

Determinants of scientific productivity: A study on Italian and French Physicists

Francesco Lissoni ^{1,6}, Jacques Mairesse ^{2,3,7,8},
Fabio Montobbio ^{1,4}, Michele Pezzoni ^{1,5,6}

¹ KITes/Cespri-Università Bocconi (Milano, IT); ² CREST-INSEE (FR);

³ Politecnico di Torino (IT); ⁴ Università dell'Insubria (Varese, IT);

⁵ Università di Bergamo (IT); ⁶ Università di Brescia (IT)

⁷ UNU-MERIT, Maastricht University (NL); ⁸ NBER (USA)

ABSTRACT: The paper examines the determinants of scientific productivity of academic physicists of the matter in France and Italy active in 2004-05. Productivity is measured through quantity and quality (impact factor) of publications, weighed by the number of co-authors. Different regressions are run for scientists at different ranks, by country. We find evidence of age and cohort effects on all dimensions of productivity, as well as limited evidence of stratification (scientists in top university are more productive, *ceteris paribus*). Gender effects appear noticeable only for France, while some evidence can be gathered on negative effects of stop-and-go recruitment policies in Italy.

1. Introduction

Most of the available studies on scientists' productivity and careers are based upon US data, as well as of theoretical insights which refer to or are based upon the institutional specificities of the US academic system (Long, 1978; Allison and Long, 1990; Levin and Stephan, 1994; Lee and Bozeman, 2005). However, the academic system of many countries does not share the same characteristics of the US one, especially with reference to the degree of autonomy of universities, the relative balance (in size and prestige) of universities and public research organizations, and to size and flexibility of the academic job market.

In this paper we deal with France and Italy, two of the countries whose universities stand at the opposite end of the US, both in term of standing within the public research system and recruitment or financial autonomy. We ask whether in such systems individual scientists' productivity is affected by the same forces that shape it in the US, such forces being age, career incentives, university stratification and individual fixed effects, either observable (such a gender) or not (such as creativity). We also ask whether the frequent stops-and-goes in the recruitment process typical of such two centralized systems have led to the existence of noticeable cohort effects, such as the self-selection of especially productive (or unproductive) scientists in specific cohorts. We do so by examining the productivity profile of French and Italian academic physicists of the matter, either employed in universities (this draft) or in CNRS and CNR (to be included in the next draft). In particular, we observe the scientists who were on active duty in 2004-2005, and their yearly publications (on selected journals) to that date (starting from either 1975 or their first year of career). In section 2 we revise the existing literature on age, cohort and stratification effects on scientific productivity. In section 3, we discuss briefly 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 collection methodology, and the econometric issues related to the model. In section 5, we comment on estimated separate regressions for various academic ranks in France and Italy. Section 6. concludes.

2. Background literature

Interest in the determinants of individual scientists' productivity dates back to the XIX century. 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 on systematized by Alfred Lotka's well-known "power law"; [1].

From the very start, enquiries on scientific productivity, which as Cattell's, 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 country's scientific performance, as measured by

the number of outstanding scientists from that country ([2]; [3]; [4])¹. 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 been busy testing 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 ([5]; [6]; [7]). And although eugenics has long disappeared from the list of policy recommendations derived from studies on scientific productivity, genetic explanations have not, witness the lively debate sparked in 2005 by the then-president of Harvard, Lawrence Summers, with his remarks on the issue of gender bias in science ([8]; [9]; for less contingent studies see: [10]; [11]).

As for the relationship between academic institutions, incentive schemes, and scientists' performance, this is a topic whose importance has greatly increased 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 public funding of science, and possibly to economize on that; 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 ([12]).

An important line of research has referred to the related issues of age and tenure, the latter being a peculiar institution of the academic work environment. In particular, many studies have explored the possibility that individual scientific productivity follows a life cycle, such that productivity increases when the scientist is young, reaches a peak at/before middle age, and declines afterwards ([13]). 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 pressure to publish (and more generally to be productive) felt by scientists who have reached a secure position in their universities. In addition, some institutional variables, such as the stratification of universities according to prestige and funding, may lead us back to the study of increasing returns in science, to the extent that, other things being equal, graduates from prestigious universities have a higher chance to get their first job at institutions in the same league, which comes to their advantage in terms of present and future research productivity ([14]).

2.1 Age, Cohort and Time Effects

Economists and sociologists of science who have produced quantitative studies on productivity have all faced the problem of identifying at the same time age, cohort and period effects in a panel of individual, the so called *identification problem* [25].

Discussion of background literature (as in Mairesse and Turner, 2007) – to be added later

¹ 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, Moody 2007).

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 barely representative of university systems worldwide². 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 ([15]; [16]). 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, with most universities often relegated at a secondary role or just to their teaching duties; 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; and professors are university employees who can bargain for their wages and working conditions on an individual basis; the existence of a proper academic job market also allows scientists with a strong publication record to move from one university to another in search of better paid or better funded research positions ([17]). This is hardly the case in France and Italy, universities there being forbidden to issue new posts without ministerial consent. In these centralized recruitment systems university staff are considered civil servants, who are employed by the government and selected by peers' commissions elected by national members of the relevant discipline, or nominated by the government itself. In these countries, competition between universities for the recruitment of the most promising or productive scientists is unheard of. At the same time, strict dependence of universities from government funding and regulation has often provoked time irregularities in the recruitment process. Italy, for example, has a long history of prolonged "dry" spells (during which universities did not recruit any new scientist, due to funding shortages or ongoing policy revisions of recruitment rules) followed by sudden waves of mass recruitment, often under the political pressure exerted by large number of young scientists seeking a secure position after years of precariousness.

