almost 0.9 paper less than a man on average per year.

This last aspect of measuring sex differences in scientific productivity has also captured the attention of researchers. Several studies have found that female scientists publish at lower rates than male scientists. Using a sample of American biochemists, Long (1992) finds that sex differences in the number of publications and citations are bigger during the first decade of the career but are reversed later. Specifically, he finds that during the first three years after graduating, males have 26% more publications; between years three and four the percentage difference jumps to 66% and increases until 91% by the ninth year. Then, it steadily declines to 59% by the year seventeen. He attributes the lower productivity of females to their overrepresentation among non-publishers and their underrepresentation among the extremely productive. In a more recent study, Xie and Shauman (1998) find that sex differences in research productivity have declined, with the female-to-male ratio increasing from about 60 % in the late 1960's to 75 to 80% in the late 1980's and early 1990's. Their research also uses a sample of American scientists.

A slightly different perspective on understanding productivity dynamics is presented by Allison and Stewart (1974). Instead of looking at productivity over the life cycle, they use a cross-section of chemists, physicists, and mathematicians in the US, and find that the highly skewed distributions of productivity among researchers can be explained by a process of accumulative advantages. Highly productive researchers maintain or increase their productivity because they receive recognition and resources, while those scientists who do not, become less productive or have to leave their career as researchers. This inequality becomes increasingly unequal as career age increases. In an extension of this study, Allison et al. (1982) examine cohorts of biochemists and chemists, and they confirm that increasing inequality is observed for counts of publications but not for counts of citations to all previous publications, suggesting that scientist' older publications are cited with less inequality than their more recent work.

Buchmueller et al. (1999) develop a productivity model function where publications depend on the graduate school, the type of the first job after graduating, some personal characteristics, and unobserved factors. They empirically test their model using data of American PhD economists. Their findings suggest that graduates of the top ten programs who had some experience as research assistants are more productive. When they take into account the initial job, they find that economists employed in research universities are also more productive.

In Table 1 a summary of the determinants of scientific productivity that have been studied and the key findings is presented.

Table 1. Main Findings of Studies Related to the Determinants of Research Productivity.

Determinant	Author	Main findings
Age	Bernier et al. (1975)	Life cycle with a peak in the 40-44 group.
	Cole (1979)	Life cycle with a peak before 50's
	Levin and Stephan (1991)	Life cycle with a peak at 45
	Turner and Mairesse (2003)	Life cycle with peak at 50
Gender	Cole and Zuckerman (1984)	Women publish 57% as many papers as men.
	Long (1992)	Males publish between 26 and 91% more than women.
	Xie and Shauman (1998)	Female-to-male ratio from 60 to 80%
	Tuner and Mairesse (2003)	Women publish 0.9 less papers per year.
Cumulative advantage	Allison and Stewart (1974)	Inequality between productive and unproductive researchers increases with age
	Cole (1979)	Inequality between productive and unproductive researchers increases with age
Education	Buchmueller et al. (1999)	Graduates from top schools, with research assistant experience and employed in research universities are more productive
	Turner and Mairesse (2003)	Graduates from Grande Ecoles are more productive
Income	Stephan (1996)	Suggest that salary is positively related

		to both article and citation counts.
Cohort effect	Levin and Stephan (1991)	No evidence

The methodologies, the size of the samples, the period of time of study, and the limitations vary greatly among these studies. As was mentioned above, one of the most important limitations is the use of cross sectional data since it does not allow separating age and cohort effects. Moreover, most of these studies analyze cohorts before the 1990's and the dynamics of publishing activity has changed dramatically during the last twenty years. As was mentioned in the introduction, the pressure for publishing has increased, as well as the number of journals. In addition, some of the studies analyze publications in a short period of time (between 2 and 6 years), so some conclusions could be weak and inconclusive. The source of publications also varies among the studies. Some studies use the field abstracts, and others use self reported number of publications. Furthermore, as was mention before, all of these studies focus on developed world.

This paper tries to correct for most of the critical limitations of previous work, expanding in new directions. First, we are using a recent extensive panel data that allow us to look at the dynamics of productivity over the life cycle and for a comparison among different research areas of knowledge. In addition, we explore differences in scientific productivity considering gender, country where PhD was earned, cohort and the total number of researchers and publications in the same area of knowledge.

2. The data

We had access to information on 14,328 researchers, in all fields of knowledge, who have been part of the Mexican National System of Researchers (SNI) from 1991 to 2002.

The National System of Researchers was created in 1984 to enhance the quality and productivity of researchers in Mexico. It gives pecuniary compensation, as a complement of salary, to the most productive researchers[2]. SNI grants represent on average 30% of the income of researchers in the program. In 2001, 31% of researchers in Mexico were in SNI (Conacyt, 2002).

It is important to note that the researchers in SNI are not randomly distributed and do not represent the whole population of researchers in Mexico. On the contrary, given the characteristics of the program, it is expected that the most productive researchers are in the system.

The data are classified in two categories:

1. Characteristics of the researchers:

---

[2] To be included in the system, researchers apply to the program, then are evaluated by peer committees that determine which researchers can receive the benefits of the program and the appropriate level. The program ranks researchers in 4 levels, with compensations depending on level (Candidate, Level 1, 2, and 3, where Level 3 is the highest).

    - Gender
    - Institution
    - Age
    - Years since PhD was earned
    - Country where PhD was earned
    - Area and discipline[3]
2. SNI variables:
    - Their presence (or absence) and level in system for each year
    - Budget and lagged budgets of Conacyt[4]
    - Total number of researchers in SNI by area and discipline
    - Total number of publications by SNI researchers by area and discipline.

The source of publications and citations is the Science Citation Index produced by the Institute of Scientific Information (ISI). The publications were obtained by matching the data base of the researchers in SNI with Mexican articles from ISI database from 1981-2002. We also have access to all the citations that were made to each of the papers until 2002 (ISI, 2003).

3. Descriptive analysis

The mean of publications per year for the entire population is 0.174 with standard deviations of 0.62 (overall), 0.32 (between) and 0.52 (within) [N= 14,328[5]]. The distribution of publications is highly skewed. The lower extreme of productivity is represented by those who never have published. [n=5900][6]. Lack of publication may

---

[3] We use a broad classification of fields of knowledge as area. In this paper, we are using 6 different areas: Exact Sc., Biology and Chemistry, Health Sc., Social and Humanities, Agricultural Sc and Biotechnology, and Engineering. We refer to discipline as a field of knowledge, for example: Exact Sciences concentrate Mathematics, Physics, Earth Sc., Astronomy, and Material Sc.

[4] Conacyt is the National Council for Science and Technology in Mexico. This agency is in charge of the Science and Technology policy in Mexico and manages the SNI program among other programs to support science and technology activities.

[5] The statistics for the entire population are: 73% are males ; 95% are in public institutions; the distribution by level is the following: 7% level 3, 15% level 2, 54% level 1 and 24% candidates; 49% got their PhD in Mexico, 17% in the United States, and 30% in Europe; 0.5% got their PhD before 1960, 2% between 1961 and 1970, 8% from 1971 and 1980, 20% from 1981 to 1990 and 69% from 1991 to 2002. The average age is 41.11 years old (std. dev. 10 years). Related to areas of knowledge, the larger area is Social and Humanities with 25%, followed by Biology and Chemistry with 17%, and Exact Sc., Agricultural Sc. and Biotechnology, and Engineering with 16% each one; finally Health Sc. has the smallest number of researchers 10%.

