Scientific Productivity and Academic Promotion: A Study on French and Italian Physicists[⊗]

Francesco Lissoni [1,2], Jacques Mairesse [3,4,5,6], Fabio Montobbio [1,7], Michele Pezzoni [1,2,8]

[1] KITes/Cespri-Università Bocconi (Milano, IT); [2] Università di Brescia (IT); [3] CREST-INSEE (FR); [4] Politecnico di Torino and ICER (IT); [5] Maastricht University (NL); [6] NBER (USA); [7] Università dell'Insubria (Varese, IT); [8] Università di Bergamo (IT)

ABSTRACT: The paper examines the determinants of scientific productivity using a panel of 3633 physicists active in 2004-05 in France and Italy. Endogeneity problems are taken care by a two-stage Heckman estimation, with a selection equation that estimates the individual scientist's probability of promotion to her present rank; and a productivity equation, conditional on the scientist's rank. Productivity is measured both in terms of number of articles and journal impact factor, and it is explained by age, gender, entry cohort, and environmental variables. We show that scientists recruited or promoted *en masse* over a short period of time (after prolonged spell with no new recruitments or promotions) are on average less productive than scientists belonging to other entry cohorts, especially in Italy. The selection equation bears results of interest also on the determinants of academic careers.
.

[⊗] Financial support to Jacques Mairesse from the Italian Ministry of Education is gratefully acknowledged (FIRB Project "La ristrutturazione del sistema della ricerca pubblica per il trasferimento tecnologico: governamce, strumenti ed interventi attuativi"). Michele Pezzoni benefitted from support by COST, the European Cooperation Programme in Science and Technology (Action IS0604: Science and Technology Research In a Knowledge-based Economy -STRIKE). Francesco Lissoni and Fabio Montobbio received support from the European Commission (NEST-2006-PATH-Cul, CID - Cultural and Innovation Dynamics, Contract n.: FP6 – 043345). Early drafts of the paper benefitted from comments by three anonymous referees and by participants to several CID workshops and to: the DIME conference on "Knowledge in space and time: economic and policy implications of the Knowledge-based economy" (Strasbourg, April 2008); the Brick-DIME workshop on "The Organisation, Economics and Policy of Academic Research" (Moncalieri, July 2008); the NBER Productivity Seminar (Cambridge MA, March 2009). All mistakes remain ours.

1. Introduction

Countries differ substantially in their scientific productivity and impact (King, 2004). It is often suggested that these large differences depend upon the institutional settings where the scientific research is done (May, 1997). In many European countries, policy makers are implementing various assessment programs in order to increase the effectiveness of public funding of science. Yet, little is known on the determinants of scientists' productivity in the specific European institutional settings. Most of the available studies on scientists' productivity and careers are based upon US data and most of the theoretical insights and policy implications refer to or are based upon the institutional specificities of the US academic system (Long, 1978; Allison and Long, 1990; Levin and Stephan, 1991; Lee and Bozeman, 2005). However, European academic systems differ substantially the US one and among themselves, the most important differences being related to: the degree of autonomy of universities; the size and flexibility of the academic job market; and the relative weight and prestige of universities and public research organizations (PROs) within the national science system.

In this paper we study the determinants of scientific productivity of physicists in French and Italian universities, and the relationships between such determinants and the academic career mechanisms typical of the two countries. In both France and Italy, professors (both junior and senior) are civil servants recruited and promoted through centralized procedures controlled by the State. In addition, universities have very limited financial autonomy and compete for human capital and resources with large PROs (which are especially strong in France). We ask whether in such systems individual scientists' productivity and careers are affected by the same forces that shape them in the US, or by different ones.

In particular we explore on the one side the impact of age, gender, and quality and quantity of research activity on the probability of promotion and, on the other side, the role of age, gender, recruitment policies and patterns of co-authorship on the quality and quantity of individual scientists' publications. Moreover we consider some institutional specificities of the French and Italian systems. In particular we analyse the recruitment processes and we ask whether the frequent observed stops-and-goes, typical of such two centralized systems, have led to the existence of noticeable entry cohort effects on productivity.

We make use of a new panel of 3633 academic physicists in France and Italy, all of them active in academic year 2004-05. The scientists' production is measured by the number of their publications on 363 high impact-factor journals (from the ISI-Web of Knowledge), where the impact factor also serves as a measure of quality. We estimate separate regressions for each position in the academic ranking of the two countries (from assistant to full professor). Sample selection issues implicit in this approach are tackled by means of a traditional Heckman two-step estimation, in which the probit selection equation gives relevant information on the determinants of the probability of promotion from one rank to another. We account for individual heterogeneity by considering the scientist's average yearly productivity before promotion (quantity/quality before promotion).

Our results show the existence of age and gender effects both on productivity and on promotion probability, although with some differences between the two countries, conditional on the academic rank. We also find that a scientist's productivity is positively affected by environmental variables, such as the productivity and affiliation (international vs. national) of her co-authors. Finally, we find that the stop-and-goes typical of the two centralized recruitment systems leave a permanent mark on selected entry cohorts. In particular, the scientists recruited or promoted in 1980 in Italy and in 1984 in France (when massive recruitment waves took place after many "dry" years) are on average less productive than scientists from other entry cohorts.

The remainder of the paper is organized as follows. Section 2 briefly outlines the background literature. In section 3, we discuss the institutional features and relevant historical developments of the French and Italian academic systems in the past 30 years. In section 4 we present our data and the econometric specification. In section 5, we comment the results. Section 6 concludes.

2. Background literature

Interest in the determinants of individual scientists' productivity dates back to the XIX century[1]. From the very start, enquiries on scientific productivity were meant to cast light on a two separate issues: the soundness of eugenic principles proposed by Sir Francis Galton, whose studies on the "hereditary genius" had been largely based on the demographics of "eminent men of science"; and the impact of academic institutional arrangements and incentive schemes on a county's scientific performance, as measured by the number of outstanding scientists from that country (Cattell, 1903; Godin, 2006; Godin 2007)[2]. Both the issues have been debated ever since and are still present in today's studies, although quite often in a disguised manner.

Since the 1960s, sociologists of science have tested whether increasing returns to scientific reputation and productivity (presently referred to as "Matthew effect") may explain Lotka's law better than widely held beliefs on the unequal distribution of intelligence in the population (Merton, 1968; Merton, 1988; David 1994). A typical result in this direction has been obtained by Long and Fox (1995), who find that, other things being equal, graduates from prestigious universities have a higher chance to get their first job at institutions in the same league, with substantial advantages in terms of present and future research productivity. As for individual determinants, gender is the one that has attracted most of the attention, with cognitive and genetic explanations of gender differences (as

[1] It was 1903 when James Cattell published the first systematic data collection on scientific papers per author and provided strong evidence of the existence of large differences across individuals, a result later systematized by Alfred Lotka's well-known "power law"(Lotka, 1926).

[2] A third purpose served by these early enquiries was the study of disciplines, from their birth to consolidation. This is still a very much beaten path, although recent studies add to simple paper counts increasingly sophisticated applications of network analysis (Crane 1972).

opposed to social ones) being still debated (Etzkowitz, Kemelgor and Uzzi, 2000; Fox, 1999; Spelke, 2005).

Understanding the relationship between academic institutions, incentive schemes, and scientists' performance has become increasingly important over the last 30 years or so. This surge of interest can be explained by the policy makers' wish to measure and increase the effectiveness of the public funding of science and by the reaction of leading research universities and scientists, whose attempt to shield themselves from threatened budget cuts has often lead to calls for a higher concentration of resources on the basis of publication (citation)-measured excellence (Graham and Diamond, 1997).

An important line of research has referred to the related issues of age and tenure. In particular, many studies have explored the possibility that individual scientific productivity follows a life cycle: productivity increases when the scientist is young, reaches a peak at/before middle age, and declines afterwards (Levin and Stephan, 1991). At the same time, studies on tenure have tried to clarify whether observable life cycles are due to biological factors or to an incentive problem, such as the reduced "publish-or-perish" pressure on senior scientists with tenured positions. In a rare study on a non-US sample, Turner and Mairesse (2007) find that while promotion has a positive effect on the quantity and quality of publications, the time spent on the same tenured senior position has a negative impact on productivity. They also show that being member of a highly productive laboratory fuels individual scientific productivity. This confirms that, some institutional variables, such as the stratification of universities according to prestige and funding, generate increasing returns in science.

3. Recruitment and Careers of Scientists in France and Italy

Most of the available literature on scientific productivity is based both theoretically and empirically on the US case. This is an important limitation, because the latter is not representative of university systems worldwide[3]. The centrality of universities for public science, the degree of academic job mobility, and the clear stratification of universities according to research intensity are typical characteristics of the US (Ben-David, 1992; Clark, 1993). In countries such as France and Germany, for example, large public research organizations such CNRS (*Centre National de la Recherche Scientifique*) or the Max Plank Institute have been regarded by policy makers as the main pillars of the public research system; for long, Italy also followed this model with CNR (*Consiglio Nazionale delle Ricerche*).

As for job mobility, we observe that US universities select candidates for professorial jobs in total autonomy, with no control from the central (federal) or state governments. Professors of all ranks are university employees who can bargain for their wages and working conditions on an individual basis; in addition, the existence of a proper academic job market allows scientists with a strong publication record to move across universities in search of better paid or better funded research positions (Ehrenberg

[3] Important exceptions are Bonaccorsi and Daraio (2003) Hall et al. (2005) and Mairesse and Turner (2007).

et al., 1990). This is hardly the case in France and Italy. In these centralized systems university staff are considered civil servants, who are employed by the government and selected by commissions of senior peers, who are elected by national members of the relevant discipline or nominated by government. In these countries, there is not a strong competition between universities for the recruitment of the most promising or productive scientists.

At the same time, strict dependence of universities on government funding (which tend to be highly pro-cyclical) and the frequent reforms of the recruitment procedures tend to make career perspectives highly erratic. Italy, for example, has a long history of prolonged periods during which universities do not recruit any new scientist, due to funding shortages or ongoing policy revision, followed by sudden waves of mass recruitment and promotions, often under the political pressure exerted by the large number of scientists seeking a tenured position after many years of temporary contracts.

3.1 Scientific careers in France and Italy

The French academic system has two main positions: "*Maitre de conference*" (MCF; roughly equivalent to the US rank of assistant professor) and "*Professeur*" (PR). In Italy there are three positions: "*Ricercatore universitario*" (RU), "*Professore Associato*" (PA) and "*Professore ordinario*" (PO). In French physics, MCFs amount to around two thirds of the tenured faculty; in Italy, each position account for around one third of the tenured faculty (see figure 1).

[FIGURE 1 ABOUT HERE]

All positions are tenured, and for all of them the teaching and research duties, as well as the wages, are defined by national laws, with limited possibility of local re-negotiation. Whatever their rank, academics are classified by the government according to their discipline. Such disciplines act very much as professional guilds, since their members (and not individual universities or department) control the recruitment process[4]. Before accessing any of the above-mentioned positions (most often RU in Italy and MCF in France) young scientists go through post-doc positions of various length. Universities that wish to recruit a young scientist or promote any member of the academic staff are subject to a series of administrative constraints, which limit their freedom to allocate funds to recruitment or career progress in the absence of governmental approval, and force them to follow a set of complex procedures to place a job call and collect the related applications. In addition, the job calls may be suspended by the government at any time for either

[4] Notice also that the US system is openly stratified according to the research *vs.* teaching intensity of institutions. On the contrary, both the Italian and French laws forbid universities to differentiate openly their mission, and to assign different research vs. teaching loads to the faculty.

financial or regulatory reasons, forcing universities to delay their recruitment or promotion plans.