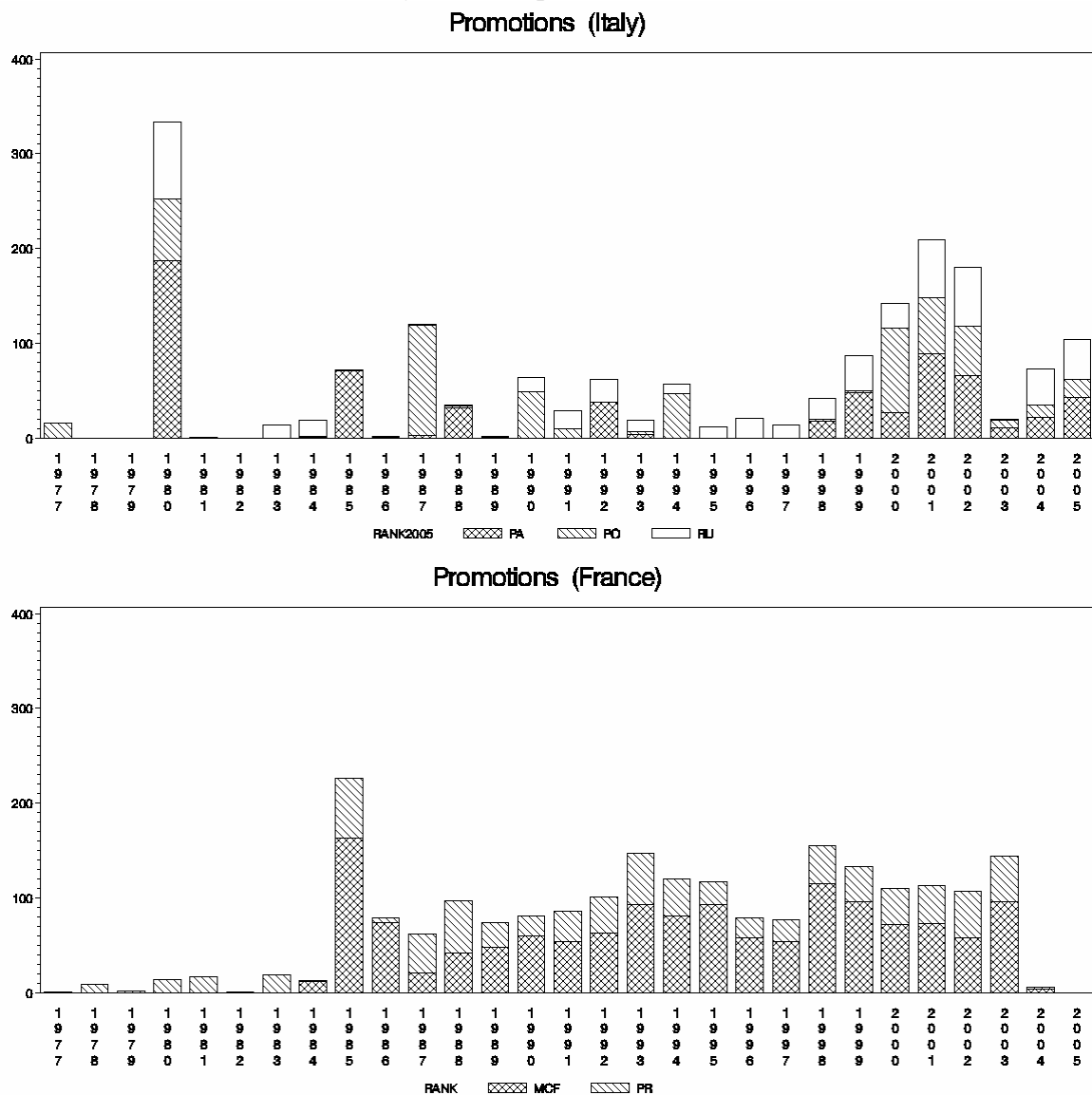
3.1 Scientific careers in France and Italy

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, which reproduce a hierarchy apparently similar to US assistant-associate-full professorship scale: "*Ricercatore universitario*" (RU), "*Professore Associato*" (PA) and "*Professore ordinario*" (PO). 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 it is their members, and not individual universities or department, who control the

² An important exceptions is [21]

recruitment process³. Before accessing any of the above-mentioned positions (most often RU in Italy and MCF in France) young scientists go through post-doc spells of various length, depending on whether the universities wishing to recruit them will obtain enough funds for issuing a new position, on whether the government will give its consensus to the issuing of such position (which may be suspended for either financial or regulatory reasons), and on how long it will take to the various ministerial and university commissions to go through the entire recruitment procedure. Years spent by candidates as *temporary assistants* are limited in France to one year to avoid the abuse of this kind of fixed-term contracts [20].

Figure 1. Physicists of the matter in Italy and France, active in 2004-05; distribution by year of recruitment (MCFs; RUs) or year of last promotion (PAs; POs; PRs)



Source: French and Italian ministries of education

³ Notice also that the US system is openly stratified according to the research *vs.* teaching intensity of institutions, and the latter's wealth. 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.

Historically, Italian and French universities faced an increasing demand for higher education instruction during the 60s and 70s due to the booming European economy. Both countries answered to the growing number of students hiring a large number of young assistants with non-tenured positions and fixed-terms contracts (*Assistants* in France and *Professori incaricati* but also *contrattisti* and *assegnisti* in Italy). These categories fought for tenured positions during the 1970s and they succeeded in early 80s, when two reform acts were passed: law 382 in Italy (in 1980) and the Higher Education Act, passed in 1984 in France. Both laws reformed the recruitment process by introducing the ranking system we described above, and by changing the hiring rules. At the same time, the two laws provided for most of the *professori incaricati* and *assistants* of the time to obtain a tenured positions as *ricercatori* or *professori associati* in Italy, or *maitres de conference* in France. As a result, each country saw a massive recruitment wave (respectively in 1980 and 1985), which was followed by a prolonged dry recruitment spell in Italy, and a sensible decline in France. Figure 1 illustrates this effect for the field of physics of the matter, as it reports 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 peaks are clearly visible.

The 1980 and 1984 laws also changed the rules for the selection of new academic staff.

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 recruited for individual universities by a committee of three PAs and POs, all appointed by the ministry.

Together with universities' lack of financial autonomy, this highly centralized recruitment system was deemed responsible for the difficulties met by universities in speeding up the recruitment process, which languished throughout the 1980s (see again figure 1). In 1993, universities were finally given some degree of financial autonomy, which included some freedom to allocate resources to their preferred expense targets. In 1998 a new recruitment system came into place, which is still in place. 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 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. Notice that the disciplinary classification is very detailed, and it is negotiated periodically between the Ministry and the leading senior professors of the country⁴. 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

⁴ Overall there are no less than 300 disciplines; some of them, such as "Ship Architecture" and "Ship and Marine Plants" – notice they are considered different - counted less than 30 affiliates in 2005.

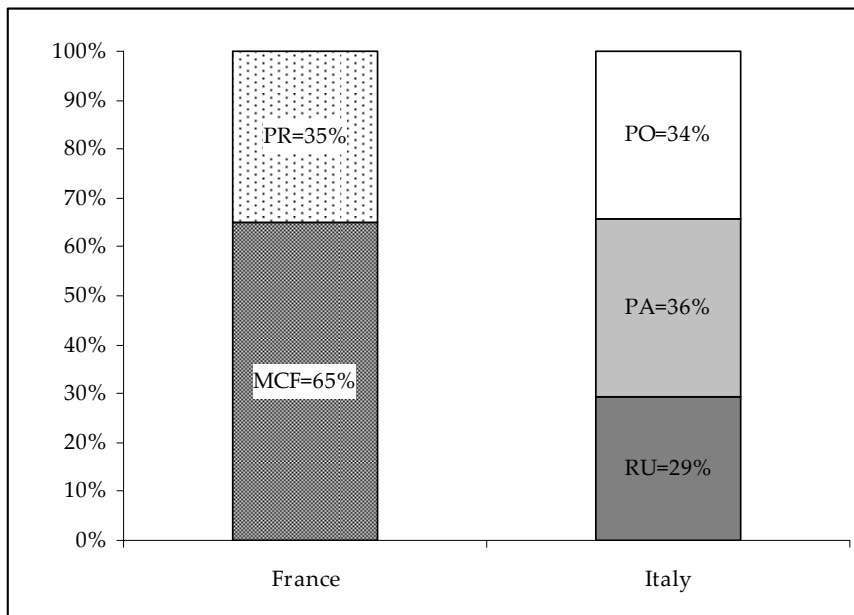
can always refuse to nominate him/her and launch a new job call. In practice, most commissions try to steer the selection process towards candidates who they know will be palatable to the university.⁵ Once again the introduction of the new law coincided with a new wave of recruitment; although less dramatic than the 1980 one, it is still visible in figure 1, over the years between 1999 and 2001.