[6] This subsample has similar distribution of the entire population. The average age is 41.34 years (std dev 10.1), 72% are males [there is no overrepresentation of nonproductive women as in the case of Long study], 94% are in public institutions, 4% are level 3, 11% are level 2, 53% are level 1, and 32% are candidates; 17% got their PhD in Mexico, 17% in the United States and 34% in Europe [Within the productive group there is an overrepresentation of researchers who got their PhD in Mexico and in other countries]. 0.5% got their PhD before 1960, 1% between 1961 and 1970, 6% between 1971 and 1980, 16% between 1981 and 1990, and 76% between 1991 and 2002. [It is possible that the latest educated are the less productive confirming a cumulative advantage process, however, given the missing data, we prefer not to make conclusions about this]. 7% are in Exact Sc., 7% in Biology and Chemistry, 3% in Health Sc., 45% in Social and Humanities, 21% in Agricultural and Biotechnology and 17% in Engineering. It is important

represent at least three types of activity: people may no longer be active in science; people are devoting most of their time in other type of activity such as teaching or consulting; finally, it may also mean that they have other type of output such as books or publications in journals not in ISI.

Because the main purpose of this paper is the analysis of the dynamics of productivity[7] over the life cycle, we limited our study to those researchers who have at least one publication in the ISI[8]. Given that we do not have information of those researchers who were in the system before 1991 and to have a balance panel, we are considering eleven years of publications from 1991 to 2001. As a result, our final sample consists of 7,793 researchers in all disciplines[9].

The mean of publications per year of this subsample is 0.489 with standard deviation of 0.965. The average age (40.3 years old) is slightly smaller than that of the entire population (41.11). The proportion of men (73%) is the same than the entire population. Related to the country where the PhD was earned in the subsample, we have that 51% got their PhD in Mexico, 17% in the United States and 27% in Europe[10]. In addition, most of the researchers in our sample got their PhD after 1990 (65%), 23% got their PhD between 1981 and 1990, 9% between 1971 and 1980, and only 3% before 1970[11]. Biology and Chemistry is the area of knowledge that concentrates the larger number of researchers (24%), followed by Exact Sc. that has 22% of the researchers. On the contrary, Social and Humanities is the smallest area with 10%. Table I in the appendix shows the main statistics of the variables used in the model.

Figure 1 presents the mean of publications per year for the entire population, from 1981 to 2001, showing the differences among areas of knowledge. As it can be seen, there has been an enormous increase in the mean of publications by Mexican researchers in SNI over the period studied. Although there is an increased tendency over the whole period, as of 1990 this growth is more noticeable. Health Sciences is the area of knowledge that concentrates the most productive researchers, followed by Exact Sciences, and Biology and Chemistry. On the other hand, Social and Humanities is the area with less productive researchers.

---

to mention that the most productive researchers are in Health Sc. and the less productive in Social and Humanities.

[7] In this paper we use publications in the Science Citation Index as a proxy of productivity.

[8] The bias that is caused because of this situation affects the peak of publications but not the shape of the curve.

[9] When the number of citations per four year is used, a subsample of 5,658 researchers and 7 years of publications (1991-1997) is drawn on.

[10] There is 27% of missing data, so it is possible to have a bias in the estimation of this parameter.

[11] There is 45% of missing data, so it is possible to have a bias in the estimation of this parameter.

4. The model

We assume that the function determining publishing proficiency P_{it} is given by:

$P_{it} = F(X_{it}, Z_i, c_i, u_{it})$, i identifies researchers and t year.

Z_i: Variables that are stable across time but not across researchers
area, gender, institution, country of PhD, cohort

X_{it}: Variables that vary in both dimensions
age, age^2, level, budget, total number of researchers, total number of publications.

c_i: is the individual unobserved effect which is stable across time but not across researchers

u_{it}; is the unobserved effect that varies in both dimensions

We use the negative binomial fixed effects model proposed by Hausman, Hall and Griliches (1984) because of the panel nature of our data. This model allows for both the possibility of permanent unobserved individual effect as well as the possibility that some unobserved effects may be correlated with publications and other explanatory variables. We chose the Negative Binomial distribution over the Poisson because the latter imposes a constant variance. This is not true for the data used in our study where the variance of productivity far exceeds the mean. One drawback of the Negative Binomial distribution is that our conclusions may be less precise since the estimated standard errors tend to be larger than in the alternative Poisson model. Another drawback comes from using fixed effects estimator because it does not take into account the between estimator, which considers the variation between the cross section observations. Specifically the model that we use is:

$$E(y_{it}|X_{it}, Z_i) = \exp(\mu + \gamma Z_i + \beta X_{it} + c_i + u_{it}) \qquad (1)$$

With y_{it} as a Negative Binomial

We assume that the errors are not serially correlated and allow for correlation between the unobserved time variant effect and the explanatory variables. This assumption is supported by the studies of Zener (1968 and 1970) who suggests that productivity is the result of many “mental factors” that vary across time. In addition, we assume that the specific individual constant effect is not correlated with the individual characteristics of the researcher. The only variable that could be endogenous is the institution, but since there is almost no mobility across institutions, and currently there are no vacancies in the biggest institutions, few researchers can choose the institution where they would like to work. In addition, we assume that there is no multicollinearity among the explanatory variables that vary across time and among time invariant explanatory variables. Specifically our assumptions are:

$$E(X'_{it} u_{it}) \neq 0$$
$$E(Z'_i c_i) = 0$$
$$E(u_{it} \mid X_{it}, c_i) = 0, \ t = 1,2,\ldots,T$$

$$\text{Rank } [E(X'_{it}X_{it})] = K$$
$$E\,(c_i \mid Z_i) = 0$$
$$\text{Rank } [E(Z'_iZ_i)] = K$$

We estimate β by the Conditional Maximum Likelihood Estimation (CMLE) proposed by Hausman, Hall and Griliches (1984), and γ by the Non Linear Least Square Method in a second step proposed by Turner and Mairesse (2003), replacing β by its CMLE estimate and estimating equation (2).

$$y_{it} / \exp\,(\beta\text{hat } X_{it}) = \exp\,(\mu + \gamma\, Z_i + c_i) \qquad (2)$$

Because one of our main interests is the analysis of the dynamics of productivity over the life cycle[12], we decided not to include time dummies since the estimation method that we are using estimates variables in deviation from their mean, thus time dummies variables and the age variable are collinear. Also, in previous results we included time dummies in the regressions and the change in the variables was not significant, so we report only the regression without considering time dummies.