In Italy and France, these constraints have often clashed against an increasing demand for higher education instruction, whose growth was particulary intense in the 1960s and 70s. Both countries answered to the growing number of students of those years by hiring a large number of young assistants with non-tenured positions and fixed-terms contracts (*Assistants* in France and *Professori incaricati* as well as *contrattisti* and *assegnisti* in Italy). In early 1980s, two reform acts were passed: law 382 in Italy (in 1980) and the Higher Education Act in France (in 1984). Both laws reformed the recruitment process by introducing the ranking system we described above, and by changing the hiring rules. At the same time, a number of *ad hoc* measures were passed along with the new laws, which were meant to allow many *professori incaricati* and *assistants* to obtain tenured positions as *ricercatori* or *professori associati* in Italy, or *maitres de conference* in France. In Italy, the *ad hoc* procedures were merely formal and candidates did not face any selective competition[5] (Moscati, 2001). As a result, each country saw a massive recruitment wave (respectively in 1980 and 1985), which was followed by a prolonged period without any recruitment in Italy, and with a sensible decline in recruitment in France. Figure 2 and 3 illustrates this effect for the field of physics: they report the distribution of scientists who were in active duty in 2004-05, by year of recruitment (for RUs and MCFs) or year of the last promotion (for those in professorial positions). The 1980 and 1985 entry/promotion peaks are clearly visible. It is important to notice that the promotion procedures to the highest academic ranks (PR in France and PO in Italy) were not much affected by these legislative actions.

[FIGURE 2 AND FIGURE 3 ABOUT HERE]

The 1980 and 1984 laws also had lasting effects on the selection procedures. In Italy, it was established that the recruitment of PAs and POs had to occur at a national level, for a number of positions issued every other year by the ministry, and managed by a professors' committee whose members were first chosen by all peers in each discipline, and then further selected by the ministry. As for RUs, these were selected at the university level by a committee of three PAs and POs, all appointed by the ministry. Together with the universities' lack of financial autonomy, this highly centralized recruitment system soon became responsible for a number of difficulties in speeding up the recruitment process by universities throughout the 1980s and early 1990s (see again Figure 2).

In 1998 a new recruitment system was introduced, which was still in place at the time of our study[6]. The system allows each university to offer new positions by launching its own call for applications (*concorso*), and to set up an examination committee. All the

[5] See article 59 Décret n 84-431 du 6 juin 1984 for France and article 50 and 57 DPR 382/1980 for Italy. These laws can be found (in original language) on the following French and Italian websites: http://www.legifrance.gouv.fr/ and http://www.pd.infn.it/infn_ric/GruppiLavoro/Stato_Giuridico/Stato%20Giuridico%20Universitari_DPR382_1980.html

[6] While we revised the paper, the Ministry of Education announced a new change in the recruitment laws, which by June 2009, however, had not yet been entirely disclosed to the Parliament.

committee members, however, must belong to the same discipline for which the position is offered and (with the exception of just one member) not selected by the university, but elected by all the discipline affiliates at the national level. Nominally, the commission has not the task to pick the most suitable candidate for the university that launched the call (on the basis, for example, of the coincidence of the candidate's and the university's research interests), but the best possible candidate in absolute terms, who should be the one with the best publication record (called "*idoneo*", which means *fit-for-the-job*). In principle, if the university does not like this candidate, it can always refuse to nominate him/her and launch a new job call. In practice, most commissions try to steer the selection process towards candidates who they know will be palatable to the university.[7] Once again the introduction of the new law coincided with a new wave of recruitment; although less dramatic than the 1980 one, it is still visible in Figure 2, over the years between 1999 and 2001.

The French recruitment system is also very centralized and discipline-centred[8]. Every year, the central government issues a list of vacancies, by discipline and institution, both for the MCF and PR positions. The applicants need first to get a *qualification* certificate, which is granted by the CNU (*Conseil National des Universités*) – whose members are partly elected and partly designed by the Ministry of Education. Once obtained, the *qualification* which is valid for four years and the qualified candidates are the only ones who can apply for filling the vacant positions. Applications will then be examined at the university level, by disciplinary commissions, composed both of local and non-local members[9].

Both the Italian and the French recruitment systems have undergone severe criticism over the years, which resulted in a succession of reforms that shifted the balance of decision power in recruitment matters back and forth between the national and local level (Musselin, 2005; Moscati, 2001; Chevaillier, 2001; see also footnotes 6 and 8). None of these reforms, however, has gone as far as to grant universities total freedom in recruitment matters, nor it has diminished much the power of disciplines. More importantly, the procedural uncertainty created by this succession of reforms, combined with repeated cuts in the public spending for universities, has made the recruitment process very irregular over time.

Time irregularities in recruitment and career advancement mechanisms determine the age composition of the academic workforce. Table 1 reports the distribution of the latter by rank and date (decade) of birth. As expected, the lower academic ranks are populated by younger scientists in both countries. However, in Italy the top academic rank (PO) is populated by much older scientists than its French equivalent (PR): while in

---

[7] This requires many behind-the–scene manouvres by the university, aimed at steering the election of the committee members in favour of friendly candidates. For a more in-depth discussion of this point, and of its consequences for academic careers in Italy, see Pezzoni et al. (2009)

[8] In August 2007, the Sarkozy presidency introduced a number of reforms aimed at shifting powers from disciplines and the State to university administrations, but their effects cannot yet be appreciated.

[9] "Each committee is made up of from 10 to 20 elected members drawn from the institution's faculty and of members from other institutions or from other disciplines of the same institution. These committees, in which professors and lecturers sit in equal number, are elected for 4 years and meet as often as is necessary to deliberate on staff recruitment and promotion" (Chevaillier, 2001, p. 61).

France 19.59% of the full professors (PRs) was born in the 1960s, in Italy the percentage of POs from the same decade is only 2.57%. Although less striking, such differences in age composition hold also for the lower ranks: only 11.63% of Italian RUs were born in the 1970s, as opposed to 24.97% of French MCFs. This is a very likely consequence of the higher emphasis on seniority-career links in Italy, and of the higher irregularity of the Italian recruitment process.

[TABLE 1 ABOUT HERE]

Finally, an important institutional feature of both the French and the Italian research systems is the role of large PROs such as the CNRS and the CNR, respectively. In France, CNRS has been traditionally regarded as the most important actor of the research system (even more so in physics, with the exception of nuclear physics, where that role was contended by CEA, a special agency for atomic energy). As such, it has often out-competed universities in attracting the best and most motivated young scientists, who may perceive the academic position of MCF as overloaded with teaching and administrative duties (not so much the professor's position, which is often targeted by CNRS *chargé de Recherché* willing to enter the university from the top). As for Italy, the history of CNR is one that starts similarly to that of CNRS (which indeed was taken as a model for its creation, in the late 1930s), and ends up differently. Badly hit by successive budget cuts, the CNR has progressively lost its centrality in the Italian research system, as well as any possibility to offer permanent positions to young scientists. This implies that while French universities have at least one competitor in the academic labour market, Italian universities have none.

3.2 Implications for scientific productivity

The analysis of the recruitment process in the two countries, and of the relative balance of universities and PROs inside the research system, suggests some observations on the factors affecting scientific productivity. In principle, productivity is a key determinant of career advancement in both countries, at least for senior positions. In a related paper we show that this is indeed the case, at least for moves up to PR positions in France, and, in Italy, for moves from the RU to the PA position (less so for moves up to the top PO position; see Pezzoni *et al.*, 2009). As a consequence, other things being equal, we should expect scientists who are currently on higher positions to be more productive than those in lower ones. We should also expect such scientists to exhibit a less pronounced life cycle, that is to be highly productive over a longer time spell, and to incur in diminishing productivity rates at a later age. However, at least for Italy, informal recruitment and promotion practices push in the direction of career by seniority and give considerable advantages to local candidates, no matter their productivity. If these effects were dominant, they could cancel the formal incentives to high productivity.

We also expect to find strong effect for the different years of entry/promotion both in Italy and France, for at least three reasons. First, access to tenured academic positions has increasingly become more difficult over the 1980s and 1990s, so we may presume that scientists who have been recruited recently are more productive than their predecessors. Second, late generations of scientists have more international experience and may be expected to find it easier to publish in international journals, which are better represented, in our data, than French or Italian ones. Last, cohorts of scientists recruited *en masse* over short periods of time (after periods of no recruitment) could be either more productive or less productive than the average. As for RUs, PAs and MCFs entered and/or last promoted in 1980 and 1985 (respectively in Italy and France), we expect them to be less productive, due to the explicit lack of selective pressure for filling those positions in those two years. We do not have any *a priori* on the effect of the 1980 and 1985 entry waves on the productivity of scientists in higher positions (PO and PR), which were much less affected by legislative changes of those years.

We expect productivity to grow over time (calendar years) for at least three reasons. First, both in Italy and in France public research funding has been increasingly distributed on the basis of competitive grants. Second, physics, as any other discipline, has enjoyed decreasing publishing cost, thanks to new procedures and media. Finally, and especially in physics, "big science" and large projects have favoured teamwork, which tend to increase productivity, as measured by publishing: *"..It is especially noteworthy that nobody who worked without collaborators or with only one co-author succeeded in producing more than four papers in the five-year period, whereas everybody with more than twelve collaborators produced fourteen or more papers in the same time.."* (Price and Beaver, 1966, p. 1014).

As for the gender effect, no apparent reason exists to think of peculiarities for France and Italy with respect to the US, so we expect to find evidence of it in both countries. Stratification effects may also exist, which we try to capture with information on the affiliation and productivity of our scientists' co-authors.

4. Data and Methodology

4.1 Data collection and sample

Our database comes from records of all tenured academics from Italian and French universities, active in academic year 2004/2005. Both lists were provided by the Ministries of Education of the two countries, alongside with the disciplinary affiliation of the individuals. The disciplines we selected are reported in Table 2: they cover all fields of physics with the exception of astro-physics and nuclear physics, although some nuclear physicists may belong to the selected fields[10].

[10] Not all the scientists classified in one discipline, in fact, have homogeneous research interests. In addition, it is often the case that some scientists are pushed to join (nominally) the discipline with the highest opportunities of promotion at a given point in time, and then trying to puruse their own interests anyway. Face-to-face interviews to Italian physicists show that some nuclear physicists are classified as *Fisica Sperimentale* (Field FIS/01), which enters our study, instead of being classifies as *Fisisca Nucleare e Subnucleare* (Field FIS/04), which we excluded from our study.

Our data include the scientists' dates of birth and last promotion, but not the date of recruitment, if different from last promotion. For example, for a scientist who in 2004/05 held the position of PA, PO (if Italian) or PR (if French) we know in what year she was promoted to that rank, but we do not know anything about the date of her first recruitment as RU or MCF (on the contrary, for scientists in the position of RU of MCF the dates of recruitment and last promotion coincide, so we know them).

Information on scientific publications was gathered from the ISI web of science, published by Thompson-Darwent. In particular, we selected all international physics journals with a current 5-year average impact factor higher than 0.5, for a total of 363 journals. The selected journals cover uniformly our disciplinary fields and can be considered the main international journals in physics[11]. Cross-discipline generalist journals like Science and Nature are not included[12]. The collected publications span from 1975 to 2005, and they have at least one author whose surname and initial(s) match those of at least one scientist in our lists.

[TABLE 2 AND TABLE 3 ABOUT HERE]

Such a collection strategy, the only one compatible with our data, is inevitably plagued by problems of homonymy, which may lead to over-estimating the productivity of scientists with common surname or to find it impossible to attribute papers with authors whose surname and initial(s) match more than one scientist's surname and name(s) in the ministerial databases. To avoid these problems we dropped from our search all homonyms in the ministerial lists and all scientists with both common surnames and an apparent productivity far too high for being credible[13]. This operation led to dropping about 3% of scientists in the original lists, and left us with 2151 French and 1769 Italian scientists.

We could not find any publication in the selected journals for almost 2.9% (52) of the selected scientists in Italy and 10.9% (235) in France. Two possible explanations for this significant difference across countries are that either French academics write more than Italians on journals not included in the ISI web of science or that it is more common for French academics to dedicate themselves exclusively to teaching. Given the impossibility

[11] The list of the 20 journals with the highest number of publications in our database is reported in the Appendix (Appendix 4). Interviews have confirmed that no top journal of physics is missing from the list.