The French recruitment system was also (until the recent reform introduced by the Sarkozy presidency, in August 2007) a very centralized and discipline-centred. As in the Italian system, professors are classified according to their disciplines, although the latter are not as much finely specified (there are 73 of them overall). Every year, the central government issues a list of vacancies, by discipline and institution, both for the positions of *maitre a conference* and *professor*. Applicants for positions need first to get a *qualification* certificate from the Ministry of Education (which is granted on the basis of their CV) and then apply to one of the universities which has been granted a vacancy; there, a commission of local members of the discipline will choose among the applicants. CNU (*Conseil National des Universités*), made by a part of member elected and part designed by Ministry of education, screens the candidates for tenured academic staff positions giving the *qualification* valid for four years. Every spring a list of qualified candidates is drawn up in every academic discipline. Only candidates “*qualified*” by CNU may apply to vacant positions in specific institution. Applicants are processed by institutions by a *commission de spécialités* organized around the disciplines and drawn among elected internal and external to university members. 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 ([18]; [19]; [20]). None of these reforms, however, has gone as far as to grant universities total freedom in recruitment matters, nor it has diminished much the role of disciplines in it. In addition, 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. In between 1980 and the mid-1990s reform, the Italian system was even more centralized than today: new positions for all universities were issued jointly, and the *concorso* took place at the national level, nominally every other year for each academic rank, but most often only if and when the public finances allowed an increase in spending for academic staff by the Ministry of Education.

⁵ This requires a lot of political background work by the university, in order to steer the election of the committee members in favour of friendly candidates. Friendly ministers may also help in this direction by creating *ex novo* new disciplines with very few affiliates, wherein electoral deals are easier to strike. Once a friendly commission has been elected, the commissioners will take care of letting some perspective candidates to know that their participation to the competition would be of embarrassment to the perspective winners, and to eliminate other candidates on the basis of legal formalities. On the long-standing importance of such practices in the Italian academic systems, see Clark (1977). In some job competitions, the commissions are also let free to declare two winners (two *idonei*), only one of which will be selected by the university. In this way, the commissioners are more at freedom to let both the strongest candidate and a local insider win (then the latter will be chosen for the job, and the former will be possibly called by some other university). Needless to say that this possibility is often exploited for more political bargaining among the professoriate: elected commissioners trade the nomination of a local candidate liked by the university issuing the job call in exchange for a ‘fit-for-the-job’ certificate for another candidate, who they wish to recruit in their own university (Moscati, 2001).

Seniority and progress in the academic hierarchy are the only way a scientist can increase his/her wage both in France and Italy: salaries are fixed by law and cannot be negotiated on an individual basis, nor they change from one institution to another⁶. In Italy, such an exclusive link between wage and academic rank, and the low level of initial wages, has often produced a strong pressure in favour of the advancement of senior RUs to the associate (PA) position, and from here to full professorship (PO). This pressure is reflected in the even distribution of scientists across the three ranks: instead of being the minority at the top of the academic ladder (as in France), Italian full professors (POs) amount to little more than one third of the whole academic personnel of universities, with PAs and RUs holding similar shares (figure 2).

Figure 2. Physicists of the matter in Italy and France, 2004-05; distribution by rank (France: MCFs; RUs; Italy: PAs; POs; PRs)



Source: French and Italian ministries of education

Time irregularities in recruitment and career advancement mechanisms determine the cohort and rank composition of active scientists. We define cohort of birth as the decade when individual was born, starting from 1930. Table 1 shows, as expected, that the lower academic ranks are populated by younger cohorts of birth in both countries, while higher ranks host scientists from more senior cohorts. Notice however that the top academic rank in Italy (PO) is composed of many more scientists from the oldest cohort than its French equivalent (PR). At the same time, only 11% of Italian RUs belong to the youngest cohort (1970s), as opposed to 26.8% 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.

⁶ Admittedly, jobs in higher positions or in more prestigious universities are more visible and offer better chances of consultancy. The latter may be an important, possibly dominant source of income for professional faculty (law, medicine and engineering). However, this is not (or not often) the case for physics, which we consider here.

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 Recherche* 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. In terms of staff, CNR nowadays weighs much less in Italy than CNRS does in France, and much less than it used to be. The average age of its research staff is also considerably high, and increasing ([21]). In the past decade, CNR has undergone recurrent restructuring process, and proposals have been put forward to close it down altogether.

Table 1. Physicists of the matter in Italy and France, 2004-05; by cohort and rank

	COHORT					Total
	1930s	1940s	1950s	1960s	1970s	
<i>France – Ranks:</i>						
MCF	7	241	111	528	325	1,212
PA	65	204	169	177	3	618
<i>Italy – Ranks:</i>						
PO	159	283	130	16	0	588
PR	21	372	166	143	3	705
RU	1	48	112	290	60	511
<i>Total</i>	253	1,148	688	1,154	391	3,634

3.2 Implications for scientific productivity

This brief examination of the recruitment process in the two countries, and of the relative balance of universities and PROs inside the research system, suggest a few observations on the factors affecting scientific productivity. In principle, productivity is a key determinant of career advancement in both countries, more so in Italy than in France, 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 (much less so for moves up to the top PO position; Pezzoni et al., 2008). As a consequence, other things 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 cohorts effects 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 US-based 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 a prolonged dry spell, could be either highly productive (in case the long wait ended up selecting only the most motivated scientists) or quite un-productive (in case the long wait pushed the more brilliant scientists – who attach higher opportunity costs to waiting – to emigrate or to quit the academic career altogether).

We expect productivity to grow over time (calendar years) for at least three reasons. First, both in Italy and in France public funding for research 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 big 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..” [27].

In principle, one should expect stratification by university or department to exist both in France and in Italy. In France, stratification is induced by CNRS’ decisions to open laboratories within selected universities (UMR), which bring with to the latter both human and financial resources for research. In Italy, for physics, a similar role is played by INFN (the national institute of physics of the matter) and INFN (the national institute of nuclear physics, which is part of CNR). However, both INFN and INFN bear much less weight, in financial terms, than CNRS, and tend to support existing academic research teams, rather than supplementing existing or creating new ones with their own personnel. So, we expect France to exhibit stronger stratification effects. As for gender and Matthew effects, no apparent reason exist to think of peculiarities for France and Italy with respect to the US-based evidence at hand. That is, we expect to find evidence of both effects in both countries.

4. Data and Methodology

4.1 Data collection and sample

The paper is based on the complete lists of tenured scientists from Italian and French universities, active in academic year 2004/2005 in a number of fields related to physics of

the matter. Both lists were provided by the Ministries of Education of the two countries with different disciplinary classification. The information coming with the lists include the date of last promotion (e.g. for an Italian scientist who in 2004/05 held the position of PA we also know in what year he was promoted to that rank; but we do not know anything about the date of his first recruitment as RU).

Table 2 reports the selected fields, and it shows that, for comparability purposes, we had to stretch our definition of physics of the matter to the point of including all academic physicists, with the only exception of astro-physicists and nuclear physicists. 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 5-year average impact factor higher than 0.5. We downloaded all articles published in between 1975 and 2005, with at least one author whose surname and initial(s) matched those of at least one scientist in our lists.