The same model and assumptions are used to adjust for quality. In this case the dependent variable is the number of publications plus the number of citations per four years. This approach was used because on average publications receive 70% of cites in 4 years after publication. In addition, there are many publications that do not have any citation, so that to take into account all the publications in the period 1991-1997 for the subsample, we decided to add publications to citations per four years. One alternative approach was to consider the average number of citations per year, however, one inconvenient is that the latest publications could have smaller average number of citations because of the time since publication; this could cause an important bias in our results. Another approach could be to use the expected number of citations produced by ISI. We decided to use the number of citations per four years instead of this factor is because the correlation between these 2 indicators is not very high (0.6).

5. Results

5.1 Results for publications

The results of our regressions are shown in Table II in the appendix. In preliminary work we estimated publications regressions by Poisson and Least squares. The qualitative results i.e. the sign and significance of key parameters did not vary substantially across the different specifications. For considering that the negative binomial is the best model given the nature of our data, we restrict our attention and our comments to the negative binomial estimates.

---

[12] An alternative approach is the analysis of the dynamics of productivity considering years since PhD was earned instead of age. The results are similar. We report only the results of the dynamics over the life cycle to compare with other studies that have been done.

Age

As seen in Figure 2, the estimation confirms the quadratic relation between age of researchers and publications per year that other authors have found. According to our model, researchers are productive between 29 and 77 years old, publishing a peak of 1.5 papers. The decline occurs almost when researchers are 57 years old, this is 10 or 15 years later than what other studies have shown, an important difference. We believe that at least three different factors, not mutually exclusive, can help explain this dissimilarity. The first one is that the National System of Researchers encourages researchers to continue publishing; this is related to the fact that in Mexico the base salary of researchers is about one third of what they actually receive. The other two thirds are given in the form of pecuniary compensations from SNI and from the university where researchers work. Moreover, the system does not guarantee 'tenure'. Evaluation is done regularly and the researcher can be demoted if the panel evaluating him or her assesses lower research productivity. Thus, researchers are reluctant to retire and they have to continue showing some productivity in order to receive these complements of salary. A second explanation could be that the eldest researchers of a research group or of a laboratory tend to appear as coauthors in the publications of their colleagues more frequently than what would be found in more developed research systems such as in the U.S[13]. Finally, a third explanation could be that Mexican researchers get their PhD when they are older, at least in comparison to the US, the country used for most previous productivity studies. In our sample, the average age of PhD graduation is 36.1 years old with a standard deviation of 6.9 years. So that, it is possible that Mexican researchers start and finish their careers some years later than their colleagues in the developed world.

When looking at differences among areas of knowledge (Table 2), we see that the decline of productivity in Exact Sciences starts when researchers are 62 years old, although they got their PhD on average when they were 33.5 years old (5.6)[14]. They reach a maximum of 1.6 publications per year. For researchers in Biology and Chemistry, we can observe a smaller dispersion on their productivity than in other areas; they also reach the highest number of publications per year among all areas, about 2.4. The decline starts when they are 58 years old and the average age of graduation is 34.6 (5.5). In addition, researchers in Exact Sc. and Biology and Chemistry produce 0.37 papers more than Social and Humanities scientists. Researchers in Health Sciences have the smallest dispersion of all disciplines. They reach a maximum of 2 publications per year and publish almost half paper per year (0.46) than researchers in Social and Humanities. The decline starts at the age of 56, and the average age of graduation is 36.4 (6.8). Researchers in Social Sciences and Humanities reach a maximum of 0.9 publications per year, the decline starts when they are 58 and the average age of graduation is the highest 39.2 (8.2). Researchers in Agricultural Sciences and Biotechnology are the less productive of all the areas reaching a maximum of 0.8 publications per year, the decline starts when they are 56 and the average age of graduation is 36.9 (5.7). Finally, researchers in Engineering reach a

---

[13] This means that less senior researchers tend to include older researchers as coauthors, even if they do not contribute much to the work.

[14] Standard deviations are presented in parenthesis.

maximum of 1.2 publications per year, the decline starts when they are 55, and the average graduation age is 34.2 (5.3).

As it can be observed, other important difference is that researchers in Biology and Chemistry are the most productive, not only because they reach the highest peak of publications (2.4 publications per year), but also because their productive cycle is the longest (58 years). Researchers in Health Sc. are the second more productive group. They reach a peak of 2 publications per year in a cycle of 50 years. Besides, these two groups are the ones to start publishing at a younger age, 25 and 28, respectively. (Table 2). Similar behavior is seen amongst researchers in Exact Sc., as they reach a peak of 1.6 publications per year and are productive for 48 years. However, they start publishing at 34 years old.

The other three areas of knowledge, Social and Humanities, Agricultural Sc. and Biotechnology, and Engineering are less productive. Researchers in Social and Humanities have the shortest productive cycle (37 years) and are the ones that have the latest start in publishing (35 years old). Researchers in Agricultural Sc. and Biotechnology are the ones that reach the lowest peak of publications (0.8 publications per year). Reaching a peak of 1.2 publications per year at 55 years old, the productive cycle of engineers ends earlier, when they are 70 years old.

This situation could suggest that knowledge in Exact Sc., Biology and Chemistry, and Health Sciences is more codified, and its transmission is easier. Thus, researchers can publish more, for more years, and since they are younger. Another possibility is that researchers in Social and Humanities, Agricultural Sc. and Biotechnology, and Engineering choose research topics that are of local interest (regional or country level), and their results tend to be diffused in other media. On the contrary, knowledge in Exact Sc., Biology and Chemistry, and Health Sciences tends to be more universal.

Although the analysis of all areas of knowledge is interesting, we decided to analyze some disciplines in different areas in order to have a better understanding of the differences among areas of knowledge. The disciplines that were analyzed are Physics, Biology and Materials Engineering. The reason of choosing Physics is that we can compare[15] our results with those of Levin and Stephan (1991), and Turner and Mairesse (2003). Biology was chosen because it is the most productive discipline. Materials Engineering was also chosen because is the discipline in Engineering that has the greatest number of publications.

Researchers in Materials Engineering are increasingly productive until they are 57, reaching a maximum of publications of 1.5 publications per year. The average age of graduation is 34.6 (5.4). Biologists reach a maximum of 2.7 publications per year, they are increasingly productive until they are 58, and their average age of graduation is 34.8

---

[15] Since we are considering all physics and not the subdisciplines of physics, this is not totally comparable. Besides, Levin and Stephan (1991) use a pooled cross sectional Tobit model with two years of publications in the 1970's. In the case of Turner and Mairesse (2003), they use a Poisson model with publications from 1986 to 1997.

(5.49). Physicists are also increasingly productive until they are 62, reaching a maximum of 1.6 publications per year, and the average graduation age is 32.9 (5.2). These figures are quite different from the results of Levin and Stephan (1991) (model B in their paper), where they find that solid state and condensed matter physicists reach a peak of 2 papers at 45. It is also different from Turner and Mairesse (2003), where condensed matter physicists reach a peak of 2.9 publications at 52 years old[16]. All the results for the Mexican researchers show that their publishing peak occurs much later, and the decline is quite subtle until reaching the retirement age (65 in Mexico), where they are as productive as when they were 41 and more than when, on average, they got their PhD.

Table 2. Publishing Dynamics in the Different Areas of Knowledge.