[12] ISI records report only the surname and initials of authors, which we had to match to surnames-plus-names from ministerial records. In multi-disciplinary journals there is no indication of the field to which the article belongs, so by including such journals we would have run the risk to count several publications outside the realm of physics as authored by scientists in our sample with common surnames and names. At the same time, it is very unlikely that a physicist with a meager publication record in specialized journals will ever make it to Science or Nature and compensate through those journals for her lack of specialized publications. That is, we think we introduced no bias by excluding generalist journals, while we would have introduced one otherwise.

[13] The large part of the scientists are dropped because they have at least one homonym in the ministerial lists. This clearly should not introduce any bias, unless one thinks that the spelling of surname is in some way related to the scientific productivity. Only 4 scientists in France and 2 in Italy are dropped because of the productivity too large to be credible, moreover we carefully checked the diffusion of surnames before dropping.

to test these explanations, we simply dropped these individuals from the following econometric models[14]. Therefore, the final sample is made by 1916 French and 1717 Italian academics (Table 3). They have produced respectively 46684 and 55596 publications for a total of 102280 publications in our database (respectively 44100, 52919, and 97019 up to 2005). All the statistics reported in the following sections of this paper are referred to the scientists not affected by homonymy and zero-productivity problems.

4.2 Reduced form model: estimation problems and solutions

Our data are an unbalanced panel: scientists are observed from the year of first publication to 2005 (i.e. the last year of observation). If a scientist has no publications before the year of promotion (to their present position), she enters the panel in the year of promotion. To evaluate the determinants of scientific productivity we have to address three kinds of problems. First, it is extremely difficult to identify precisely the effect of promotion on productivity because of endogeneity issues. In order to solve this problem, we estimate five different models, one for each position (PO, PA, RU, PR, and MCF), and we account for the productivity only after (conditional to) promotion.

This solution rises however a second problem, that is sample selection (Heckman, 1979; Dubin and Rivers, 1989). Our five regression models are not based on a random sample because we consider the scientists that in 2005 had been promoted to one of the five positions (and not yet to a higher one, if any). One way to correct for this is to adopt a standard Heckman procedure. We use a probit selection equation to estimate the inverse Mills ratios and then we use them in the second step to correct for sample selection (Heckman, 1979). In our framework the probit selection equation gives information on the determinants of the probability of promotion[15]. The estimated inverse Mills ratios[16] are then used to assess the determinants of the individual scientists' productivity.

This model can be estimated either by a two-step Heckman's procedure (Heckman, 1979) or by maximum likelihood. We use the former because the large size of our panel and the large number of explanatory variables create problem of convergence with the latter. However, the two-step procedure is not applicable in two out of five regressions, namely those referred to RUs and MCFs. This is due to the lack of observations needed to estimate the selection equation in the first step. In most cases, in fact, MCFs and RUs do not have a publication record before being promoted. To avoid this issue we simply applied OLS to the productivity equation of MCFs and RUs, without any correction.

---

[14] Following an anonymous referee's request, we re-run all regressions after re-including all scientists with zero publications. Our results do not change substantially, with the exception of the effects of the 1980 and 1985 entry waves on productivity, which appear to be stronger. The regression tables are available on request.

[15] Academics are all promoted in 2005 to the rank we observe in 2005. Precisely, the selection equation is conditional on being promoted in 2005. Therefore, among the individuals observed, are excluded the academics who are not promoted. The selection equation is useful to correct for the selection bias but it gives also information about the determinants that help to be promoted faster.

[16] We have tried also to estimate different Inverse Mill's ratios by pooled probit across individuals *i* and times *t*, but the results do not change much.

The third problem is to control for individual heterogeneity. Individual characteristics such as ability and effort are not directly observable, but have a crucial impact on promotion and on productivity after promotion. We account for individual heterogeneity by considering the scientist's average yearly productivity before promotion.

4.3 The determinants of selection and scientific productivity

Table 4 lists all the variables we consider for our econometric exercise. The exercise requires estimating two equations: a selection equation in which the dependent variable is promotion of the scientist to the present position (PO, PA, RU, PR, or MCF), with past scientific productivity as the key explanatory variable; and a productivity equation in which the dependent variable is productivity, for all scientists in the same position. Each estimation exercise of the productivity equation is run twice, for two different dependent variables that measure respectively the quantity and quality of the scientific productivity of each individual scientist:

- *Quantity:* log(1 + number of articles in year *t*)
- *Quality:* log(1 + average 5-year impact factor of journals with articles in year *t*)

We do not correct the number of articles, and the related quality measure, by the number of co-authors of each article; that is, we do not try to capture the individual scientist's contribution to the article. We have two reasons for doing so. First, almost all articles in physics are co-authored, often by quite a large number of scientists. Papers written in isolation or by a few authors may be either theoretical ones or, if applied, they may reflect the isolation of their authors from the rest of the scientific community, rather than a larger individual contribution. Introducing arbitrary corrections by the number of authors would have meant giving greater weight to theoretical papers or, possibly, to papers by peripheral authors. Instead, we control for each scientist's number of co-authors before t, which we expect to be correlated to the number of papers and authors at time t[17].

The key explanatory variables of both promotion and productivity are related to the scientists' age, the historical conditions of the academic labour market at the time of their entry or promotion, and time trends in productivity[18]. Additionally, promotion is explained by productivity, which enters the regression through two distinct variables (*Quantity flow* and *Quality flow*) which consists respectively of the average number of papers published by the scientist in between *t-4* and *t-2*, and the related quality measure[19]. Finally, promotion is also a function of the *academics per year*, which is the number of individuals who achieved promotion in year *t*.

[17] We have also produced a series of ancillary regressions with author-weighted measures of quantity and quality, whose results do not differ much from those reported in this paper, and are available on request.

[18] It is common knowledge that bibliometric measures of productivity exhibit an increasing trend over time. Our data do not escape this regularity, both for quantity and quality.

[19] Publications in *t-1* could not be considered exogenous to promotion in *t*, because the recruitment procedure (*concorso* in Italy or *concour* in France) can take many months, up to one year or so. Therefore, some publications appear after a candidate has filed for the job and are not considered by the commissions are relevant for the appointment.

Introducing age, cohort, and historical time in the same equation, whether the selection (promotion) or the productivity one, creates a identification problem discussed by Hall *et al.* (2005). In particular, it is impossible to take simultaneously into account age, time and birth cohorts, one variable being a linear combination of the other two (age = year – birth cohort). In line with Hall *et al.* (2005) we attack this problem by using a semi-parametric model in which these effects enter linearly, and are represented by variables the least affected by identification problems.

First, we do not consider *birth* cohorts, but only *entry* cohorts. That is, our *Cohort* dummies refer to the scientists' years of entry in the academic workforce, namely the first year in which the individual scientist published an article or, if the scientist does not have publications before last promotion, the date of last promotion. This is similar to the approach used by Levin and Stephan (1991) which define as cohort the year in which a scientist received her PhD. It also captures more precisely than the year of birth the influence of changes in the importance of research fields or legislation on a cohort's productivity. The entry cohort dummies are interacted with a full set of *Period* dummies that refer to calendar years; in other words, our semi-parametric model includes a dummy variable for each cohort-period combination (*Cohort* x *Period*). In addition, we estimate the impact of being hired/promoted in 1980 in Italy and 1985 in France. We expect scientists belonging to these entry waves to display, other things being equal, lower productivity levels than scientists from all other cohorts, due to lack of selective pressure. Notice that the share of scientists recruited as RUs, PAs, and POs in 1980 in Italy and as MCFs and PRs in 1985 in France account for respectively 34%, 52%, 19%, 20% and 13% of the observations in the productivity regression samples (see Table 6 and 7), and therefore they represent a substantial share of the academics active in 2005[20]. In both the selections and the productivity equation we identify these scientists with the *wave 1980* and *wave 1985* dummy variables. Besides, in the selection equation, we control for the ease of entry into academia by counting, in each year *t*, the number of scientists promoted in that year and still active in the same rank in 2005.

As for the scientist's age, we address the above-mentioned identification by using age groups; more precisely, we introduce in both the selection and productivity equations five *Age* dummy variables, representing ten-years age-intervals of our scientists (see Table 4)[21]. Our expectation is to find a negative impact of scientist's age on productivity (Levin and Stephan, 1991; Hall *et al.*, 2005). On the contrary, when considered as a determinant of promotion (selection equation), *age* is expected to have a positive impact: the older the scientists, the higher her chances to be promoted (as in Long, 1993).

[TABLE 4 ABOUT HERE]

---

[20] It's worthwhile remarking that we do not have the full career profile of the scientists and therefore we do not know how many PRs, POs and PAs - that were promoted after 1980 in Italy and after 1984 in France – were previously recruited or promoted in the *wave 1980* and in the *wave 1985*.

[21] We have also tried different specifications using directly the age variable and age squared and the results do not change substantially.

In both the equations for selection and productivity we test for gender effects (*gender* dummy equal 1 if the physicist is a woman). In our sample, women represent respectively the 23% and 27% of the French MCFs and Italian RUs. At the same time they are the 7% and 6% of the French and Italian full professor. The literature typically indicates that women publish less than men (Levin and Stephan, 1998). However this result often depends on cross-sectional data that cannot control for individual heterogeneity.

Moreover, in both the selection and productivity equations, we control for the scientist's specific *research fields* through a set of dummies reflecting ministerial classifications of disciplines *(Field 28* and *Field 30* for France; *Field FIS/01*, *Field FIS/02*, and *Field FIS/03* for Italy). We expect the probability to be recruited or promoted to be field-specific for two reasons. First, resources to hire new employees are not distributed homogeneously among all the disciplines. Second, some disciplines could be more prolific in terms of new discoveries and research paths, thus attracting more junior scientists. As for productivity, this can also be affected by the scientist's specific field of research, due to differences in the resources needed to produce a paper.

In order to characterize the work environment in the productivity equations, we also control for the relationship of our scientists with other members of the scientific community. First, we consider the past productivity of each scientist's co-authors in year *t (Co-author's quantity* and *quality)*. The co-authors we consider are only those who also belong to our sample (that is, we ignore all co-authors who are neither French nor Italian physicsts, for whose publication records we have no data). Their overall *Quantity* and *Quality* is calculated over the interval *[t-3,t-1]* and divided by the number of years in the interval. We expect both variables to bear positive effects on the scientist's productivity, however measured, to the extent that working contacts with other productive individuals provide access to knowledge and information[22].

Second, we consider the size and geographical reach of the projects in which our scientists are involved. In order to do so we produce a four dummy variables that summarize information on the number and addresses of each scientist's co-authors, as reported by the ISI publication records (in this case we consider all co-authors, not only the French and Italian scientists)[23]. The first variable is the *Large Project* dummy, which takes value one for all scientists who have at least one article with 30 or more authors over [*t-3*,*t-1*][24]. All of these large project are international, that is for all papers with more than

[22] In order to avoid endogeneity problems, we do not consider the co-authors' publication to which the observed scientist also contributed.

[23] The information on affiliation and related addresses provided by ISI is not precise. For each publication, in fact, the authors' names and affiliations are reported in separate fields, with no one-to-one matching between the two. So there is no way to know how many authors of a publication come from a university or research institute among those listed; but only that at least one authors comes for sure from such university or institute.