Table 2. Selected disciplinary fields

	French UNIV	Italian UNIV
Fields (chosen according to comparability)	<ul style="list-style-type: none"> • 28 Milieux denses et materiaux • 30 Milieux dilues et optique 	<ul style="list-style-type: none"> • Fis/01 Fisica Sperimentale • Fis/02 Fisica Teorica, Modelli E Metodi Matematici • Fis/03 Fisica Della Materia
Ranks	<ul style="list-style-type: none"> • (PR) <i>Professor</i> • (MCF) <i>Maitre de conference</i> 	<ul style="list-style-type: none"> • (RU) <i>Ricercatore universitario</i> • (PA) <i>Professore Associato</i> • (PO) <i>Professore Ordinario</i>

Table 3. Selected physicists of the matter in Italy and France, 2004-05; by gender and rank

	FRANCE			ITALY			
	Women	Men	Total	Women	Men	Total	
MCF	405	990	1395	RU	129	395	524
PR	78	679	757	PA	99	540	639
				PO	40	566	606
<i>Total</i>	483	1669	2152 ⁷	<i>Total</i>	268	1501	1769

The selected journals cover uniformly the disciplinary fields, as scientists belonging to each field publish on 250/300 different journals among selected. Such a search strategy, although the only possible one given 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

⁷ Without double names and scientists whose productivity data are high to be credible.

ministerial lists and all scientists with both common surnames and an apparent productivity far too high for being credible. This operation led to dropping about 3% of scientists in the original lists, and left us with over 2000 French scientists, and little less than 1800 Italians (Table 3).

For almost 2.9% (52) of the selected scientists in Italy we could not find any publication in the selected journals, while the same figure for France is about 10.9% (236 scientists). 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 (such as French journals, as opposed to US ones) or that it is more common for French academics to dedicate themselves exclusively to teaching. Given the impossibility to test these explanations, we simply dropped these individuals from the following econometric models. Therefore, the final sample is made by 1916 French and 1769 Italian academics. 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 come as an unbalanced panel. Scientists are observed from the year of first publication to 2005 (i.e. the last year of observation). Alternatively, in the case of absence of at least one publication before the year of promotion to their present position, the first year of observation is the same as the year of promotion.

The implementation of an econometric model to evaluate determinants of productivity to this kind of data rises three kind of problems.

First, the endogeneity between promotion and productivity. It is impossible to establish a one-way causal relationship between the two variables. In order to solve this problem, we consider five equations, one for each rank, and we account for the productivity only after promotion. This solves also the problem of data design, such that the date of last promotion cannot be later than 2005.

This solution rises however a second problem, that of sample selection ([22]; [23]; [24]). None of our five regression is based on a random sample, because it is not based on a random sample, but only those who, 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 think of it as caused by an omitted variable in the productivity equation and to include in the model a selection equation ([22]). Such equation gives information on the determinants of promotion.

$$y_1 = x_1 \beta_1 + \mu_1 \quad \text{Productivity equation}$$

$$y_2 = 1[x_2 \beta_2 + \mu_2 > 0] \quad \text{Promotion/selection equation}$$

This model can be estimated either by a two-step Heckman's estimation ([22]) or by maximum likelihood. We use the first because of problems of convergence, due to the large size of our panel and the large number of explanatory variables we will make us of.

However, the two-step procedure is not applicable in two out of five of our 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 problem we simply applied OLS to the productivity equation of MCFs and RUs, without any correction.

The third problem is to control for unobservable heterogeneity. Individual characteristics such as ability and effort are not directly observable, but have a crucial impact on productivity after promotion and for promotion too. We account for individual heterogeneity by considering the scientist's the average yearly productivity before promotion (*quantity/quality before promotion*). On the side of promotion/selection equation we introduce an index called *quantity/quality flow* to evaluate the impact of past scientific productivity. The *flow* index in year t is the average productivity (in terms of quantity and quality) in the three years $t-2$, $t-3$ and $t-4$.

4.3 The determinants of selection and scientific productivity

We run five separate regression exercise, one for each academic positions under study: RU, PA and PO for Italy; MCF and PR for France. Each exercise requires estimating two equations, one in which the dependent variable is promotion to the observed position, with past scientific productivity as the key explanatory variable; the other in which the dependent variable is productivity, for scientists in the observed position. Each exercise is run twice, one for each proxy through which we can measure productivity, respectively:

- *quantity* (log of the number of articles per year)
- *quality* (log of the average impact factor of the journals in which a scientist has published, per year)

The panel includes N individuals from C cohorts, observed in P periods (years). The semi-parametric model we applied for our analysis includes a dummy variable for each cohort-period combination (CxP).

Individuals included in the selection equations are all those whose

Cohort dummies refer to entry years, which we define as the first year in which the individual scientist published an article (i.e. the year in which she first joined the scientific community) or, if the scientist does not have publications before last promotion, the date of last promotion.

Period dummies in the CxP combinations refer to calendar years, as we expect each individual scientist's productivity to grow over time, other things being equal, for the reasons explained in section 3.2.

Table 4 lists all the variables we consider for either the productivity or selection regressions.

Table 4. Explanatory variables for productivity and/or selection regression

CO-VARIATES	EXPLANATION
<i>In both the selection and productivity equations:</i>	
<i>Age dummies</i>	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)
<i>Wave 1980/1985</i>	Dummy=1 if academic experienced last promotion in 1980 in Italy or in 1985 in France
<i>Woman</i>	Gender dummy variable
<i>Quantity (number of papers) before promotion</i>	Logarithm of the productivity before the year of professor's last promotion
<i>Quality (average) before promotion</i>	Logarithm of the average quality of the publications before the year of professor's last promotion
<i>Zeros before promotion</i>	Dummy=1 when <i>Quantity before promotion</i> =0 (and <i>Quality before promotion</i> =0)
<i>Field</i> (28 30 FIS/01 FIS/02 FIS/03)	Country specific field where professor is classified by the Ministry of education
<i>Only in the productivity equations:</i>	
<i>Co-authors quantity</i>	moving average of productivity of co-authors in the three years before year of observation
<i>Co-authors quality</i>	moving average of average quality of the articles published by the co-authors in the three years before year of observation
<i>Co-authors zeros</i>	Dummy=1 when <i>co-authors quantity</i> is 0 (and <i>co-authors quality</i>)
<i>Zero after promotion</i>	Dummy=1 when scientists published 0 articles post promotion
<i>Promotion before 1975</i>	dummy variable for professor promoted before 1975, without any information about the publication history before promotion (no information about <i>Quantity before promotion</i>)
<i>Quantity flow</i>	Average flow productivity of the three years $t-2, t-3$ and $t-4$
<i>Quality flow</i>	Average impact factor of the paper published in the three years $t-2, t-3$ and $t-4$
<i>Zero flow</i>	Dummy=1 when <i>Quantity flow</i> =0 (and <i>Quality flow</i> =0)
<i>Promotion</i>	Step dummy for being promoted
<i>Quantity</i>	flow of articles year by year
<i>Quality</i>	average quality of the papers published year by year
<i>Year X cohort of entry</i>	Interactions between dummies for calendar year of publication and cohort of entry defined before as the first year we see the scientist publish or, in case of zero publications before promotion/entering, the year of promotion
<i>Coauthors' location/affiliation</i> (Italy, France, US, CNR, CNRS, CERN, INFN)	Dummies=1 in period t if in three years before the authors have at least one author affiliated to a PRO or coming from <i>Italy, France</i> or the <i>US</i> .
<i>Only in the selection equations:</i>	
<i>Academics per year</i>	Number of academics nominated in year t

We test for *age* effects by introducing in both the selection and productivity equations five dummy variables, representing as many decades in which our scientists were born. Our expectation is to find a negative impact of scientist's age on productivity ([13]; [25]). 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 (this is especially true in France and even more for Italy, as opposed to the US, as we discussed in section 3; [26]).

In both the equations for selection and productivity we test for gender effects (*woman* dummy), and expect them to be negative.