Area of knowledge	Researchers are productive between: (years old)	Peak of number of publications per year	The peak in productivity is at age: (years old)
All	29-77	1.5	57
Exact Sc	34-82	1.6	62
Biol & Chem	25-83	2.4	58
Health Sc.	28-78	2	56
Social and Hum	35-72	0.9	58
Agric & Biotech	33-72	0.8	56
Engineering	32-70	1.2	55
Disciplines			
Physics	34-82	1.6	62
Biology	25-82	2.7	58
Materials Eng	29-75	1.5	57

As mentioned earlier, our results suggest three possible explanations. The first one, related with the salary structure in Mexico, cannot easily be tested because of the lack of an adequate control sample. The second one, that suggests that the eldest researchers of a research group or a laboratory tend to appear as coauthors in the publications of their colleagues, will be tested in future research. The third one, that supposes that Mexican researchers start and finish their careers some years later than their colleagues in other countries, was tested by running separate regressions depending on the age when the PhD was earned. We divided the sample into two groups: those researchers who got their PhD when they were 30 years old or younger, and those who were above 30.

The results suggest that the decline for those in the first group is even two years later. In addition, they reach a peak of 0.3 publications per year more than the second group. A break down of the estimation by areas shows that in Exact Sciences, Biology and Chemistry, and Social and Humanities researchers who got their PhD younger are more productive and the decline starts a couple of years later. This evidence contradicts our hypothesized explanation. However, in Health Sciences, Agricultural and Biotechnology

---

[16] This result was gotten from an annex regression in which age and age squared replaced the age cohorts.

as well as in Engineering, we find the opposite. The productivity decline for researchers who got their PhD younger starts almost ten years before the second group, while those who got their PhD later are more productive.

Thus, these results suggest that although the age of graduation is more important for some areas of knowledge than for others, there are not systematic differences, suggesting that such aspect is not the critical factor determining the life cycle research productivity effects.

Gender

Our results suggest that there is not a big gender difference in scientific productivity. As noted above, Mexican female scientists are not overrepresented among the non publishers, and they produce only slightly fewer papers (0.08 paper less than men on average per year). The proportion of female scientists varies a lot among areas of knowledge. The areas with greater representation are Health Sc. with 39%, Social and Humanities with 38%, followed by Biology and Chemistry with 36%. In Agricultural Sc. and Biotechnology 22% are women and in Exact Sc. 15%. Although, only 12% of engineers are women, our results suggest that women in Engineering are slightly more productive than men[17]. The biggest gap is in Health Sciences where women publish 0.24 paper less than men. Women physicists publish 0.16 paper less than men per year. Comparing this result with that of Turner and Mairesse (2003), they found that a woman publishes almost 0.9 paper less than a man on average per year. In their study the proportion of women is 18% and in our case is 11%. However, we do not have any explanation about this situation. As was mentioned by Cole and Zuckerman (1984), sex differences in research productivity continues being “the productivity puzzle”.

Country of PhD

Our results suggest that researchers who got their PhD in the United States or Europe publish about 0.2 paper less than those who got their PhD in Mexico or in other countries. Although this is a very small difference, it could be interpreted in several ways. First, it could be that the most productive researchers who got their PhD outside Mexico stay in other countries (brain drain). Another possibility could be that it is easier for those researchers who got their PhD in Mexico to find partners in Mexico whom to publish with, or have a better knowledge of how to obtain research resources. Definitely these suggestions must be interpreted very carefully since there is not enough evidence to support them.

It is important to note that there is a difference in the average age of graduation depending on the country where PhD was earned. For those who got their PhD in Mexico, the average age of graduation is 38.2 (7.4). For those who graduated from American or European universities, the average age is 34.8 (5.4) and 33.6 (5.7), respectively. This could be explained if we look at the proportion of researchers in each

---

[17] In Social and Humanities the parameter is also positive but it is not significant at 10%.

area of knowledge who got their PhD abroad. The two most productive areas Biology and Chemistry, and Health Sciences have the less proportion of researchers who got their PhD outside of Mexico, 30% and 32%, respectively. The largest proportion is in Engineering (69%), followed by Exact Sc. and Agriculture Sc. and Biotechnology with 58% each one. 48% of the researchers in Social and Humanities got their PhD abroad.

Level in SNI

The results of the regression suggest that researchers in level 3 produce 0.39 papers more than candidates and are slightly less productive than researchers in level 1 or 2; researchers in level 2 publish 0.45 papers more than candidates; and researchers in level 1 publish 0.42 papers more than candidates. The difference between level 2 and level 1 is small, this could signify that the big jump in research productivity is seen when researchers change from candidate to level 1. However, when adjusting for quality, our results suggest that researchers in level 3 receive the larger number of citations, (see results for citations). It is important to stress that in a survey that was sent to the peer committees of SNI, it was found they take into account the entire history of publications to determine the level. Because we are not evaluating the adequacy of the program, the cumulative number of publications is not considered in this work. The reason to include the levels in the program was to control for the fact that researchers are classified in the program. In addition, it is highly possible to have confounded effects between age and level in the program. The average age of researchers in level 3 is 58.7 (10.6), in level 2 is 50.6 (8.9), in level 1 is 44.4 (7.9), and in level candidate is 34.8 (3.8).

Cohort Effects

We created five cohorts according to the decade when the PhD was earned, i.e., we created dummy variables to control for those researchers who got their PhD before 1960, for those between 1961 and 1970, for those between 1971 and 1980, for those between 1981 and 1990, and for those who got their PhD after 1991. Our results suggest two types of cohort effect. On one side, researchers who got their PhD before 1960 are more productive (0.46 paper more) than the latest educated, i.e. those who got their PhD between 1991 and 2002. This could be associated with the cumulative advantage, where the past publications matter to get access to research resources and be more productive. This could be also related to the possibility that the more experienced researchers are the leaders, and they appear as coauthors of younger researchers. Considering that only 0.5% of the researchers in SNI got their PhD before 1960, we could also say that only the most productive researchers in each cohort stay in the system.

On the other side, our results suggest that the latest educated are the best educated due to the change in the knowledge base. We find that researchers in the cohort 1971-1980 publish 0.2 paper less than those in the cohort 1991-2002, and the cohort 1981-1990 are slightly less productive than the latest educated (0.08 papers less)[18]. It may also reflect that this cohort is under greater pressure to publish.

---

[18] Our result for the cohort 1961-1970 suggest that also this generation is less productive, however this result is not significant at 10%.

In the analysis of areas of knowledge, our results vary a lot in significance for this control. The areas of knowledge where we get significant results in most of the cohorts are Exact Sc. and Social and Humanities. For researchers in Exact Sc. our results support the general presumption that the latest educated are the best educated and the more productive. The cohort 1991-2002 is more productive than any other cohort, being the greatest difference with the cohort that got their PhD before 1960. This latter cohort publishes 0.57 paper less. In Social and Humanities our results are similar than our results for all the areas together. The cohort 1991-2002 is more productive than cohort 1971-1980 (0.34 paper less) and cohort 1981-1990 (0.07 paper less). However, the cohort before 1960 is more productive than the cohort 1991-2002, publishing 0.47 paper more.