[24] Share of articles with 30 or more coauthors, considered *Large Project* outputs, is higher in Italy (8.57%) then in France (5.95%). Appendix 1 shows the percentage of large projects in the two countries and the information available about the affiliations (*addresses*) form the ISI database. Moreover, Appendix 2 reports the average number of authors (per article)

30 co-authors ISI reports several non-Italian and non-French addresses. Alternatively, in case the scientist is not involved in any large project, we move on to check whether at least some international projects, although of a smaller scale. To this end the *Small_Project_with_Foreign_co-author* dummy takes value one if the scientist has no publications with 30 or more authors over [*t-3*,*t-1*], but she has at least one publication with foreign co-authors[25]. The case of scientists who have only publications over [*t-3*,*t-1*]with less than 30 authors, none of them with a foreign affiliation, a third dummy is captured by a third dummy variable, named *Small_ Project_with-only _national_co-authors*)[26] . Since in many cases the information on affiliation and addresses reported by ISI is incomplete, we also control for the possibility that none of a scientist's articles have information on its authors' affiliation or addresses (*Small Project with co-authors of unknown affiliations* dummy). Scientists with no publications over [*t-3*,*t-1*] constitute the reference case.

In order to capture unobserved heterogeneity of professors' skills in the productivity equations, we control for the scientist's productivity before promotion, that is the yearly average quantity and quality since entry year *(Quantity before promotion; Quantity before promotion*).

Finally, we control for problems arising from data design. First, in both the productivity equation we deal with the left truncation problem due to the unavailability of information on publications before 1975. To this end, a dummy variable *Promotion before 1975* is inserted, which takes value one for scientists who reached the present position before the first year of observation. As for the selection (promotion) equation, we simply do not consider observations affected by this problem. Second, we deal the large presence of zero values in all measures of quantity and quality (when used as explanatory variables) by inserting a number of dummies which take value 1 in case of absence of any publication[27].

Table 5,6 and 7 report summary statistics for all the regressors, by rank.

[TABLE 5, TABLE 6 AND TABLE 7 ABOUT HERE]

---

according to a list of 22 nationalities and PROs. Both in Italy and France US, German and English scientists are quite frequent co-authors in *Large Projects* publication class.

[25] Appendix 3 shows the average number of authors belonging to 22 nationalities and PROs per article, in small project publications with at least one foreign author.

[26] The dummy has to be considered as an approximation due to the fact that not all the addresses of the authors are always reported on the articles.

[27] In particular, in the selection equation we insert a *Zero flow* dummy, which takes value one when the scientist has no publications over the time interval [t-3, t-1]. The rationale for this correction is as follows. Productivity may affect the probability of promotion in two ways: first, only active researchers (who have at least one publication over the considered time interval) are likely to be considered for recruitment or career advancements; second, any extra publication (or increase in quality) has a positive marginal effect on the promotion probability. Although we have excluded from our sample all scientists with no publications after 1975, many are left who have dry spells longer than three years, for which the two effects have to be distinguished. Along similar lines, we build two dummy variables that enter the productivity equation (*Zero in preceding year*s and *Zero after promotion*). In this case, the two dummies improve upon all explanatory variables of present productivity based upon past level of quantity and quality of publications.

5. Results and discussion

In what follows we comment separately the results from the estimation of the selection (promotion) equation and of the productivity equation. Notice that the selection equation could be estimated only for Italian PAs and POs, and for French PRs. In the case of RUs and MCFs we do not see enough history of the individual to explain the promotion with a *probit* model, as necessary with the Heckman procedure. For example, many RUs and MCFs have very few or no publications before being promoted to their present rank; and the latter may have been achieved mainly thanks to the candidates' graduate work or educational attainments, which we also do not observe.

5.1 Selection equation

Results from the selection equation for Italy and France are reported in Table 8. Cohort-period interactions are included in the estimation, but not displayed. As expected, the promotion probability of both Italian and French scientists increases with age, which confirms the role of seniority in both academic systems. Note however that for full professorship in both countries (PO in Italy and PR in France), the estimated coefficient for *Age group 5* (that is, for academics who are more than 60 years old) is smaller than the coefficient for *Age group 4*. This suggests a non-monotonic effect of seniority on promotion probability

[TABLE 8 ABOUT HERE]

The positive sign and absolute size of the coefficients for *wave 1980* and *wave 1985*, confirm that in those two years a scientist's chances of being promoted were much higher than at any other time in recent history. This result holds despite we also controlled for the (positive) effect of *Academics per year* on the probability of promotion. For example, the estimated marginal effect for *Academics per year* in the PO equation suggests that one more PO promoted in year *t* gives any (Italian) candidate to the same position 0.18% more chances of success (0.12% for PA and 0.48% for PR). On top of this, being a candidate to the same position in 1980 meant to have 43% more probability of success (43% also for PA and 27%, in 1985, for PR). Overall, in 1980, 62 individuals reached the PO position, for a combined increase in the probability of promotion of around 54%, other things being equal (43% from the *wave 1980* dummy, plus 62x0.18% from *academics per year*).

The *gender effect* is strongly negative in both countries and confirms that, other things being equal, women in physics face more difficulties than men to be promoted. Field effects on promotion are also significant, with some relevant differences across ranks. In Italy, promotion to PO appears to be easier in *Fisica teorica (Field FIS/02)* rather than *Fisica della material (Field FIS/03,* the reference dummy*)*. On the other hand the chances of being promoted to a PA position are higher in *Fisica sperimentale (Field Fis/01)*. In France, promotion to PR are more likely in *Milieux dilues et optique (Field 30)*, compared to *Milieux denses et materiaux (Field 28)*.

As for productivity, its effect on the probability of promotion is captured by the two variables *quality flow* and *quantity flow*, whose estimated coefficients differ across countries. In Italy, promotion seems to be affected only by quantity, the impact of quality being negligible for promotions to both PA and PO positions[28]. In France both quality and quantity flow affect positively and significantly the probability of promotion to PR.

5.2 Productivity equation

Results from the productivity equations for Italy and France are reported in Tables 9 an 10. In Table 9 the dependent variable is *quantity* while Table 10 refers to *quality*. As expected, the age of academics has a negative impact on both the quantity and quality of articles published. In all equations, this is captured by the age groups coefficient, which are all negative and significant (with the exception of *Age group 1*) and increasingly larger when moving from *Age group 3* to *Age group 5*[29]. Only POs' quality does not conform to this general trend.

[TABLE 9 AND TABLE 10 ABOUT HERE]

Elaborations on the estimated coefficients in table 9 tell us that over-60 Italian POs, produces around 1.73 articles less than their colleagues (also POs) in their thirties, per year; a similar comparison suggests that over-60 French PRs produce 2.04 articles less per year[30]. As for quality (table 10), French PRs seem the most affected by age, with over-60s exhibiting a loss of average impact factor per year equal to 2.97. Bearing in mind that we observe the productivity conditional to the scientist's ranking 2005, we can argue that our estimates on age effects capture two different phenomena: *i)*. the life cycle effect, as discussed in section 4.3; *ii)* a self-selection effect, by which less productive scientists get stuck at lower academic ranks. The self-selection effect is best understood by going back to table 8 and related comments in section 5.1, which show that scientists with lower

---

[28] Variables quality and quantity flow are highly correlated

[29] In a separate regression exercise we have tested the effect of age on productivity of Italian Pos, in the absence of controls for participation to large and/or international projects. In this case, the impact of age on productivity (quantity) appears to be negligible. One interpretation is that POs suffers of a productivity loss at the individual level, which they compensate by participating to large international projects.

[30] Throughout section 5 we express the impact of the covariates on quantity (the number of articles) or quality, (average impact factor) by means of marginal effects, which we compute on the basis of the estimated coefficients in tables 9 and 10. The marginal effect of any dummy covariates is given by: $\frac{d(1+y)}{d(1+x)} = \frac{dLog\,(1+y)}{dx} * \overline{(1+y)}$ where $\overline{(1+y)}$ is the average value of the dependent variable and $\frac{dLog\,(1+y)}{dx}$ is the coefficient in the regression tables. As for the marginal effects of a continuous or discrete covariate, this is given by: $\frac{d(1+y)}{d(1+x)} = \frac{dLog\,(1+y)}{dLog\,(1+x)} \overline{\left(\frac{1+y}{1+x}\right)}$ where $\overline{\left(\frac{1+y}{1+x}\right)}$ is the average value of the ratio between the dependent variable and the covariate and $\frac{dLog\,(1+y)}{dLog\,(1+x)}$ is the coefficient (elasticity) estimated in the regressions.

productivity are less likely to be promoted; we now see, from tables 9 and 10, that such scientists also tend to be the least productive among those in the same position.

Gender impacts differently in Italy and in France, and across academic ranks. With reference to quantity estimates in table 9, we observe that Italy exhibit a negative gender effect only for RUs: Italian *ricercatrici* publish 0.45 papers less than their male peers, per year; on the contrary, women in PA or PO position are not less productive than their male equivalents. As for France, we observe a negative gender effect on quantity both for PRs and for MCFs, where women produce respectively 0.46 and 0.17 papers less than men, per year. Results for quantity are similar, with women's gap in average impact factor per year as follows: -0.94 for PRs, -0.4 for MCFs, and -0.61 for RUs. In the absence of additional data, we can can provide intuitive explanations for such results. In particular, we notice that, for Italy, the existence of a gender effect only for the most junior position (RU) is consistent with the possibility that gender matters more at the early career stages, when familiar engagements may also be demanding. Thus, a self-selection process may take place, by which only the best or most motivated researchers try to access the higher ranks, at which the gender effect appears to be less remarkable or even not significant. If this was the correct interpretation, we need then to explain why we observe a gender effect also at the senior level of French PRs. In this respect we note that in France 10% of PRs are women (as opposed to 6% of Italian POs). Moreover going back to the selection equation (table 8), we notice that the negative gender effect on promotion is smaller for PRs relatively to POs. One possibility is therefore that French women face less difficulties in the promotion to full professorship (thanks, for example, to the first centralized step managed by CNU), which may explain why we observe the negative gender to persist at the top academic rank.

The estimated coefficient of the *wave 1980* and *wave 1985* dummies is negative and significant in the quantity model, for all positions (with the exception of French PRs; see table 9), while it is still negative, but never significant in the quality model (with the exception of Italian RUs; see table 10). As for quantity, POs promoted to such rank in 1980 publish 0.42 articles less per year than other POs. Similarly, both PAs and RUs last promoted in 1980 publish 0.6 articles less per year than colleagues last promoted in more recent years. The effect for France is still noticeable, but less pronounced: MCFs who were recruited in 1985 (and never promoted to PR) produce 0.06 articles less than other MCFs, per year. Bearing in mind that we do not observe how many scientists, among those who entered the academic system as RUs or MCFs in 1980 and 1985, have been subsequently promoted to higher positions (as opposed to having retired or left Italy), the interpretation of our results may be as follows: the 1980 and 1985 big recruitment waves in the two countries filled up the RU, PA, and MCF positions with less productive scientists, who did not make much further progress in their careers, and are now responsible for the negative signs we observe in table 9 and (for RUs) in table 10[31].

[31] We can estimate the total scientific productivity loss due to the entry waves in terms of number of papers published by scientists last promoted during the entry waves, compared to the estimated number of papers published by scientists from other entry cohorts, other things being equal. This exercise suggests that the 62 Italian POs promoted in 1980 have published about 651 articles less than what expected by similar POs from other cohorts, over the same years of activity.

Some descriptive statistics are useful to reinforce the interpretation: no less than 10.5% (62/588) of POs, 26.8% (166/618) of PAs and 14.28% (73/511) of RUs active in Italy in 2005 were recruited such position in 1980. In France, only 7.8% (55/705) of PRs and 9.9% (121/1212) of MCFs reached their present position in 1985. In other words, the 1980 wave in Italy was bigger and has produced more enduring effect than its French equivalent[32]. As for Italian RUs, this effect is also visible for the quality-based measure of productivity.

As explained in the previous section, we control for the individuals' unobservable heterogeneity by means of lagged measures of productivity (*Quantity before promotion* and *Quality before promotion*). As expected, all such controls bear a positive and significant sign. The only exception is the negative impact of *Quality before promotion* on POs' quality-adjusted productivity (table 10).