Historical time of recruitment is also a variable of substantial interest. We test the hypothesis that being hired or promoted in years of massive recruitment, such 1980 in

Italy or 1985 in France, may impact negatively on the recruited/promoted scientist's expected future productivity. This is because a sudden "opening of the gates", typical of centralized systems such as the Italia and the French, allow in the system scientists of minor quality or dedication. We notice that this variable may also be affected by the problem of selection due to the combination of two effects. One is "wave" effect just described, the other is due to the fact that less productive scientists may never be promoted and stay in the same rank up to 2005, the year in which we observe them. Notice however that scientists recruited as RUs or MCFs or promoted to PA in 1980 in Italy and 1985 in France, and who are still in the same rank after 20/30 years, are quite a substantial share of the academics active in 2005. Therefore, although we do not have information on who and how many of those recruited in 1980 moved on to higher positions, we know how many of them in 2005 are still in the same rank as 1980. In both the selection and the productivity equation we identify them 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.

Finally, 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 2* and *Field 30* for France; *Field FIS/0*, *Field FIS/02*, and *Field FIS/03* for Italy). We expect also that propensity of being promoted are field-specific for two reasons. First resources to hire new employees could not be distributed homogeneously among all the disciplines and second, some disciplines could be more prolific in terms of new discovery and research paths requiring more academics to work on. 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 the productivity equations, we also control for the relationship of our scientists with the rest of the scientific community. First, we consider the past productivity of each individual scientist i 's co-authors (*Co-author's quantity and quality*⁸). We presume that having worked with other productive individuals up to year t may have a positive effect on a scientist's productivity in that year, because first-hand contacts with productive colleagues allows to tap in their knowledge and experience. Notice that from co-authors' productivity measures we exclude the articles written with the scientist i under analysis. In addition, we consider the co-authors' affiliations and geographical location, as listed on the ISI publication records: a set of dummy variables tell us whether the individual scientist i has signed at least one article with at least one co-author affiliated either to CNRS, CNR, INFN (*Istituto Nazionale Fisica della Materia*), UMR (*Unite Mixte de Recherche*) or CERN, and with at least one co-author from the US, France or Italy.

⁸Productivity in terms of moving average of the average number of publication (or quality of publications) in the 3 years before the year of observation, from $t-2$ to $t-4$

In order to capture unobserved heterogeneity of professors' skills, on the side of productivity equation, we look at productivity before promotion⁹ (*quantity/quality before promotion*) and, on the side of selection equation, we keep the productivity in the years before observation (*quantity/quality flow*).

Finally, the dataset needs some dummies to control for the design of the data. Such design suffers of a left truncation problem due to the unavailability of information about publications before 1975. The variable *promotion before 1975* is a dummy for academics promoted before the first year of observation of the publications. Another problem is that all the variables are affected by a large presence of zero values. Dummy variables are used to correct for the large presence of zeros and different truncation problems are not reported in the results.

Table 5 reports summary statistics for all the regressors, by regression exercise.

Table 5a Summary statistics of dependent variables and covariates: selection equation

	MCF				PA				PO				PR				RU			
	mean	sd	min	max	mean	sd	min	max	mean	sd	min	max	mean	sd	min	max	mean	sd	min	max
Promotion	0.89	0.31	0.00	1.00	0.66	0.47	0.00	1.00	0.59	0.49	0.00	1.00	0.64	0.48	0.00	1.00	0.80	0.40	0.00	1.00
Age group 1	0.02	0.14	0.00	1.00	0.01	0.09	0.00	1.00	0.00	0.07	0.00	1.00	0.01	0.08	0.00	1.00	0.02	0.13	0.00	1.00
Age group 3	0.29	0.45	0.00	1.00	0.37	0.48	0.00	1.00	0.35	0.48	0.00	1.00	0.38	0.49	0.00	1.00	0.32	0.47	0.00	1.00
Age group 4	0.19	0.39	0.00	1.00	0.23	0.42	0.00	1.00	0.32	0.47	0.00	1.00	0.29	0.45	0.00	1.00	0.10	0.30	0.00	1.00
Age group 5	0.04	0.19	0.00	1.00	0.08	0.27	0.00	1.00	0.16	0.37	0.00	1.00	0.06	0.23	0.00	1.00	0.01	0.08	0.00	1.00
Wave 1980	0.00	0.00	0.00	0.00	0.35	0.48	0.00	1.00	0.12	0.32	0.00	1.00	0.03	0.16	0.00	1.00	0.27	0.45	0.00	1.00
Wave 1985	0.21	0.41	0.00	1.00	0.12	0.33	0.00	1.00	0.00	0.00	0.00	0.00	0.10	0.30	0.00	1.00	0.00	0.06	0.00	1.00
Woman	0.23	0.42	0.00	1.00	0.15	0.35	0.00	1.00	0.06	0.25	0.00	1.00	0.07	0.25	0.00	1.00	0.27	0.44	0.00	1.00
Quantity flow	0.55	0.54	0.00	4.03	0.97	0.74	0.00	3.65	1.21	0.71	0.00	3.78	0.87	0.61	0.00	4.01	1.00	0.77	0.00	3.53
Quality flow	1.04	0.82	0.00	3.06	1.46	0.81	0.00	3.02	1.72	0.70	0.00	3.16	1.42	0.75	0.00	3.00	1.53	0.80	0.00	3.05
Zero flow	0.30	0.46	0.00	1.00	0.16	0.36	0.00	1.00	0.08	0.27	0.00	1.00	0.14	0.35	0.00	1.00	0.14	0.34	0.00	1.00
Promotion before 1975	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.22	0.00	1.00	0.02	0.14	0.00	1.00	0.00	0.00	0.00	0.00
Academics per year	51.9	32.9	0	121	22.2	30.2	0	166	19.1	29.4	0	109	27.0	15.8	0	55	25.2	18.8	0	73
Field Fis/01	0.00	0.00	0.00	0.00	0.62	0.49	0.00	1.00	0.52	0.50	0.00	1.00	0.00	0.00	0.00	0.00	0.52	0.50	0.00	1.00
Field Fis/02	0.00	0.00	0.00	0.00	0.18	0.38	0.00	1.00	0.23	0.42	0.00	1.00	0.00	0.00	0.00	0.00	0.23	0.42	0.00	1.00
Field 30	0.35	0.48	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30	0.46	0.00	1.00	0.00	0.00	0.00	0.00

⁹ We found that the various measures of productivity (quantity, quality) before promotion are correlated mainly due to the correspondence of periods with zero articles and zero average quality. The correlation without zeros between quality and quantity is positive for every rank even if with lower coefficients. This result exclude a trade off between quantity and quality of scientific production.