Critical Mass

Our estimates suggest that although small, there is a positive and significant externality effect in the productivity when we consider the total number of publications in the same area of knowledge in all areas except in Physics. On the contrary, it seems that there is a little effect of saturation of the system when we control for the total number of researchers. Nevertheless, all significant results, including those for individual areas of knowledge, are small and negative. This is a surprising result since the scientific community in Mexico is relatively small and one could, in principle, expect the presence of some type of critical mass effect as the system grows. But it is important to note that the evaluation was done at a very aggregate level, which may not be the best to observe researcher critical mass effects. Researchers tend to cooperate more with colleagues in the same institution or in the same region, not necessarily in this broad area. Hence, this broad level of the variable may instead be measuring some degree of competition effect. Although the community is small, researchers compete for scare resources needed for publishing in a particular area, so that more people means less individual funding. Nevertheless, we feel that any potential interpretation of this variable should be taken with care. Future work will specifically address potential critical mass aspects. There, specific controls for institution and area will be used to better identify these aspects.

Budget

We tried to see the impact of lagged budget of S&T[19] in the productivity of researchers. In previous work we ran regressions considering different lags and all but the one year lag were not significant. Thus, we decide to run the final regression only considering the one year lag. Our results suggest that there is a significant and positive relation between the budget of the previous year and publications. However, this control does not make a distinction of the resources that are invested in each area of knowledge or, even better, the resources that each researcher has. So, our conclusions in this topic are weak.

---

[19] We are using the annual budget of Conacyt, since is the main agency in Mexico that supports S&T activities.

5.2 Results for citations

The results of the regressions using the number of citations per four years are shown in Table III in the appendix[20].

The estimation also confirms a quadratic relation between age of researchers and citations that was found using publications. The peak of citations per four years is 0.3 and is reached when researchers are 58 years old; this is only one year older than what was found for citations. This result could suggest that the SNI researchers continue publishing not only because of the incentives of receiving the support of the program but also for the satisfaction of "solving the puzzle".

Our results also suggest that publications of researchers in level 3 receive more citations than any other level (0.28). Publications of researchers in level 2 and in level 1 receive 0.21 and 0.14 more citations than candidates, respectively. When the areas of knowledge are analyzed, the most important difference among levels in the program is seen in researchers in Exact Sc. Researchers in level 3 receive 0.52 more citations, in level 2 0.45 and in level 1 0.31, than candidates. On the contrary, the area with less difference is Agricultural Sc. and Biotechnology. In this case, researchers in level 3 receive 0.12 more citations, in level 2 0.08 and in level 1 0.06, than candidates. This is an important result because as researchers become more senior, it is not the count of publications that matters, but rather the impact of those publications.

Our results also suggest that researchers in Health Sc. receive the largest number of citations per four years. They have received 0.44 more citations than researchers in Social and Humanities. Researchers in Exact Sc. and Biology and Chemistry have received 0.33 and 0.34 more citations than Social and Humanities' researchers, respectively. Engineers and researchers in Agriculture Sc. and Biotechnology have received only 0.15 and 0.16, more citations, respectively, considering the same reference These results could suggest, as in the case of publications, that knowledge in Exact Sc., Biology and Chemistry, and Health Sciences is more codified and universal, and its transmission is easier; and researchers in Social and Humanities, Agricultural Sc. and Biotechnology, and Engineering choose research topics that are of local interest (regional or country level), and their results tend to be diffused in other media.

As in the case of publications, there is only a slight difference between Mexican female scientists and their male colleagues. They receive only 0.06 less citations. The area of knowledge that has the largest gap is Health Sc. where women receive 0.14 less citations than men. However, in Social and Humanities, as well as in Engineering, women receive slightly more citations than men, 0.02 and 0.04, respectively.

Our results also suggest that there is almost no difference in citations depending on the country where PhD was earned. However, there is some difference depending on the

---

[20] In preliminary work we dropped those publications that have more than 50 citations over 4 years. However, the results do not vary substantially. The final regressions were done including the entire subsample of 5, 658 researchers and 7 years of publications (1991-1997)

cohort. Those researchers who got their PhD before 1960 receive 0.1 less citations than those who graduated after 1991; those who got their PhD between 1961 and 1970 receive 0.11 more citations, and those who got their PhD between 1971 and 1990 receive 0.12 more citations. This could be associated with the cumulative advantage, suggesting that eldest researchers are more prestigious and have published more, so they receive more citations than the latest educated. However, researchers in Social and Humanities who got their PhD before 1960 receive more citations (0.11) than the latest educated; and those who got their PhD between 1960 and 1970, and between 1970 and 1980 receive 0.12 and 0.16 less citations than the last cohort. These effects in conjunction with the results considering only publications suggest that there are potential cohort effects. On one side, senior researchers may have more prestigious and get access to more financial resources to do research. On the other hand, the latest educated can be the best educated due to the change in the knowledge base and may be under greater pressure to publish.

6. Conclusions

An analysis of the determinants of researcher productivity using a data base of the most productive researchers has been presented. We find that although there is a quadratic relation between age and researcher productivity, the effect of age is not very important: SNI researchers at 65 years old are as productive as those at 41 and more than when they finish their PhD. This result is consistent not only for count in publications but also for citations. Thus, we can conclude that researchers in retirement age have the same satisfaction of "solving the puzzle" and the same economic incentive that their colleagues who are 25 years younger. This is a very different conclusion to other authors. For example, Levin and Stephan (1991) find that the age effect is very important. This could either suggest that research activity over the life cycle is not investment-motivated or that it is, but the system in Mexico has created incentives for people to continue such investment until the end of their working lives. These findings could have important implications for the design of public policy, mainly in countries where the research system is funded by public funds, because it seems that it is possible to create programs to enhance productivity until the end of researchers' working lives.

We also find that there are significant differences in research productivity among areas of knowledge, not only in the peak of publications and citations but also in the productive cycle. In addition, the proportion of non-productive researchers is also dissimilar. These results suggest that researchers in different areas of knowledge have singular incentives, at least to publish in ISI journals. Besides, it seems that the nature of knowledge is different. There are some areas, such as Exact Sc., Biology and Chemistry, and Health Sc., where knowledge is more codified and is more universal, so that the outcomes of research could be publish and diffuse more extensively than knowledge in other areas, such as Engineering, Social and Humanities and Agricultural Sc. and Biotechnology.

When comparing with similar studies in the U.S and France, we found that the difference in the peak of publications between Mexican and American physicists is only 0.4 publications per year. However, when comparing with the French, Mexicans are 1.3

papers per year less productive. Considering that the study for the U.S. was done with data of the 1970's and the one in France with data of the late 1980's and 1990's, we could conclude that the pressure for publishing has increased in the last decade, both in developed and developing countries.

Our findings related to age, in conjunction with the results of the other variables, mainly the cohort effects, have important implication in terms of the planning of higher education institutions. For example, the concern that the average age of researchers has increased, could be something that is not a major distress in terms of count of publications.

However, these results must not be considered completely conclusive since, as we mentioned before, our sample includes only the most productive researchers. In addition, it must also be taken under consideration that publications are not the only product of research. Although number of publications and citations are the most commonly used measures, other factors such as the number of coauthors, author position, and other type of productivity output such as books and patents, need to be considered to get stronger results.