Productive co-authors impact positively on individual scientists' productivity. For quantity measures of productivity, estimates in table 9 suggest that Italian POs and French PRs and MCFs, produce 0.1 to 0.17 extra articles per year for every additional article produced by their co-authors over the three preceding years. The marginal effect of co-authors' quality on the scientist's quality-adjusted output per year positive for alla ranks and countries, and it ranges between 0.74 for French MCFs and 1.5 for POs (elaborations on estimated coefficients from table 10).

Involvement in large and/or international projects affects strongly a scientist's yearly productivity. This is especially true for Italian POs, where the dummy *Large project* produces a marginal effect on quantity of more than 5.5 articles per year. Similar results hold also for PAs and RUs grows strongly their productivity (4.52 and 4.05 more articles per year; the reference group is given by scientists with no publications in the three years before *t*). The effect of *Large projec*t for France, albeit high if compared to the effect of others covariates, is weaker than in Italy, the marginal effects for PRs and MCFs being respectively equal to 3.55 and 2.42. Even quality is affected by participation to large international projects, with increases in the average impact factor per year comprised between 3.56 (for MCFs) and 8.5 (for POs). Scientists who are not involved in large projects, but had at least one foreign co-author in the three years before *t*, also have larger quantity and quality scores than the reference group (but lower than their peers participating to large projects). Italian POs with dummy *Small_project_with_foreign_co-authors* equal to one, for example, publish 3.25 more papers per year than the refrence group; those who participate only to On the other hand, POs with participation to *Small_project_with_only_national_co-authors* publish 1.84 papers more than the reference group. Similar results hold for all position in both countries.

Controls for field effects also prove signficant. Italian POs belonging to *Fisica Sperimentale (Field Fis/01)* and *Fisica Teorica, Modelli E Metodi Matematici (Field Fis/02)*

Similar calculations suggest a production gap of 1740 articles by PAs and of 1095 articles by RUs from the 1980 wave, for a total scientific loss of around 3486 articles.

[32] Notice that our calculations may under-estimate the effects of the 1980 and 1985 entry waves, since they are based on a sample which include only academics who produced at least one article since 1975. A large share of the French and Italian unproductive scientists comes precisely from the 19890 and 1985 waves [respectively 17% (40/235) and 36.5% (19/52)], so that their inclusion would increase further the size of the estimated coefficient for the two wave dummies.

publish respectively 0.9 and 1 articles less than colleagues from *Fisica Della Materia (Field Fis/03)*, per year. The quality of the publications is also lower. In France MCFs belonging to *Milieux dilués et optique (Field 30)* publish more and higher quality articles if compared to *Milieux denses et matériaux (Field 28).*

Table 9 and 10 show also the estimated inverse Mill's ratio produced by the Heckman procedure to correct for the problem of selection bias. Its negative sign means that the probability of not being promoted affects negatively the individual scientific productivity. The coefficient is significant for both PRs, POs and PAs in the quantity regressions, and for PRs only in the quality regressions. Lack of significance for POs and PAs in the quality regression means that the two equations, selection and productivity, could have been estimated separately, without any kind of correction for endogeneity of the academic rank.

6. Conclusions

In this paper we have investigated the determinants of scientific productivity for a large number of physicists active in French and Italian universities in 2005, conditional on their academic position in the same year [*Ricercatore* (RU), *Professore Associato* (PA) and *Professore Ordinario* (PO) for Italy; *Maitre de conference* (MCF) and *Professeur* (PR) for France]. In order to do so, we have applied a two-stage Heckman estimation procedure. First, we have estimated a selection equation, in order to calculate a scientist's probability of being promoted to a specific position; second, we have estimated a productivity equation, with either the quantity or the quality of the scientist's publications per year as the dependent variables, conditional on promotion to the scientist's position in 2005.

In the selection equations we find that Italian or French scientist's chances of being promoted grow with age, although not monotonically (promotion chances decline for scientists over 60). Women, other things being equal, have a significantly lower probability of being promoted both in Italy and France. In Italy, promotion is influenced only by the quantity of a scientist's publications while in France both quality and quantity play a statistically significant role. In both academic systems we observe big recruitments waves, in 1980 and 1985, which affect significantly the probability of promotion.

In the productivity equations we show that the age of academics has a negative impact on the quantity and quality of articles published by French PRs, but not on the quality of Italian POs' publications. For the other academic ranks the age effect is mixed with the other effect of not being promoted, due to the database design. French MCFs and Italina PAs and RUs are characterized by a strong negative impact of age.

As for gender effects, Italian, women at the early stage of their careers (for example RUs) are penalized in their publication activity. However, if Italian women manage to be promoted to higher ranks, then they publish as much as their male colleagues. On the contrary, in France we observe a negative gender effects across all ranks. We find evidence that the work environment is very important for individual scientists' productivity. In particular being involved in large projects, or at least having an international collaboration has a strong and positive effect on quantity and quality of published articles

Finally we show that the big recruitment and promotion waves of 1980 and 1985 had negative and lasting effects on the average scientific productivity of the two countries, especially for Italy and for quantity-based measures of productivity. Such big waves came after a prolonged spell with no recruitment of new scientists and few promotions, and were the result of policy provisions aimed at providing many scientists with precarious and temporary assignments with a permanent positions. Although never equalled in terms of size, such provisions are still typical of the French and Italian systems, due to the tight control over recruitment and promotion exercised by the government, and the little autonomy of universities. The 1980 and 1985 allowed many low productive scientists to obtain tenure as RUs or PAs in Italy, or MCFs in France. Many of these scientists appear not to have progressed much in their careers, and to have persistently scored as the least productive of scientists in their position.

Our results on age, gender and the importance of promotion probability confirm many of the results found in the literature, as surveyed in section 2. Findings on the enduring negative effects of *en masse* recruitment, on the contrary, are more specific of France and, most notably, Italy. They suggest that tight governmental control over academic careers, such as when governments can prevent universities from recruiting or promoting scientists for a long time, and then intervene with sudden and massive waves of job creation, can create persistent damage to the national academic system.

References

Allison, P.D. and Long, J.S. (1990), 'Departmental effects on scientific productivity', *American Sociological Review*, **55**(4), 469–478.

Ben-David, J. (1992), *Centers of Learning: Britain, France, Germany, United States*. Transaction Publishers.

Bonaccorsi, A. and Daraio, C. (2003), 'Age effects in scientific productivity', *Scientometrics*, **58**(1), 49–90.

Cattell, J.M. (1903), 'Statistics of American Psychologists', *American Journal of Psychology*, **14**, 310–328.

Chevaillier, T. (2001), 'French academics: Between the professions and the civil service', *Higher Education*, **41**(1), 49–75.

Clark, B.R. (1977), *Academic Power in Italy: Bureaucracy and Oligarchy in a National University System*. University of Chicago Press.

Clark, B.R. (1993), *The Research Foundations of Graduate Education: Germany, Britain, France, United States, Japan*. University of California Press

Crane, D. (1972), *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*. University of Chicago Press.

David, P.A. (1994), 'Positive Feedbacks and Research Productivity in Science: Reopening Another Black Box', *The Economics of Technology*

Dubin, J.A. and Rivers, D. (1989), 'Selection Bias in Linear Regression, Logit and Probit Models', *Sociological Methods and Research*, **18**(2), 360–90.

Ehrenberg, R.G., Kasper, H. and Rees, D.I. (1990), 'Faculty Turnover at American Colleges and Universities: Analysis of AAUP Data', *Working Paper No. 3239. Available at SSRN*

Etzkowitz, H., Kemelgor, C. and Uzzi, B. (2000), *Athena Unbound: The Advancement of Women in Science and Technology*. Cambridge University Press.

Fox, M.F. (1999), 'Gender, Hierarchy, and Science', *Handbook of the Sociology of Gender*, 441–457.

Godin, B. (2006), 'On the origins of bibliometrics', *Scientometrics*, **68**(1), 109–133.

Godin, B. (2007), ' From Eugenics to Scientometrics: Galton, Cattell, and Men of Science', *Social Studies of Science*, **37**(5), 691.

Graham, H.D. and Diamond, N. (1997), *The Rise of American Research Universities: Elites and Challengers in the Postwar Era*. Johns Hopkins University Press.

Hall, B.H., Mairesse, J. and Turner, L. (2005), 'Identifying Age, Cohort and Period Effects in Scientific Research Productivity: Discussion and Illustration Using Simulated and Actual Data on French Physicists', *NBER Working Paper N° W11739.*

Heckman, J. (1979), ' Sample Selection Bias as a Specification Error', *Econometrica*, **47**(1), 153–161.

King, DA. (2004); 'The Scientific Impact of Nations', *Nature*, **430**, 311-316

Lee, S. and Bozeman, B. (2005), 'The Impact of Research Collaboration on Scientific Productivity', *Social Studies of Science*, **35**(5), 673.

Levin, S.G. and Stephan, P.E. (1991), 'Research productivity over the life cycle: evidence for academic scientists', *American Economic Review*, **81**(1), 114–132.

Levin, S.G. and Stephan, P.E. (1998), 'Gender Differences in the Rewards to Publishing in Academe: Science in the 1970s', *Sex Roles*, **38**(11), 1049–1064.

Long, J.S. (1978), 'Productivity and academic position in the scientific career, '*American Sociological Review*, **43**(6), 889–908.

Long, J.S. (1993), 'Rank Advancement in Academic Careers: Sex Differences and the Effects of Productivity', *American Sociological Review*, **58**(5), 703–22.

Long, J.S. and Fox, M.F. (1995), 'Scientific Careers: Universalism and Particularism', *Annual Reviews in Sociology*, **21**(1), 45–71.

Lotka, A.J. (1926), 'The frequency distribution of scientific productivity', *Journal of the Washington Academy of Sciences*, **16**(12), 317–323.

Mairesse, J. and Turner, L. (2007), 'Individual Productivity Differences in Scientific Research: An Econometric Exploration of Publications by French Physicists', mimeo.

May, RMM. (1997), 'The Scientific Wealth of Nations', *Science*, **275**, 793-796.

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

Merton, R.K. (1988), 'The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property' *Isis*, **79**(4), 606.

Moscati, R. (2001), 'Italian university professors in transition', *Higher Education*, **41**(1), 103–129.

Musselin, C. (2005), *Le marché des universitaires: France, Allemagne, États-Unis*. Presses de Sciences Po.

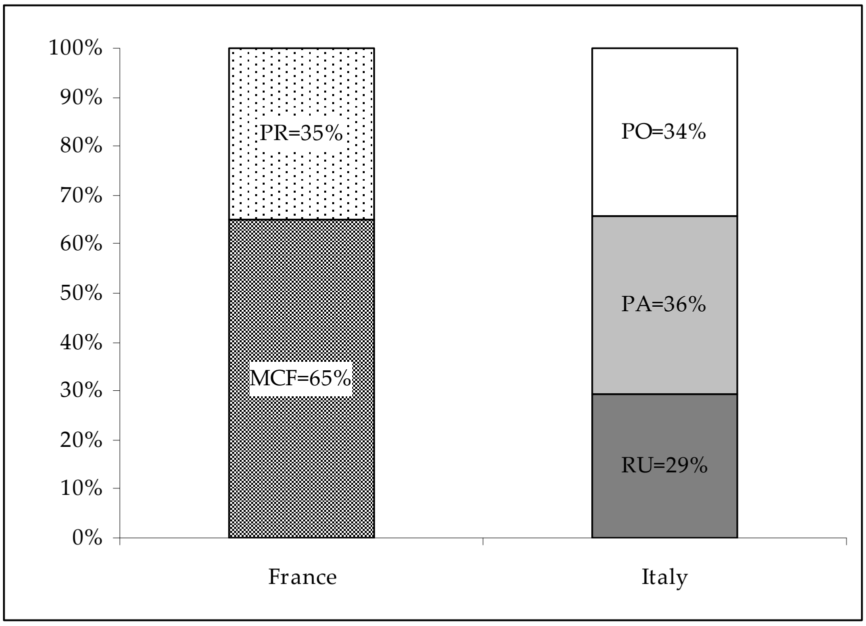
Pezzoni, M., Sterzi, V. and Lissoni, F. (2009), 'Career progress in centralized academic systems: an analysis of French and Italian physicists', mimeo

Price, D.J. and Beaver, D.D. (1966), 'Collaboration in an invisible college' , *American Psychologist*, **1966**, 1011–8.