Table 5b Summary statistics of dependent variables and covariates: productivity equation

	MCF				PA				PO				PR				RU			
	mean	sd	min	max	mean	sd	min	max	mean	sd	min	max	mean	sd	min	max	mean	sd	min	max
log(quantity)	0.48	0.64	0.00	4.89	0.82	0.88	0.00	4.38	1.16	0.89	0.00	4.36	0.81	0.75	0.00	4.26	0.94	0.88	0.00	4.43
log(quality)	0.81	0.98	0.00	3.18	1.16	1.02	0.00	3.17	1.55	0.95	0.00	3.76	1.21	0.98	0.00	3.19	1.34	1.02	0.00	3.57
Age group 1	0.02	0.13	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.00	1.00
Age group 3	0.28	0.45	0.00	1.00	0.44	0.50	0.00	1.00	0.25	0.44	0.00	1.00	0.40	0.49	0.00	1.00	0.38	0.49	0.00	1.00
Age group 4	0.21	0.40	0.00	1.00	0.34	0.48	0.00	1.00	0.47	0.50	0.00	1.00	0.43	0.50	0.00	1.00	0.13	0.33	0.00	1.00
Age group 5	0.04	0.20	0.00	1.00	0.12	0.33	0.00	1.00	0.27	0.44	0.00	1.00	0.09	0.28	0.00	1.00	0.01	0.09	0.00	1.00
wave 1980	0.00	0.00	0.00	0.00	0.52	0.50	0.00	1.00	0.19	0.39	0.00	1.00	0.04	0.19	0.00	1.00	0.34	0.47	0.00	1.00
wave 1985	0.20	0.40	0.00	1.00	0.16	0.37	0.00	1.00	0.00	0.00	0.00	0.00	0.13	0.33	0.00	1.00	0.00	0.06	0.00	1.00
woman	0.23	0.42	0.00	1.00	0.13	0.34	0.00	1.00	0.05	0.21	0.00	1.00	0.06	0.23	0.00	1.00	0.28	0.45	0.00	1.00
co-authors																				
quantity	0.48	0.72	0.00	4.19	0.81	0.83	0.00	3.89	1.06	0.79	0.00	3.50	0.58	0.75	0.00	4.04	1.00	0.85	0.00	3.43
quality	0.68	0.95	0.00	3.11	1.09	1.03	0.00	3.12	1.47	0.96	0.00	3.62	0.84	0.98	0.00	3.11	1.32	1.01	0.00	3.11
zeros	0.64	0.48	0.00	1.00	0.46	0.50	0.00	1.00	0.28	0.45	0.00	1.00	0.56	0.50	0.00	1.00	0.35	0.48	0.00	1.00
quantity																				
before																				
promotion	0.60	0.44	0.00	2.81	0.86	0.53	0.00	3.32	1.00	0.59	0.00	3.24	0.81	0.47	0.00	2.82	0.82	0.63	0.00	3.31
quality																				
before																				
promotion	1.34	0.74	0.00	3.11	1.52	0.66	0.00	2.80	1.65	0.81	0.00	3.11	1.63	0.65	0.00	2.73	1.35	0.80	0.00	2.86
zeros before																				
promotion	0.15	0.36	0.00	1.00	0.08	0.27	0.00	1.00	0.15	0.36	0.00	1.00	0.09	0.29	0.00	1.00	0.17	0.37	0.00	1.00
Promotion																				
before 1975	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.28	0.00	1.00	0.03	0.18	0.00	1.00	0.00	0.00	0.00	0.00
Zeros after																				
promotion	0.09	0.28	0.00	1.00	0.02	0.15	0.00	1.00	0.01	0.09	0.00	1.00	0.03	0.18	0.00	1.00	0.01	0.08	0.00	1.00
Field Fis/01	0.00	0.00	0.00	0.00	0.64	0.48	0.00	1.00	0.49	0.50	0.00	1.00	0.00	0.00	0.00	0.00	0.50	0.50	0.00	1.00
Field Fis/02	0.00	0.00	0.00	0.00	0.20	0.40	0.00	1.00	0.27	0.44	0.00	1.00	0.00	0.00	0.00	0.00	0.24	0.43	0.00	1.00
Field 30	0.34	0.47	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.29	0.46	0.00	1.00	0.00	0.00	0.00	0.00
French co-																				
authors	0.46	0.50	0.00	1.00	0.18	0.38	0.00	1.00	0.29	0.45	0.00	1.00	0.65	0.48	0.00	1.00	0.21	0.41	0.00	1.00
US co-																				
authors	0.09	0.29	0.00	1.00	0.17	0.38	0.00	1.00	0.23	0.42	0.00	1.00	0.13	0.34	0.00	1.00	0.22	0.41	0.00	1.00
CNR co-																				
authors	0.02	0.13	0.00	1.00	0.24	0.43	0.00	1.00	0.30	0.46	0.00	1.00	0.03	0.16	0.00	1.00	0.23	0.42	0.00	1.00
CNRS co-																				
authors	0.42	0.49	0.00	1.00	0.11	0.31	0.00	1.00	0.17	0.37	0.00	1.00	0.56	0.50	0.00	1.00	0.12	0.32	0.00	1.00
CERN co-																				
authors	0.01	0.11	0.00	1.00	0.12	0.32	0.00	1.00	0.21	0.40	0.00	1.00	0.02	0.13	0.00	1.00	0.14	0.34	0.00	1.00
INFM co-																				
authors	0.02	0.14	0.00	1.00	0.20	0.40	0.00	1.00	0.23	0.42	0.00	1.00	0.02	0.15	0.00	1.00	0.26	0.44	0.00	1.00
Italian co-																				
authors	0.06	0.24	0.00	1.00	0.71	0.45	0.00	1.00	0.86	0.35	0.00	1.00	0.08	0.27	0.00	1.00	0.79	0.41	0.00	1.00
UMR co-																				
authors	0.22	0.42	0.00	1.00	0.02	0.14	0.00	1.00	0.03	0.16	0.00	1.00	0.25	0.44	0.00	1.00	0.03	0.16	0.00	1.00

5. Results and discussion

In what follows we follow the Heckman procedure and comment separately the results from the estimation of the selection equation and the productivity equation. Notice however that the Heckman procedure was not applied to MCFs and RUs simply because we do not have enough information to apply the first step. MCFs and RUs are observed during the starting period of their careers and often we do not have information before their hiring in the academia. In other words we do not see enough history of the individual to explain the promotion with a *probit* model.

5.1 selection equation

Results from selection equation estimates for Italy and France are reported in tables 6 and 7, respectively.

As expected, an Italian or French scientist's chances of being promoted grow with age, which confirms the role of seniority in both academic systems.

Table 6. Italy, selection equation

COEFFICIENT	PO Marginal Effects prom	PA Marginal Effects prom
Age group 1		
Age group 3	0.50*** (0.019)	0.29*** (0.013)
Age group 4	0.68*** (0.013)	0.38*** (0.011)
Age group 5	0.56*** (0.0094)	
Wave 1980	0.43*** (0.0085)	0.43*** (0.011)
Woman	-0.15*** (0.025)	-0.094*** (0.016)
Quantity flow	0.051*** (0.012)	0.087*** (0.011)
Quality flow	-0.0033 (0.015)	-0.013 (0.013)
Zero flow	0.18*** (0.026)	0.16*** (0.012)
Academics per year	0.0018*** (0.00026)	0.0012*** (0.00031)
Field Fis/01 (<i>sperimentale</i>)	0.0030 (0.014)	0.027** (0.013)
Field Fis/02 (<i>teorica</i>)	0.13*** (0.015)	-0.0035 (0.017)
Observations	12431	9979

The positive sign and significance of *wave 1980* and *wave 1985*, respectively, confirms that big recruitment waves increase a scientist's chances of being promoted. The overall "wave" effect has to be calculated also taking into account the positive effect of *academics*

per year. For example, one more PR promoted in a year t gives a candidate to the same position 0.48% more chances of success (0.18% for PO and 0.12% for PA).

Gender effect is strongly negative in both countries and confirms that physics is a discipline that historically penalises the careers of women.