Furthermore, it is possible that the reward structure of the research establishment responds to initiatives like SNI, and researchers choose to allocate time to those activities that maximize future rewards. However, it is possible that those activities not necessarily are the ones that produce greater social payoff. For example, a researcher can choose between allocating time to publishing or to teaching. Perhaps these activities complement each other but if not the current reward system could not be giving more incentives to those activities that produce the higher social payoff.

In future work we are going to include an analysis of co-authorship. We will explore if researchers publish with partners in the same institution and in the same area or discipline, and how important the multidisciplinary work is. As part of this, we will analyze the critical mass effects. In addition, this work will include an exploration of the international cooperation of Mexican researchers.

We are also interested in studying the effects of budget in publications. This has important public policy implications since the problem that policy-makers face is not only how much resources should be invested in S&T, but also how to manage the most efficient portfolio of programs.

Acknowledgement

Support from the Mexican Council for Science and Technology (Conacyt) is gratefully acknowledges.

7. References

Allison, P., Long, S. and Krauze, T., 1982, "Cumulative Advantage and Inequality in Science", American Sociological Review, 47, 5, 615-625.

Allison, P. and Stewart J., 1974, "Productivity Differences Among Scientists: Evidence for Accumulative Advantage", American Sociological Review, 39, 596-606.

Arora, A., David, P, and Gambardella, A. 1998, "Reputation and Competence in Publicly Funded Science", *Annales d'Economie et de Statistiques*, No. 49/50, 163-196

Arora, A. and Gambardella, A., 1998, "The Impact of NSF Support for Basic Research in Economics", 1-24.

Arundel, A. and Geuna, A., 2001, "Does Proximity Matters for Knowledge Transfer from Public Institutes and Universities to Firms? SPRU, University of Sussex.

Bernier, C., Gill, W. and Hunt, R., 1975, "Measures of Excellence of Engineering and Science Departments: A Chemical Engineering Example", Chemical Engineering Education, 194-197.

Buchmueller, T., Dominitz, J. and Hansen, L., 1999, "Graduate Training and the Early Career Productivity of PhD Economists", Economics of Education Review, 14, 65-77.

Bush, V., 1945, "Science, The Endless Frontier", Washington, D. C., Office of Scientific Research and Development.

Cole, S., 1979, "Age and Scientific Performance", The American Journal of Sociology, 84, 4, 958-977.

Cole, J. and Zuckerman, H., 1984, "The Productivity Puzzle: Persistence and Change in Patterns of Publications of Men and Women Scientists", Advances in Motivation and Achievement, 2, 217-258.

Conacyt, 2002, Indicadores de Actividades Científicas y Tecnológicas 2002, Mexico.

Dasgupta, P., and David, P, 1994, "Towards a New Economics of Science", Research Policy, 23, 487-521.

David, P., 1994, "Positive Feedbacks and Research Productivity in Science", in Grandstrand, O. (ed.), Economics of Technology, North-Holland, Amsterdam and London.

Diamond, A. M., 1986, "The Life-Cycle Research Productivity of Mathematicians and Scientists", Journal of Gerontology, 41, 520-525.

Hausman, J., Bronwyn, H. and Griliches, Z, 1984, "Econometric Models for Count Data with an Application to the Patents-R&D Relationship", Econometrica, 52, 909-938.

Institute of Scientific Information, 2003, "National Citation Report for Mexico", Articles from ISI database years 1981-2002 cited in the period 1981-2002.

Jaffe, A., 1998, "Measurement Issues", in Branscomb, L. and Keller, J. (ed.), Investing in Innovation: Creating a Research and Innovation Policy that Works, MIT Press, Cambridge, Massachussets, and London, England, 64-84.

Jaffe, A., 2000, "Building Program Evaluation into the Design of Public Research Support Programs". Prepared for Technological Policy and Innovation: Economic and Historical Perspectives.

Kuhlmann, S., 2002, Strategic Evaluation: Concepts and Experiences with Regard to Priority Setting and Funding Mechanisms, Presentation in Berlin, May 6/7, 2002, Fraunhofer-Institute System and Innovation Research and Universiteit Utrecht-Copernicus Institute for Sustainable Development and Innovation.

Levin, S. and Stephan, P., 1991, "Research Productivity Over the Life Cycle: Evidence for Academic Scientists", The American Economic Review, 81, 114-132.

Long, S., 1992, "Measures of Sex Difference in Scientific Productivity", Social Forces, 71, 1, 159-178.

Nelson, R.(Editor), 1993, "National Innovation Systems: A Comparative Analysis", Oxford University Press.

OECD, 2002, Ad Hoc Working Group "Steering and Funding of Research Institutions" Governance of Science Systems: New Challenges. Detailed Outline of the Final Project Report. Paris.

Salter, A. and Martin, B., 2001, "The Economic Benefit of Publicly Funded Basic Research: A Critical Review", Research Policy, 30, 509-532.

Stephan, P., 1996, "The Economics of Science", Journal of Economic Literature, XXXIV, 1199-1235.

Turner, L. and Mairesse, J., 2003, "Individual Productivity Differences in Scientific Research: An Econometric Study of the Publications of French Physicists", Working Paper.

Weiss, Y. and Lillard, L., 1982, "Output Variability, Academic Labor Contracts, and Waiting Times for Promotion", Research in Labor Economics, 5, 157-188.

Xie, Y. and Shauman, K., 1998, "Sex Differences in Research Productivity: New Evidence About an Old Puzzle", American Sociological Review, 63, 847-870.

Zener, C., 1968, “An Analysis of Scientific Productivity”, Texas A&M University, College Station.

Zener, C., 1970, “Statistical Theories of Success”, Proceedings of the National Academy of Sciences, 66, 1, 18-24.