Spelke, E.S. (2005), 'Sex Differences in Intrinsic Aptitude for Mathematics and Science?: A Critical Review', *American Psychologist*, **60**(9), 950.

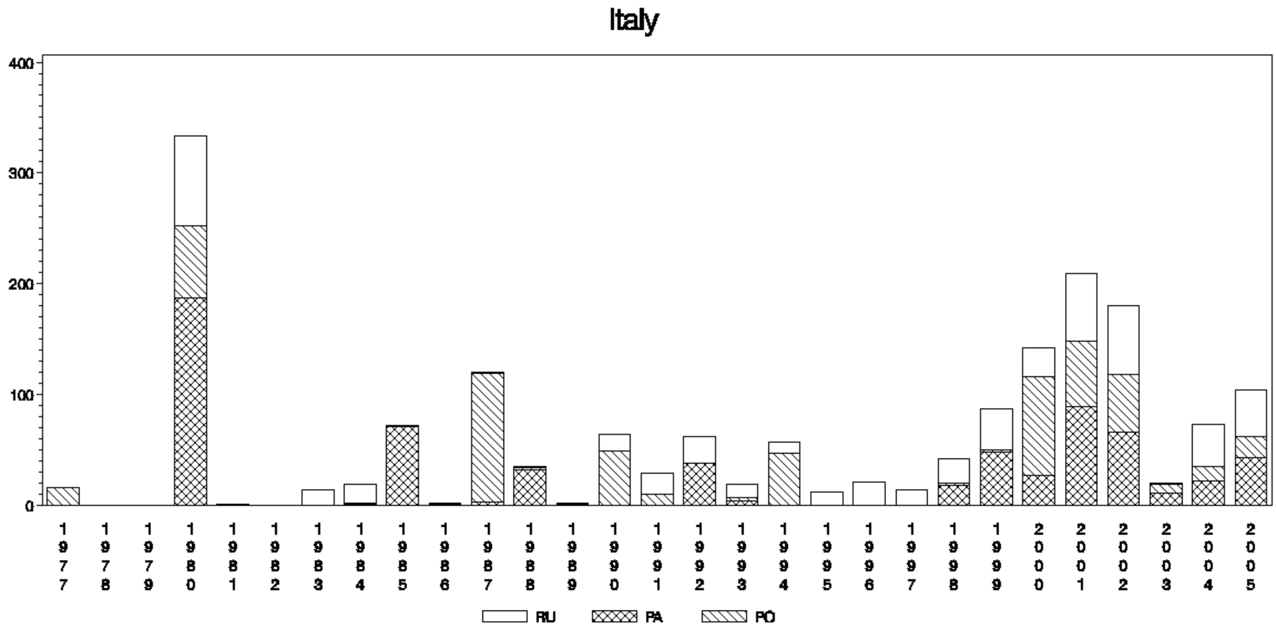
Figures and Tables

Figure 1: Rank distribution of physicists Rank in 2004-05; France (MCF, PR) and Italy (RU, PA, PO)



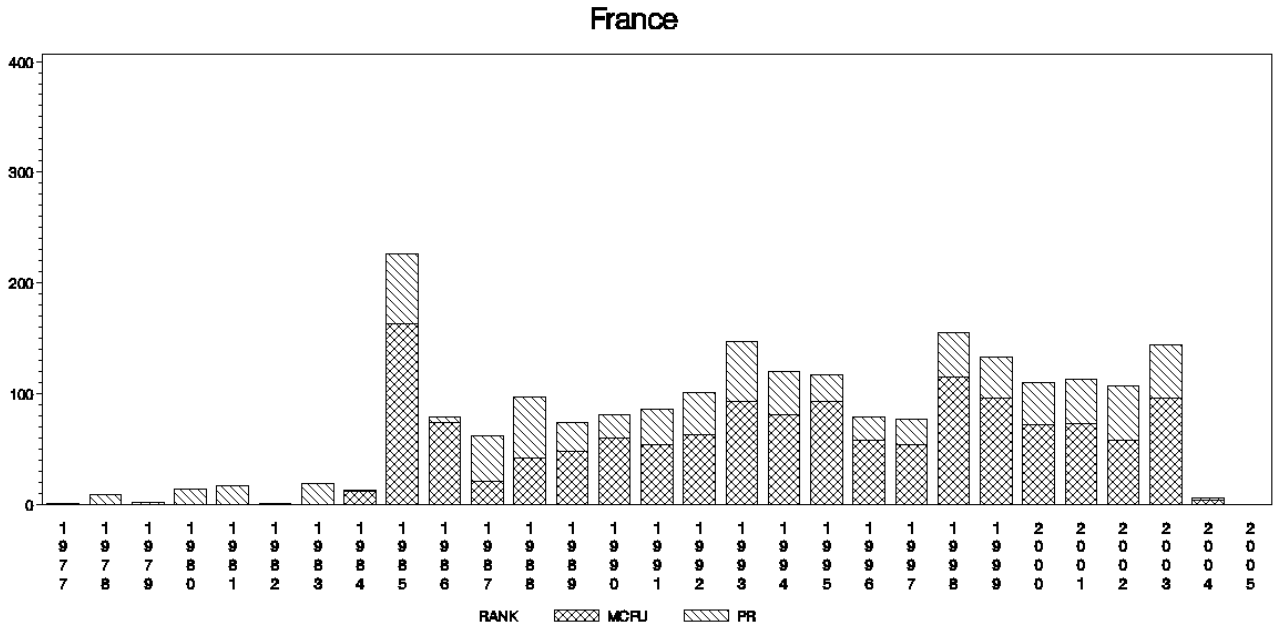
Source: own elaboration on ministerial records

Figure 2. Italian physicists, active in 2004-05; distribution by year of recruitment (RUs) or year of last promotion (PAs; POs)



Source: own elaboration on ministerial records

Figure 3. French physicists, active in 2004-05; distribution by year of recruitment (MCFs) or year of last promotion (PRs)



Source: own elaboration on ministerial records

Table 1. Numbers of Physicists in France and Italy in 2004-05, by cohort of birth and rank

	COHORT					
	1930s	**1940s**	**1950s**	**1960s**	**1970s**	**Total**
France-Ranks:						
PR	23	423	179	153	3	781
% by cohort	*2.94*	*54.16*	*22.92*	*19.59*	*0.38*	*100*
MCF	12	341	134	586	357	1430
% by cohort	*0.84*	*23.85*	*9.37*	*40.98*	*24.97*	*100*
						2211
Italy-Ranks:						
PO	168	301	137	16		622
% by cohort	*27.01*	*48.39*	*22.03*	*2.57*		*100*
PA	72	222	177	187	3	661
% by cohort	*10.89*	*33.59*	*26.78*	*28.29*	*0.45*	*100*
RU	1	54	115	301	62	533
% by cohort	*0.19*	*10.13*	*21.58*	*56.47*	*11.63*	*100*
						1816
Total	276	1341	742	1243	425	*4027*

Note: Computed on overall French and Italian samples (i.e. before cleaning).

Table 2. Selected disciplinary fields

	French Universities	Italian Universities
Fields	• 28 Milieux denses et matériaux • 30 Milieux dilués et optique	• Fis/01 Fisica Sperimentale • Fis/02 Fisica Teorica, Modelli E Metodi Matematici • Fis/03 Fisica Della Materia
Ranks	• (PR) *Professor* • (MCF) *Maitre de conference*	• (RU) *Ricercatore universitario* • (PA) *Professore Associato* • (PO) *Professore Ordinario*

Table 3. Overall sample and selected physicists in Italy and France, 2004-05; by gender and rank

FRANCE	Women	Man	Total
MCF	410	1020	1430
PR	80	701	781
Total	490	1721	2211

ITALY	Women	Men	Total
RU	132	401	533
PA	102	559	661
PO	40	582	622
Total	274	1542	1816

Note: Computed on overall French and Italian samples (i.e. before cleaning).

FRANCE	Women	Men	Total
MCF	299	912	1211
PR	50	655	705
Total	349	1567	1916

ITALY	Women	Men	Total
RU	124	387	511
PA	93	525	618
PO	38	550	588
Total	255	1462	1717

Note: Computed on overall French and Italian study samples (i.e. after cleaning for homonymy and "zero-productive' academics).

Table 4. Dependent and explanatory variables in promotion and productivity equations

VARIABLE	DEFINITION
Dependent variables in promotion and productivity equations	
Promotion	Step indicator for being promoted: Not promoted=0, Promoted=1
Productivity (Quantity)	Logarithm of the numbers of articles published by year, plus 1
Productivity (Quality)	Logarithm of the average five years impact factor per year for the journals in which articles have been published (productivity equation), plus 1
Explanatory variables in the promotion equations only:	
Quantity flow	Logarithm of the average number of articles in the three years *t-2*,*t-3* and *t-4*, plus 1
Quality flow	Logarithm of the average five years impact factor of the journals in which articles have been published in the three years *t-2*, *t-3* and *t-4*, plus 1.
Zero flow	Dummy=1 when *Quantity flow*=0 (and =0 if not); and Dummy=1 when *Quality flow*=0 (and =0 if not)
Academics per year	Number of academics nominated in year t
Explanatory variables in the productivity equations only:	
Quantity (number of papers) before promotion	Logarithm of the average number of articles per year published before the year of last promotion, plus 1
Quality (average) before promotion	Logarithm of the average impact factor of the of the journals in which articles have been published before the year of last promotion, plus 1
Zeros before promotion	Dummy=1 when *Quantity before promotion*=0 (and =0 if not); and Dummy=1 when *Quality before promotion*=0 (and =0 if not).
Co-authors quantity	Logarithm of 1 plus the moving average of co-authors' number of articles (with other scientists than the author) in the three years *t-1*, *t-2* and *t-3* before year of observation
Co-authors quality	Logarithm of 1 plus the moving average of average impact factor of the articles published by the co-authors (with other scientists than the author) in the three years *t-1*, *t-2* and *t-3* before year of observation
Co-authors quantity and quality zeros	Dummy=1 when *co-authors quantity* is 0 (and=0 if not); and Dummy=1 when *co-authors quality* is 0 (and =0 if not).
Co-authors number and affiliation dummies	Large Project Dummy=1 if in the three years before *t* (*t-1*,*t-2* and *t-3*) scientists have at least one article with more than 29 co-authors (and=0 if not); Small Project with Foreign co-authors Dummy=1 if in the three years before *t* (*t-1*,*t-2* and *t-3*) scientists have no articles with more than 29 co-authors and have at least one article with an address indicating at least one foreign co-authors (and=0 if not); Small Project with only national co-authors Dummy=1 if in the three years before *t* (*t-1*,*t-2* and *t-3*) scientists have no articles with more than 29 co-authors and have no articles with addresses indicating foreign co-authors (and=0 if not); Small Project with co-authors of unknown affiliations Dummy=1 if in the three years before *t* (*t-1*,*t-2* and *t-3*) scientists have no articles with more than 29 co-authors and have only articles with missing addresses (and=0 if not);
Zero in preceding years	Dummy=1 when scientists published 0 articles in the three years *t-1*, *t-2* and *t-3* before year of observation (and=0 if not).
Zero after promotion	Dummy=1 when scientists published 0 articles after promotion (and=0 if not).