Field influence on promotions do not give univocal results for all ranks. POs have more chances of being promoted in disciplines such as *Fisica teorica* compared to *Fisica della materia*. On the other hand PAs have more chances of being promoted in *Fisica sperimentale*. PRs have more chances of being promoted in *Milieux dilues et optique* if compared to *Milieux denses et materiaux*.

We suppose that all the non-observable individual characteristics and all the omitted variables end up being captured by the two variables *quality flow* and *quantity flow*. They measure the productivity of each scientist in terms of quantity and quality of articles over three years before t . In Italy, the selection process seems to be affected only quantity, the impact of quality being negligible for both PAs and POs¹⁰. In France both quality and quantity play a non negligible role in the chances of being promoted.

Cohort-period interactions do not exhibit a similar impact across the three equations.

Table 7. France, selection equation

COEFFICIENT	PR Marginal Effects prom
Age group 1	
Age group 3	0.38*** (0.013)
Age group 4	0.59*** (0.0092)
Age group 5	0.40*** (0.0067)
Wave 1985	0.27*** (0.013)
woman	-0.11*** (0.025)
Quantity flow	0.047*** (0.015)
Quality flow	0.073*** (0.015)
Zero flow	0.20*** (0.020)
Academics per year	0.0048*** (0.00060)
Field 30 (<i>Milieux dilues et optique</i>)	0.039*** (0.012)
Observations	11844
R-squared	.

¹⁰ Variables quality and quantity flow are highly correlated

5.2 Productivity equation

Results from selection equation estimates for Italy and France are reported in tables 8 to 11, respectively. Tables 8 and 10 refer to exercise with *quantity* as the dependent variable. Tables 9 and 11 refer to *quality*.

As expected, the *age* of academics has a negative impact on the quantity and quality of articles published. Only POs are excluded from this general trend. There are two possible explanations for the PO exception. First, the decreasing trend, observed also in other studies referred to US case, do not involve POs thanks to some unobserved features common among them and not to academics in other ranks. For example POs may have access to more resources for their research activity thanks to their prominent position. Alternatively, POs lead and manage active research teams, on whose publications they feel entitled to appear as authors.

Gender impacts differently in Italy and in France, and across academic ranks; some difference also arise when moving from *quantity* to *quality* as proxies for productivity. When quantity is considered, women appear to be less productive than men in both countries and all the positions but PO. As for quality, the negative sign persists from France, but in Italy remains only for RU. We do not have information which can help us in identifying through what channels gender exerts its effect (e.g., we do not have information on marital status or number of children). However, we notice that, in absolute terms, the impact is larger for more junior positions, such as RU and MCF. This is consistent with the possibility that gender affects more the scientists at the early stages of their career, at a time when familiar engagements may also be demanding. Thus, an auto-selection process may take place, by which only the 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.

Having productive co-authors is expected to impact positively on individual scientists' productivity. Our results confirm this hypothesis for all academic ranks in France and for POs in Italy. A one-percent increase in the nr of articles by co-authors over the three years before t , gives the Italian POs or French PRs and MCFs from 6.1% to 15% more quantity a time t . Co-authors' productivity index adjusted for quality produces a strong positive effect on the quality of individual scientists' production in both countries and all positions. Moreover it impacts positively also on quantity of publications in France but not in Italy. Even if the two variables (co-authors quantity and quality) are strictly correlated (especially because co-authors with zero quantity have necessarily zero quality) the larger and most univocal effect among ranks is played by co-authors' quality with respect to individual scientists' publication quality. In a few cases, co-authors' quantity impacts negatively on individual scientists' quality, possibly because it pushes the latter to add publications to her record, which appear on journals with lower impact factor.

Table 8. Italy, productivity equation (QUANTITY)

COEFFICIENT	PO	PO	PA	PA	RU
	OLS	Heckman	OLS	Heckman	OLS
	Log(Quantity)	Log(Quantity)	Log(Quantity)	Log(Quantity)	Log(Quantity)
Age group 1	0 (0)	0 (0)	0 (0)	0 (0)	0.045 (0.16)
Age group 3	-0.024 (0.083)	-0.015 (0.093)	-0.034 (0.028)	-0.073** (0.035)	-0.078** (0.031)
Age group 4	-0.040 (0.083)	-0.023 (0.11)	-0.11*** (0.032)	-0.17*** (0.042)	-0.27*** (0.052)
Age group 5	-0.17** (0.084)	-0.18 (0.11)	-0.21*** (0.038)	0 (0)	-0.48*** (0.11)
<i>Interactions</i>	<i>Not reported</i>				
Wave 1980	0.038* (0.021)	0.029 (0.029)	-0.062*** (0.021)	-0.12*** (0.028)	-0.22*** (0.047)
Woman	0.053 (0.038)	0.026 (0.041)	-0.032 (0.022)	-0.022 (0.024)	-0.12*** (0.021)
Co-authors quantity	0.074*** (0.020)	0.065*** (0.021)	0.013 (0.021)	0.015 (0.023)	0.014 (0.024)
Co-authors quality	0.030 (0.029)	0.0060 (0.030)	0.031 (0.029)	0.013 (0.031)	-0.016 (0.036)
Co-authors zeros	-0.18*** (0.061)	-0.21*** (0.065)	-0.23*** (0.059)	-0.25*** (0.063)	-0.30*** (0.076)
Quantity before promotion	0.51*** (0.023)	0.50*** (0.023)	0.37*** (0.020)	0.36*** (0.023)	0.45*** (0.024)
Quality before promotion	-0.100*** (0.023)	-0.093*** (0.023)	-0.016 (0.019)	0.022 (0.022)	0.0053 (0.023)
Zero before promotion	0.31*** (0.058)	0.30*** (0.059)	0.11** (0.055)	0.12** (0.063)	0.24*** (0.061)
Promotion before 1975	-0.094* (0.050)	0 (0)	0 (0)	0 (0)	0 (0)
Zeros after promotion	-0.19** (0.091)	-0.21** (0.096)	-0.29*** (0.049)	-0.34*** (0.058)	-0.35*** (0.13)
Field Fis/01 (<i>sperimentale</i>)	-0.064*** (0.021)	-0.11*** (0.022)	0.076*** (0.021)	0.055** (0.023)	0.047** (0.024)
Field Fis/02 (<i>teorica</i>)	-0.096*** (0.024)	-0.10*** (0.026)	0.047* (0.025)	0.047* (0.028)	-0.034 (0.028)
French co-authors	0.20*** (0.020)	0.16*** (0.021)	0.25*** (0.024)	0.20*** (0.027)	0.20*** (0.028)
US co-authors	0.39*** (0.022)	0.39*** (0.023)	0.53*** (0.024)	0.54*** (0.028)	0.38*** (0.028)
CNR co-authors	0.15*** (0.018)	0.14*** (0.019)	0.094*** (0.018)	0.11*** (0.020)	0.022 (0.023)
CNRS co-authors	0.22*** (0.024)	0.25*** (0.025)	0.12*** (0.028)	0.14*** (0.032)	0.15*** (0.035)
CERN co-authors	0.28*** (0.023)	0.29*** (0.024)	0.40*** (0.027)	0.47*** (0.030)	0.30*** (0.033)
INFN co-authors	0.19*** (0.021)	0.16*** (0.023)	0.17*** (0.022)	0.15*** (0.027)	0.092*** (0.026)
Lambda		-0.00060 (0.049)		-0.10** (0.045)	
Constant	0.53*** (0.11)	0.60*** (0.14)	0.46*** (0.076)	0.54*** (0.089)	0.76*** (0.12)
Observations	8332	7183	8089	6462	5106
R-squared	0.421	0.399	0.516	0.520	0.523