8. Appendix

Figure 1. Evolution of the mean of ISI publications per year by SNI authors

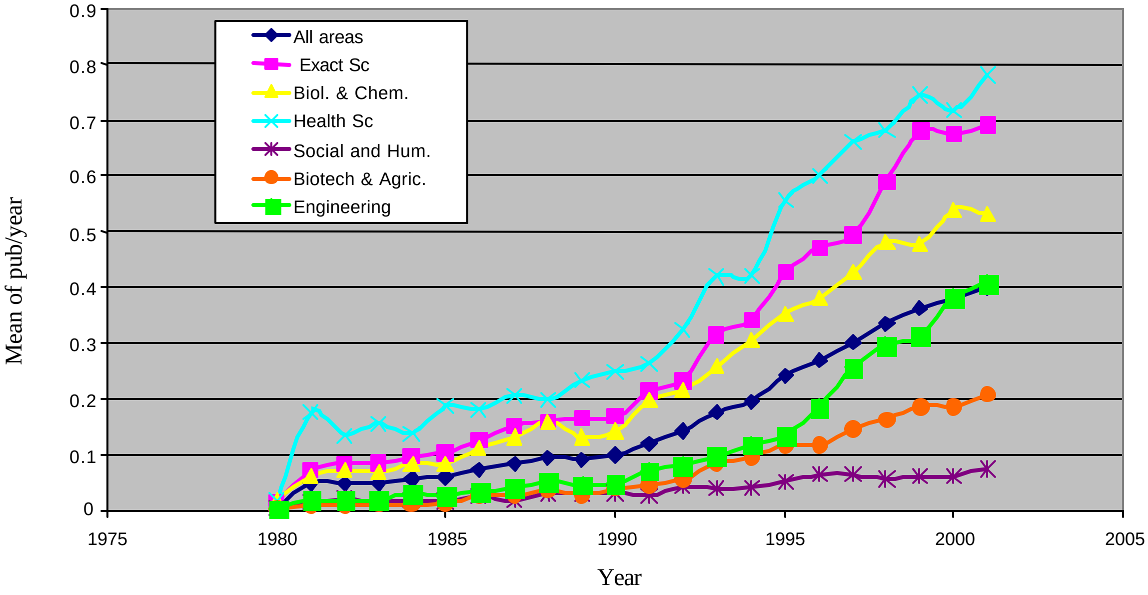

Figure 2. Research Productivity Over the Life Cycle: All Areas of Knowledge

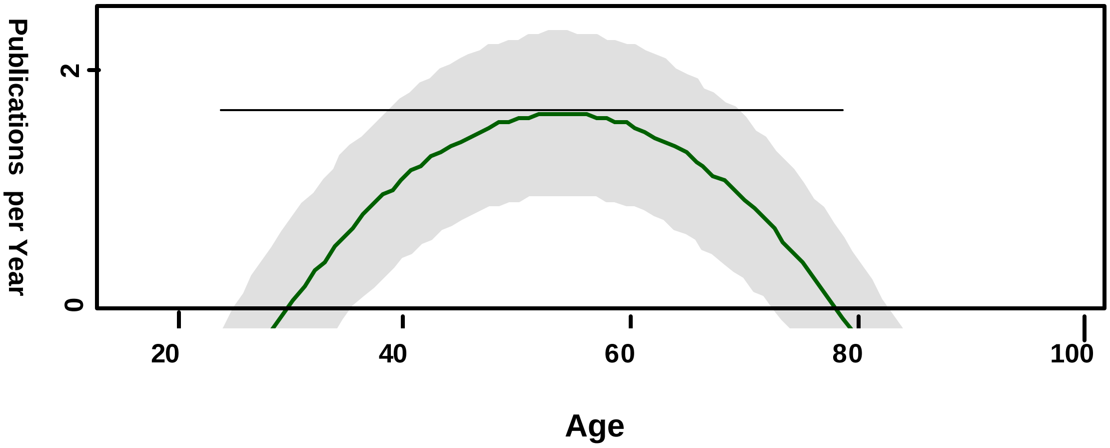

Table I.

Variable	Mean	Std. Dev.
Time variant variables		
Publications per year	0.489	0.965
Age	40.321	9.817
age_2	1722.187	884.546
Level 3	0.079	0.270
Level 2	0.171	0.376
Level 1	0.538	0.499
Level Candidate	0.211	0.408
Budget in Millions of dollars (2001)	332.351	58.593
Budget -1 year in Millions of dollars (2001)	308.723	82.168
Budget -2 years in Millions of dollars (2001)	286.996	101.046
Time invariant variables		
Exact Sc.	0.221	0.415
Biology and Chemistry	0.244	0.429
Health Sc.	0.171	0.376
Social and Humanities	0.097	0.296
Agriculture and Biotechnology	0.118	0.322
Engineering	0.150	0.357
Male	0.730	0.444
Female	0.270	0.444
Public Institution	0.967	0.178
Private institution	0.032	0.177
PhD in Mexico	0.508	0.500
PhD in USA	0.173	0.378
PhD in Europe	0.271	0.445
Cohort before 1960	0.004	0.065
Cohort 1961-1970	0.023	0.150
Cohort 1971-1980	0.092	0.289
Cohort 1981-1990	0.226	0.418
Cohort 1991-2002	0.655	0.475
Total # of years in SNI (1991-2002)	6.762	3.890

N= 7,793. Total number of publications (1991-2001) >0

Table II. Publications

Variables	Coef.	Std. Err.		Marginal Impact
Time variant variables				0.9814
Age	0.2637	0.0093	***	0.2588
Age Square	-0.0025	0.0001	***	-0.0024
Level 3	0.3950	0.0475	***	0.3877
Level 2	0.4624	0.0304	***	0.4538
Level 1	0.4302	0.0198	***	0.4222
Total # of pub	0.0005	0.0000	***	0.0005
Total # of researchers	-0.0003	0.0000	***	-0.0003
Budget -1 year in Millions of dollars (2001)	0.0009	0.0001	***	0.0009
Constant	-5.9628	0.2117	***	
Time invariant variables				
Exact Sc.	0.3755	0.0376	***	0.3686
Biology and Chemistry	0.3784	0.0371	***	0.3714
Health Sc.	0.4723	0.0390	***	0.4635
Agriculture and Biotechnology	0.0417	0.0439		0.0410
Engineering	0.1672	0.0408	***	0.1641
Female	-0.0768	0.0202	***	-0.0754
Public Institution	0.0300	0.0545		0.0295
PhD in USA	-0.2121	0.0302	***	-0.2082
PhD in Europe	-0.1895	0.0250	***	-0.1860
Cohort before 1960	0.4733	0.1463	***	0.4645
Cohort 1961-1970	-0.1337	0.0834		-0.1312
Cohort 1971-1980	-0.2104	0.0448	***	-0.2065
Cohort 1981-1990	-0.0856	0.0284	***	-0.0840
Constant	-2.0235	0.0612	***	-1.9859

The base is a male researcher in Social and Humanities, in a private institution, who got his PhD in Mexico or other country between 1991 and 2002.

N=7,793, T=1991-2001

*** Significant at 1%, ** significant at 5%, * significant at 10%.

Table II. Continuation

	Exact Sc.			Biology and Chemistry			Heath Sciences		
	Coef.		Marginal Impact	Coef.		Marginal Impact	Coef.		Marginal Impact
Time variant variables			1.059			1.336			1.350
Age	0.2535	***	0.268	0.2737	***	0.3656	0.2838	***	0.383
Age Square	-0.0023	***	-0.002	-0.0025	***	-0.0034	-0.0027	***	-0.004
Level 3	0.9264	***	0.981	0.1550		0.2070	-0.0377		-0.051
Level 2	0.9509	***	1.007	0.2183	***	0.2917	0.0908		0.123
Level 1	0.8149	***	0.863	0.2187	***	0.2922	0.2025	***	0.273
Total # of pub	0.0001		0.000	0.0003	**	0.0004	0.0008	***	0.001
Total # of researchers	0.0001		0.000	-0.0004		-0.0005	-0.0008	***	-0.001
Budget -1	0.0012	***	0.001	0.0008		0.0011	0.0009		0.001
Constant	-6.1644	***	-6.526	-5.4228	***	-7.2444	-5.8518	***	-7.899
Time invariant variables									
Female	-0.1499	***	-0.159	-0.1130	***	-0.085	-0.1808	***	-0.244
Public Institution	0.5952	***	0.630	0.4669	**	0.349	-0.2029		-0.274
PhD in USA	-0.2254	***	-0.239	-0.2458	***	-0.184	-0.1304		-0.176
PhD in Europe	-0.1544	***	-0.163	-0.1251	**	-0.094	-0.4620	***	-0.624
Cohort before 1960	-0.5388	*	-0.570	0.5689		0.426	?		?
Cohort 1961-1970	-0.2769	**	-0.293	-0.1264		-0.095	0.2319		0.313
Cohort 1971-1980	-0.4725	***	-0.500	-0.0125		-0.001	-0.0320		-0.043
Cohort 1981-1990	-0.2819	***	-0.298	0.0689		0.052	0.2504	***	0.338
Constant	-2.0730	***	-2.195	-2.4561	***	-1.838	-1.5867	***	-2.142

*** Significant at 1%, ** significant at 5%, * significant at 10%.

? Because of the small number of observations, the variable was dropped due to collinearity.