Table 4 (continued). Dependent and explanatory variables in promotion and productivity equations

Explanatory variables in both promotion and productivity equations:	
Wave 1980/1985	Dummy=1 if academic experienced last promotion in 1980 in Italy or in 1985 in France
Gender	Gender dummy variable: Female=1, Male=1
Age group dummies	Dummy= 1 if academic belongs to a specific age class: age group 1 (age<30), age group 2 (30<=age<40), age group 3 (40<=age<50), age group 4 (50<=age<60), age group 5 (age>=60)
Field (28, 30, FIS01, FIS02, FIS03)	Country specific fields of affiliation of physicists in French and Italian universities. In France 28= *Milieux denses et matériaux* and 20= *Milieux dilués et optique*. In Italy Fis01= *Fisica Sperimentale*, Fis02= Fisica Teorica and Fis03= Fisica Della Materia.
(Year) x (cohort of entry)	Interactions between dummies for year of publication and cohort of entry defined as the first year we see the scientist publish (or in case of zero publications before entering or being promoted the year of entry or of promotion)
Promotion before 1975	Dummy variable for the (small) subset of professors promoted before 1975. The inclusion of this dummy in the productivity equations controls for the fact that the publication history of this subset of professors is unknown between promotion and 1975. Its inclusion in the promotion equations is equivalent to eliminating them in the estimation of these equations (the dependent step indicator of promotion being always equal to one for them).

Table 5 Summary statistics of dependent and explanatory variables in promotion equations

	PR (obs. 14094)				PO (obs. 14114)				PA (obs. 12165)			
	mean	sd	min	max	mean	sd	min	max	mean	sd	min	max
prom	0.64	0.48	0	1	0.59	0.49	0	1	0.66	0.47	0	1
Quality flow	0.87	0.61	0	4.01	1.21	0.71	0	3.78	0.97	0.74	0	3.65
Quantity flow	1.42	0.75	0	3.00	1.72	0.70	0	3.16	1.46	0.81	0	3.02
Zero flow	0.14	0.35	0	1	0.08	0.27	0	1	0.16	0.36	0	1
Academics per year	26.97	15.84	0	55	19.12	29.40	0	109	22.24	30.21	0	166
Wave 1985 (FR)/ Wave 1980 (IT)	0.10	0.30	0	1	0.12	0.32	0	1	0.35	0.48	0	1
Gender	0.07	0.25	0	1	0.06	0.25	0	1	0.15	0.35	0	1
Age group 1	0.01	0.08	0	1	0.00	0.07	0	1	0.01	0.09	0	1
Age group 3	0.38	0.49	0	1	0.35	0.48	0	1	0.37	0.48	0	1
Age group 4	0.29	0.45	0	1	0.32	0.47	0	1	0.23	0.42	0	1
Age group 5	0.06	0.23	0	1	0.16	0.37	0	1	0.08	0.27	0	1
Field 30 (FR)	0.30	0.46	0	1								
Fis01 (IT)					0.52	0.50	0	1	0.62	0.49	0	1
Fis02 (IT)					0.23	0.42	0	1	0.18	0.38	0	1
Promotion before 1975	0.02	0.14	0	1	0.05	0.22	0	1				

Note: This table is computed without double names, scientists whose productivity data are high to be credible, and "zero-productive" scientists.

Table 6 Summary statistics of dependent and explanatory variables: productivity equation for French academics

	PR (obs. 9018)				MCF (obs. 12057)			
	mean	sd	min	max	mean	sd	min	max
Productivity(Quantity)	0.81	0.75	0	4.26	0.48	0.64	0	4.89
Productivity(Quality)	1.21	0.98	0	3.19	0.81	0.98	0	3.18
Productivity(effort)	0.15	0.14	0	0.69	0.09	0.13	0	0.69
Quantity before promotion	0.81	0.47	0	2.82	0.60	0.44	0	2.81
Quality before promotion	1.63	0.65	0	2.73	1.34	0.74	0	3.11
Co-authors quantity	0.58	0.75	0	4.04	0.48	0.72	0	4.19
Co-authors quality	0.84	0.98	0	3.11	0.68	0.95	0	3.11
Co-authors zero Dummy	0.56	0.50	0	1	0.64	0.48	0	1
Large Project Dummy	0.02	0.15	0	1	0.02	0.14	0	1
Small Project with Foreign co-authors Dummy	0.52	0.50	0	1	0.36	0.48	0	1
Small project with only National co-authors Dummy	0.24	0.43	0	1	0.23	0.42	0	1
Small Project with co-authors of unknown affiliations Dummy	0.05	0.22	0	1	0.06	0.24	0	1
Wave 1985 (FR)	0.13	0.33	0	1	0.20	0.40	0	1
Gender	0.06	0.23	0	1	0.23	0.42	0	1
Age group 1	0.00	0.00	0	0	0.02	0.13	0	1
Age group 3	0.40	0.49	0	1	0.28	0.45	0	1
Age group 4	0.43	0.50	0	1	0.21	0.40	0	1
Age group 5	0.09	0.28	0	1	0.04	0.20	0	1
Field 30 (FR)	0.29	0.46	0	1	0.34	0.47	0	1
Zeros before promotion	0.09	0.29	0	1	0.15	0.36	0	1
Promotion before 1975	0.03	0.18	0	1	0.00	0.00	0	0
Zero after promotion	0.03	0.18	0	1	0.09	0.28	0	1

Note: This table is computed without double names, scientists whose productivity data are high to be credible, and “zero-productive” scientists.

Table 7 Summary statistics of dependent and explanatory variables: productivity equation for Italian academics

	PO (obs. 8332)				PA (obs. 8089)				RU (obs. 5106)			
	mean	sd	min	max	mean	sd	min	max	mean	sd	min	max
Productivity(Quantity)	1.16	0.89	0	4.36	0.82	0.88	0	4.38	0.94	0.88	0	4.43
Productivity(Quality)	1.55	0.95	0	3.76	1.16	1.02	0	3.17	1.34	1.02	0	3.57
Productivity(effort)	0.15	0.15	0	0.69	0.13	0.16	0	0.69	0.13	0.15	0	0.69
Quantity before promotion	1.00	0.59	0	3.24	0.86	0.53	0	3.32	0.82	0.63	0	3.31
Quality before promotion	1.65	0.81	0	3.11	1.52	0.66	0	2.80	1.35	0.80	0	2.86
Co-authors quantity	1.06	0.79	0	3.50	0.81	0.83	0	3.89	1.00	0.85	0	3.43
Co-authors quality	1.47	0.96	0	3.62	1.09	1.03	0	3.12	1.32	1.01	0	3.11
Co-authors zero Dummy	0.28	0.45	0	1	0.46	0.50	0	1	0.35	0.48	0	1
Large Project Dummy	0.22	0.42	0	1	0.16	0.37	0	1	0.20	0.40	0	1
Small Project with Foreign co-authors Dummy	0.48	0.50	0	1	0.33	0.47	0	1	0.39	0.49	0	1
Small project with only National co-authors Dummy	0.19	0.39	0	1	0.26	0.44	0	1	0.24	0.42	0	1
Small Project with co-authors of unknown affiliations Dummy	0.01	0.11	0	1	0.02	0.15	0	1	0.02	0.14	0	1
Wave 1980 (IT)	0.19	0.39	0	1	0.52	0.50	0	1	0.34	0.47	0	1
Gender	0.05	0.21	0	1	0.13	0.34	0	1	0.28	0.45	0	1
Age group 1	0.00	0.00	0	0	0.00	0.00	0	0	0.00	0.06	0	1
Age group 3	0.25	0.44	0	1	0.44	0.50	0	1	0.38	0.49	0	1
Age group 4	0.47	0.50	0	1	0.34	0.48	0	1	0.13	0.33	0	1
Age group 5	0.27	0.44	0	1	0.12	0.33	0	1	0.01	0.09	0	1
Fis01 (IT)	0.49	0.50	0	1	0.64	0.48	0	1	0.50	0.50	0	1
Fis02 (IT)	0.27	0.44	0	1	0.20	0.40	0	1	0.24	0.43	0	1
Zeros before promotion	0.15	0.36	0	1	0.08	0.27	0	1	0.17	0.37	0	1
Promotion before 1975	0.08	0.28	0	1	0.00	0.00	0	0	0.00	0.00	0	0
Zero after promotion	0.01	0.09	0	1	0.02	0.15	0	1	0.01	0.08	0	1

Note: This table is computed without double names, scientists whose productivity data are high to be credible, and "zero-productive" scientists.

Table 8. Promotion Probit equations for professors (marginal effects).

	PR (FR)	PO (IT)	PA (IT)
Quantity flow	0.047***	0.051***	0.087***
	(0.015)	(0.012)	(0.011)
Quality flow	0.073***	-0.0033	-0.013
	(0.015)	(0.015)	(0.013)
Zero flow	0.20***	0.18***	0.16***
	(0.020)	(0.026)	(0.012)
Academics per year	0.0048***	0.0018***	0.0012***
	(0.00060)	(0.00026)	(0.00031)
Wave 1985 (FR)/ Wave 1980 (IT)	0.27***	0.43***	0.43***
	(0.013)	(0.0085)	(0.011)
Gender	-0.11***	-0.15***	-0.094***
	(0.025)	(0.025)	(0.016)
Age group 1			
Age group 3	0.38***	0.50***	0.29***
	(0.013)	(0.019)	(0.013)
Age group 4	0.59***	0.68***	0.38***
	(0.0092)	(0.013)	(0.011)
Age group 5	0.40***	0.56***	
	(0.0067)	(0.0094)	
Field 30 (FR) / Fis01 (IT)	0.039***	0.0030	0.027**
	(0.012)	(0.014)	(0.013)
Fis02 (IT)		0.13***	-0.0035
		(0.015)	(0.017)
Observations	11844	12431	9979

*See Table 4 for the precise definitions of the variables. We control for the (year)x(cohort of entry) interactions dummies and for the promotion before 1975 dummy variable, both not reported in the Table. Standard errors are given in parentheses, and P-values less than 0.10, 0.05 and 0.01 are respectively denoted by *, ** and ***. Some observations are dropped for PR, PA and PO because the interaction dummies between cohort of entry and calendar year (year)x(cohort of entry) sometimes predict exactly failure/success.*

Table 9. Productivity equations (QUANTITY). The dependent variable is the logarithm (plus one) of the number of articles published per year.