Table II. Continuation

	Social and Humanities			Agricultural Sc and Biotechnology			Engineering		
	Coef.		Marginal Impact	Coef.		Marginal Impact	Coef.		Marginal Impact
Time variant variables			0.457			0.690			0.451
Age	0.2274	***	0.104	0.223	***	0.154	0.243	***	0.109
Age Square	-0.0022	***	-0.001	-0.002	***	-0.001	-0.002	***	-0.001
Level 3	0.6382	***	0.292	-0.039		-0.027	0.267	**	0.120
Level 2	0.6438	***	0.294	0.069		0.047	0.550	***	0.248
Level 1	0.5762	***	0.263	0.179	***	0.123	0.566	***	0.255
Total # of pub	0.0050	***	0.002	0.002	***	0.002	0.002	***	0.001
Total # of researchers	-0.0005	**	0.000	0.000		0.000	-0.001	***	0.000
Budget -1	0.0002		0.000	0.000		0.000	-0.001		0.000
Constant	-5.6422	***	-2.579	-5.267	***	-3.632	-5.348	***	-2.410
Time invariant variables									
Female	0.0625		0.029	-0.099		-0.068	0.122	**	0.055
Public Institution	-0.2756	***	-0.126	-0.259		-0.178	0.215	***	0.097
PhD in USA	0.0363		0.017	-0.210	***	-0.145	-0.368	***	-0.166
PhD in Europe	-0.1026		-0.047	-0.243	***	-0.168	-0.341	***	-0.154
Cohort before 1960	1.0355	***	0.473	?		?	-1.481		-0.667
Cohort 1961-1970	-0.1233		-0.056	-1.409	*	-0.972	-0.055		-0.025
Cohort 1971-1980	-0.7553	***	-0.345	0.047		0.032	0.030		0.014
Cohort 1981-1990	-0.1545	*	-0.071	0.223	***	0.154	0.004		0.002
Constant	-1.6379	***	-0.749	-1.552	***	-1.070	-1.574	***	-0.709

*** Significant at 1%, ** significant at 5%, * significant at 10%.

? Because of the small number of observations, the variable was dropped due to collinearity.

Table II. Continuation

	Physics			Materials Eng.			Biology		
	Coef.		Marginal Impact	Coef.		Marginal Impact	Coef.		Marginal Impact
Time variant variables			1.147			0.761			1.597
Age	0.2219	***	0.255	0.2781	***	0.212	0.2869	***	0.458
Age Square	-0.0021	***	-0.002	-0.0026	***	-0.002	-0.0026	***	-0.004
Level 3	0.8565	***	0.983	-0.2921		-0.222	-0.0027		-0.004
Level 2	0.8602	***	0.987	0.2377		0.181	0.0144		0.023
Level 1	0.7866	***	0.902	0.3359	***	0.256	0.1085	**	0.173
Total # of pub	-0.0002		0.000	0.0015	***	0.001	0.0002		0.000
Total # of researchers	0.0003		0.000	-0.0006	*	0.000	-0.0002		0.000
Budget -1	0.0018	***	0.002	0.0002		0.000	0.0011	***	0.002
Constant	-5.5221	***	-6.335	-6.3067	***	-4.802	-5.6855	***	-9.082
Time invariant variables									
Female	-0.1855	**	-0.213	0.1665		0.127	-0.1305	**	-0.208
Public Institution	0.3337		0.383	-0.2573		-0.196	0.2774		0.443
PhD in USA	-0.0423		-0.049	-0.2824		-0.215	-0.2554	**	-0.408
PhD in Europe	-0.0740		-0.085	-0.3253	***	-0.248	-0.3021	***	-0.483
Cohort before 1960	-0.9603	*	-1.102	-	-	-	0.2685		0.429
Cohort 1961-1970	-0.2614		-0.300	-0.4484		-0.341	-0.1320		-0.211
Cohort 1971-1980	-0.4626	***	-0.531	0.1828		0.139	0.0286		0.046
Cohort 1981-1990	-0.3545	***	-0.407	0.0909		0.069	-0.0374		-0.060
Constant	-1.7583	***	-2.017	-1.1855	***	-0.903	-2.6092	***	-4.168

*** Significant at 1%, ** significant at 5%, * significant at 10%.

? Because of the small number of observations, the variable was dropped due to collinearity.

Table III. Citations per four years

Variables	***Coef.***	***Std. Err.***		***Marginal Impact***
Time variant variables				0.2509
Age	0.1446	0.0107	***	0.0363
Age Square	-0.0014	0.0001	***	-0.0004
Level 3	1.1161	0.0584	***	0.2800
Level 2	0.8348	0.0419	***	0.2094
Level 1	0.5699	0.0287	***	0.1430
Total # of pub	0.0007	0.0001	***	0.0002
Total # of researchers	-0.0002	0.0000	***	0.0000
Budget -1 year in Millions of dollars (2001)	0.0012	0.0001	***	0.0003
Constant	-5.9997	0.2336	***	
Time invariant variables				
Exact Sc.	1.2972	0.0239	***	0.3255
Biology and Chemistry	1.3596	0.0238	***	0.3411
Health Sc.	1.7412	0.0241	***	0.4368
Agriculture and Biotechnology	0.6368	0.0271	***	0.1598
Engineering	0.6097	0.0263	***	0.1530
Female	-0.2287	0.0093	***	-0.0574
Public Institution	0.4092	0.0313	***	0.1027
PhD in USA	-0.0213	0.0122	*	-0.0053
PhD in Europe	-0.2645	0.0112	***	-0.0664
Cohort before 1960	-0.4092	0.1112	***	-0.1027
Cohort 1961-1970	0.4528	0.0291	***	0.1136
Cohort 1971-1980	0.4599	0.0154	***	0.1154
Cohort 1981-1990	0.4745	0.0107	***	0.1191
Constant	-1.2633	0.0374	***	-0.3170

The base is a male researcher in Social and Humanities, in a private institution, who got his PhD in Mexico or other country between 1991 and 2002.
N=5,658, T=1991-1997
*** Significant at 1%, ** significant at 5%, * significant at 10%.