VARIABLES	PR(FR) Heckman	MCF(FR) OLS	PO(IT) Heckman	PA(IT) Heckman	RU(IT) OLS
Quantity before promotion	0.44***	0.33***	0.59***	0.37***	0.48***
	(0.023)	(0.017)	(0.023)	(0.023)	(0.023)
Quality before promotion	0.074***	0.057***	-0.071***	0.054**	0.015
	(0.024)	(0.012)	(0.023)	(0.022)	(0.023)
Co-authors quantity	0.054**	0.13***	0.064***	0.011	-0.040
	(0.022)	(0.015)	(0.021)	(0.023)	(0.025)
Co-authors quality	0.066**	0.029	0.0048	0.014	-0.018
	(0.030)	(0.021)	(0.031)	(0.031)	(0.036)
Co-authors zero Dummy	0.0020	0.10**	-0.13*	-0.10	-0.27***
	(0.063)	(0.044)	(0.066)	(0.064)	(0.077)
Large Project Dummy	1.12***	1.14***	1.12***	1.21***	0.97***
	(0.057)	(0.038)	(0.046)	(0.037)	(0.045)
Small Project with Foreign co-authors Dummy	0.55***	0.43***	0.65***	0.57***	0.39***
	(0.028)	(0.015)	(0.042)	(0.029)	(0.037)
Small project with only National co-authors Dummy	0.34***	0.21***	0.37***	0.34***	0.18***
	(0.029)	(0.016)	(0.043)	(0.027)	(0.035)
Small Project with co-authors of unknown affiliations Dummy	0.25***	0.14***	0.26***	0.23***	0.083
	(0.039)	(0.024)	(0.090)	(0.052)	(0.071)
Wave 1985 (FR)/ Wave 1980 (IT)	-0.026	-0.029*	-0.084***	-0.16***	-0.14***
	(0.027)	(0.016)	(0.029)	(0.029)	(0.047)
Gender	-0.15***	-0.079***	0.029	-0.012	-0.11***
	(0.035)	(0.012)	(0.041)	(0.024)	(0.021)
Age group 1	0	0.098**	0	0	-0.016
	(0)	(0.043)	(0)	(0)	(0.16)
Age group 3	-0.28***	-0.061***	-0.22**	-0.11***	-0.070**
	(0.059)	(0.020)	(0.094)	(0.035)	(0.031)
Age group 4	-0.47***	-0.12***	-0.26**	-0.23***	-0.22***
	(0.082)	(0.026)	(0.11)	(0.042)	(0.051)
Age group 5	-0.64***	-0.11***	-0.35***	0	-0.37***
	(0.093)	(0.036)	(0.11)	(0)	(0.11)
Field 30 (FR)/Fis01 (IT)	-0.018	0.054***	-0.18***	-0.0093	-0.014
	(0.017)	(0.010)	(0.022)	(0.023)	(0.024)
Field Fis02 (IT)			-0.20***	-0.031	-0.055**
			(0.024)	(0.027)	(0.027)
inverse Mill's ratio	-0.33***		-0.14***	-0.25***	
	(0.064)		(0.049)	(0.046)	
Constant	0.38***	-0.20***	0.42***	0.29***	0.52***
	(0.13)	(0.058)	(0.14)	(0.092)	(0.12)
Observations	6866	12057	7183	6462	5106
R-squared	0.362	0.343	0.381	0.514	0.525

*See Table 4 for the precise definitions of the variables. We control for the (year)x(cohort of entry) interactions dummies, for the Zero after promotion dummy variable and for the promotion before 1975 dummy variable, all three not reported in the Table. The Zero in preceding years is normalized to be the reference group for theco-authors number and affiliation dummies. Standard errors are given in parentheses, and P-values less than 0.10, 0.05 and 0.01 are respectively denoted by *, ** and ***. Some observations are dropped for PR, PA and PO because in the selection equation interaction dummies between cohort of entry and calendar year (year)x(cohort of entry) sometimes predict exactly failure/success.*

Table 10. Italy, productivity equation (QUALITY). Quality is the average impact factor of the journals where scientist *i* have published in a given year *t*. The dependent variable of the econometric model is the logarithm of quality plus one.

VARIABLES	PR(FR) Heckman	MCF(FR) OLS	PO(IT) Heckman	PA(IT) Heckman	RU(IT) OLS
Quantity before promotion	0.39***	0.21***	0.34***	0.20***	0.21***
	(0.030)	(0.026)	(0.026)	(0.030)	(0.029)
Quality before promotion	0.34***	0.13***	0.17***	0.13***	0.11***
	(0.031)	(0.019)	(0.026)	(0.029)	(0.030)
Co-authors quantity	-0.019	0.067***	-0.017	-0.060**	-0.082**
	(0.029)	(0.024)	(0.023)	(0.030)	(0.032)
Co-authors quality	0.18***	0.20***	0.22***	0.25***	0.20***
	(0.040)	(0.033)	(0.034)	(0.041)	(0.047)
Co-authors zero Dummy	0.18**	0.32***	0.16**	0.22**	0.064
	(0.083)	(0.070)	(0.073)	(0.084)	(0.099)
Large Project Dummy	0.98***	0.95***	1.25***	1.20***	0.95***
	(0.075)	(0.061)	(0.051)	(0.049)	(0.058)
Small Project with Foreign co-authors Dummy	0.82***	0.72***	1.02***	0.90***	0.73***
	(0.037)	(0.024)	(0.046)	(0.038)	(0.047)
Small project with only National co-authors Dummy	0.63***	0.43***	0.74***	0.59***	0.46***
	(0.038)	(0.025)	(0.047)	(0.035)	(0.045)
Small Project with co-authors of unknown affiliations Dummy	0.47***	0.29***	0.57***	0.36***	0.36***
	(0.051)	(0.038)	(0.099)	(0.067)	(0.092)
Wave 1985 (FR)/ Wave 1980 (IT)	0.037	0.0083	-0.030	-0.022	-0.18***
	(0.036)	(0.025)	(0.032)	(0.037)	(0.061)
Gender	-0.18***	-0.11***	-0.027	-0.013	-0.10***
	(0.047)	(0.019)	(0.046)	(0.032)	(0.028)
Age group 1	0	0.16**	0	0	-0.084
	(0)	(0.069)	(0)	(0)	(0.20)
Age group 3	-0.23***	-0.089***	-0.018	-0.059	-0.072*
	(0.078)	(0.031)	(0.10)	(0.045)	(0.041)
Age group 4	-0.42***	-0.15***	-0.094	-0.15***	-0.29***
	(0.11)	(0.041)	(0.12)	(0.055)	(0.067)
Age group 5	-0.58***	-0.14**	-0.20	0	-0.60***
	(0.12)	(0.058)	(0.13)	(0)	(0.14)
Field 30 (FR)/Fis01 (IT)	0.0052	0.077***	-0.19***	-0.066**	-0.081**
	(0.023)	(0.017)	(0.024)	(0.030)	(0.032)
Field Fis02 (IT)			-0.075***	-0.066*	-0.019
			(0.027)	(0.036)	(0.035)
inverse Mill's ratio	-0.23***		-0.067	-0.036	
	(0.084)		(0.054)	(0.060)	
Constant	-0.17	-0.31***	-0.17	-0.063	0.48***
	(0.17)	(0.093)	(0.16)	(0.12)	(0.16)
Observations	6866	12057	7183	6462	5106
R-squared	0.328	0.286	0.335	0.400	0.398

*See Table 4 for the precise definitions of the variables. We control for the (year)x(cohort of entry) interactions dummies, for the Zero after promotion dummy variable and for the promotion before 1975 dummy variable, all three not reported in the Table. The Zero in preceding years is normalized to be the reference group for theco-authors number and affiliation dummies. Standard errors are given in parentheses, and P-values less than 0.10, 0.05 and 0.01 are respectively denoted by *, ** and ***. Some observations are dropped for PR, PA and PO because in the selection equation interaction dummies between cohort of entry and calendar year (year)x(cohort of entry) sometimes predict exactly failure/success.*

Appendix 1

Distribution of co-authors and number of affiliations listed in the articles.

France

Authors	Addresses 0-5	6-10	11-20	>20	
1-5	65,66	0,16			*65,82*
6-10	23,25	0,85	0,01		*24,11*
11-20	2,63	0,77	0,05		*3,46*
21-29	0,32	0,27	0,06	0,00	*0,66*
Large projects	0,28	0,67	1,12	3,88	*5,95*
	92,15	*2,72*	*1,25*	*3,88*	*100,00*

Italy

Authors	Addresses 0-5	6-10	11-20	>20	
1-5	65,61	1,27	0,01	0,00	*66,89*
6-10	15,61	1,86	0,02		*17,49*
11-20	3,21	1,82	0,15		*5,18*
21-29	0,60	0,95	0,33	0,00	*1,87*
Large projects	0,65	1,27	2,56	4,08	*8,57*
	85,68	*7,16*	*3,06*	*4,09*	*100,00*

Appendix 2

Average number of authors per *Large Project* articles, according to the information available from the database.

	France		Italy	
1	USA	4.39	USA	2.60
2	GERMANY	1.70	INFN	2.30
3	INFN	1.28	ITALY	2.00
4	ENGLAND	1.28	GERMANY	1.25
5	ITALY	1.06	ENGLAND	1.08
6	CNRS	0.83	FRANCE	0.61
7	FRANCE	0.72	CNRS	0.54
8	RUSSIA	0.51	CERN	0.46
9	CHINA	0.48	JAPAN	0.35
10	BELGIUM	0.41	RUSSIA	0.34
11	CERN	0.31	BELGIUM	0.25
12	GREECE	0.29	SPAIN	0.24
13	JAPAN	0.29	SWITZERLAND	0.24
14	INDIA	0.23	CHINA	0.22
15	NORWAY	0.19	GREECE	0.21
16	SPAIN	0.19	NORWAY	0.16
17	SWEDEN	0.11	INDIA	0.13
18	SWITZERLAND	0.08	SWEDEN	0.09
19	DENMARK	0.07	DENMARK	0.08
20	CNR	0.02	CNR	0.04
21	UMR	0.00	INFM	0.00
22	INFM	0.00	UMR	0.00

Appendix 3

Average number of authors per article, according to the information available from the database. We consider only articles with at last one foreign author in small projects

	France		Italy	
1	FRANCE	0.43	ITALY	0.90
2	USA	0.23	INFN	0.39
3	CNRS	0.21	INFM	0.22
4	GERMANY	0.21	USA	0.20
5	UMR	0.16	FRANCE	0.18
6	ITALY	0.10	GERMANY	0.18
7	ENGLAND	0.09	ENGLAND	0.11
8	CHINA	0.08	RUSSIA	0.08
9	SPAIN	0.08	SWITZERLAND	0.08
10	SWITZERLAND	0.08	CNR	0.08
11	RUSSIA	0.07	SPAIN	0.07
12	JAPAN	0.06	CERN	0.07
13	BELGIUM	0.05	CNRS	0.06
14	INFN	0.03	JAPAN	0.05
15	INFM	0.03	UMR	0.03
16	SWEDEN	0.02	CHINA	0.03
17	CNR	0.01	SWEDEN	0.02
18	INDIA	0.01	BELGIUM	0.02
19	DENMARK	0.01	DENMARK	0.02
20	GREECE	0.01	INDIA	0.02
21	CERN	0.01	GREECE	0.01
22	NORWAY	0.00	NORWAY	0.01

Appendix 4

List of first 20 selected journals in terms of publications. The following table shows the first and last year when we see an article published on a journal (*Entry date* of the journal in the database and *exit date* of the journal from the database). Moreover the table shows the number of publications (Italian, French and the sum) available in the database for each journal.

				Publications		
N.	Journal	Entry date	Exit date	Italy	France	Total
1	PHYSICAL REVIEW B	1975	2005	4225	3424	7649
2	PHYSICAL REVIEW LETTERS	1975	2005	3000	2566	5566
3	PHYSICS LETTERS B	1975	2005	4129	1045	5174
4	NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH SECTION A-ACCELERATORS SPECTROMETERS DETECTORS AND ASSOCIATED EQUIPMENT	1984	2005	2446	552	2998
5	JOURNAL OF APPLIED PHYSICS	1975	2005	1245	1542	2787
6	APPLIED PHYSICS LETTERS	1975	2005	1178	1347	2525
7	NUCLEAR PHYSICS B	1975	2005	2039	308	2347
8	PHYSICAL REVIEW A	1975	2005	1281	1036	2317
9	JOURNAL OF CHEMICAL PHYSICS	1975	2005	875	1376	2251
10	PHYSICAL REVIEW D	1975	2005	1586	499	2085
11	PHYSICAL REVIEW E	1993	2005	1207	811	2018
12	SOLID STATE COMMUNICATIONS	1975	2005	831	938	1769
13	JOURNAL OF PHYSICS-CONDENSED MATTER	1989	2005	726	945	1671
14	EUROPHYSICS LETTERS	1986	2005	778	858	1636
15	JOURNAL OF MAGNETISM AND MAGNETIC MATERIALS	1977	2005	535	1060	1595
16	JOURNAL DE PHYSIQUE	1975	1990	257	1193	1450
17	OPTICS COMMUNICATIONS	1975	2005	625	810	1435
18	JOURNAL DE PHYSIQUE IV	1991	2005	264	1140	1404
19	NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH SECTION B-BEAM INTERACTIONS WITH MATERIALS AND ATOMS	1984	2005	731	652	1383
20	PHYSICS LETTERS A	1975	2005	960	349	1309