Table 9. Italy, productivity equation (QUALITY)

COEFFICIENT	PO	PO	PA	PA	RU
	OLS	Heckman	OLS	Heckman	OLS
	Log(Quality)	Log(Quality)	Log(Quality)	Log(Quality)	Log(Quality)
Age group 1	0 (0)	0 (0)	0 (0)	0 (0)	-0.038 (0.20)
Age group 3	0.075 (0.096)	0.14 (0.11)	-0.085** (0.037)	-0.0039 (0.047)	-0.087** (0.041)
Age group 4	-0.020 (0.096)	0.090 (0.12)	-0.21*** (0.042)	-0.094* (0.057)	-0.36*** (0.068)
Age group 5	-0.17* (0.097)	-0.073 (0.13)	-0.36*** (0.050)	0 (0)	-0.76*** (0.14)
<i>Interactions</i>	<i>Not reported</i>				
Wave 1980	0.030 (0.024)	0.047 (0.033)	-0.0068 (0.028)	0.033 (0.038)	-0.27*** (0.061)
Woman	0.042 (0.043)	-0.027 (0.047)	-0.014 (0.028)	-0.021 (0.033)	-0.12*** (0.028)
Co-authors quantity	-0.00034 (0.023)	0.0039 (0.024)	-0.055** (0.027)	-0.055* (0.031)	-0.054* (0.032)
Co-authors quality	0.26*** (0.033)	0.21*** (0.035)	0.25*** (0.038)	0.25*** (0.042)	0.20*** (0.048)
Co-authors zeros	-0.0027 (0.070)	-0.053 (0.074)	-0.047 (0.077)	-0.020 (0.085)	-0.096 (0.100)
Quantity before promotion	0.30*** (0.026)	0.33*** (0.027)	0.21*** (0.027)	0.24*** (0.031)	0.19*** (0.031)
Quality before promotion	0.13*** (0.027)	0.15*** (0.027)	0.074*** (0.025)	0.096*** (0.030)	0.11*** (0.030)
Zero before promotion	0.38*** (0.067)	0.38*** (0.068)	0.073 (0.072)	0.13 (0.085)	0.095 (0.080)
Promotion before 1975	-0.088 (0.058)	0 (0)	0 (0)	0 (0)	0 (0)
Zeros after promotion	-0.63*** (0.11)	-0.66*** (0.11)	-0.63*** (0.065)	-0.67*** (0.078)	-0.83*** (0.17)
Field Fis/01 (<i>sperimentale</i>)	-0.12*** (0.024)	-0.15*** (0.025)	-0.019 (0.027)	-0.015 (0.031)	-0.074** (0.031)
Field Fis/02 (<i>teorica</i>)	-0.021 (0.027)	0.0050 (0.029)	0.034 (0.033)	0.018 (0.038)	0.040 (0.037)
French co-authors	0.15*** (0.023)	0.12*** (0.025)	0.28*** (0.031)	0.27*** (0.036)	0.11*** (0.037)
US co-authors	0.19*** (0.025)	0.20*** (0.026)	0.36*** (0.031)	0.31*** (0.038)	0.24*** (0.036)
CNR co-authors	0.068*** (0.021)	0.058** (0.022)	0.093*** (0.024)	0.10*** (0.027)	0.069** (0.031)
CNRS co-authors	0.047* (0.028)	0.050* (0.029)	-0.038 (0.037)	-0.042 (0.043)	0.040 (0.046)
CERN co-authors	0.20*** (0.026)	0.19*** (0.027)	0.23*** (0.036)	0.28*** (0.041)	0.22*** (0.043)
INFN co-authors	0.15*** (0.025)	0.13*** (0.026)	0.27*** (0.029)	0.20*** (0.036)	0.17*** (0.035)
Lambda		0.058 (0.056)		0.16*** (0.061)	
Constant	0.51*** (0.13)	0.44*** (0.16)	0.62*** (0.10)	0.42*** (0.12)	1.02*** (0.15)
Observations	8332	7183	8089	6462	5106
R-squared	0.325	0.296	0.383	0.364	0.377

Table 10. France, productivity equation (QUANTITY)

COEFFICIENT	PR	PR	MCF
	OLS Log(Quantity)	Heckman Log(Quantity)	OLS Log(Quantity)
Age group 1	0 (0)	0 (0)	0.10** (0.044)
Age group 3	-0.075** (0.031)	-0.20*** (0.059)	-0.080*** (0.020)
Age group 4	-0.14*** (0.033)	-0.38*** (0.083)	-0.14*** (0.026)
Age group 5	-0.27*** (0.041)	-0.58*** (0.094)	-0.13*** (0.037)
<i>interactions</i>	<i>Not reported</i>		
Wave 1985	0.10*** (0.021)	0.031 (0.027)	-0.034** (0.016)
Woman	-0.14*** (0.030)	-0.16*** (0.036)	-0.067*** (0.012)
Co-authors quantity	0.061*** (0.019)	0.073*** (0.022)	0.15*** (0.015)
Co-authors quality	0.061** (0.026)	0.061** (0.031)	0.022 (0.021)
Co-authors zeros	-0.017 (0.055)	0.0027 (0.063)	0.083* (0.044)
Quantity before promotion	0.49*** (0.020)	0.50*** (0.023)	0.35*** (0.017)
Quality before promotion	0.081*** (0.021)	0.084*** (0.024)	0.055*** (0.012)
Zero before promotion	0.48*** (0.046)	0.46*** (0.055)	0.27*** (0.026)
Promotion before 1975	-0.090* (0.052)	0 (0)	0 (0)
Zeros after promotion	-0.27*** (0.040)	-0.30*** (0.050)	-0.19*** (0.019)
Field 30 (<i>Milieux dilues et optique</i>)	-0.013 (0.015)	-0.029* (0.017)	0.048*** (0.011)
Italian co-authors	0.24*** (0.026)	0.26*** (0.031)	0.38*** (0.022)
US co-authors	0.31*** (0.022)	0.31*** (0.027)	0.27*** (0.019)
CNRS co-authors	0.18*** (0.017)	0.18*** (0.019)	0.18*** (0.013)
UMR co-authors	0.068*** (0.020)	0.068*** (0.024)	0.057*** (0.016)
lambda		-0.22*** (0.064)	
Constant	0.18** (0.075)	0.43*** (0.13)	-0.055 (0.057)
Observations	9018	6866	12057
R-squared	0.346	0.345	0.322

Table 11. France, productivity equation (QUALITY)

COEFFICIENT	PR	PR	MCF
	OLS Log(Quality)	Heckman Log(Quality)	OLS Log(Quality)
Age group 1	0 (0)	0 (0)	0.16** (0.070)
Age group 3	-0.099** (0.042)	-0.11 (0.079)	-0.11*** (0.032)
Age group 4	-0.21*** (0.045)	-0.27** (0.11)	-0.18*** (0.042)
Age group 5	-0.36*** (0.055)	-0.47*** (0.13)	-0.18*** (0.059)
<i>interactions</i>	<i>Not reported</i>		
Wave 1985	0.16*** (0.029)	0.11*** (0.036)	-0.0015 (0.026)
Woman	-0.19*** (0.040)	-0.22*** (0.048)	-0.10*** (0.019)
Co-authors quantity	-0.062** (0.026)	-0.021 (0.029)	0.065*** (0.025)
Co-authors quality	0.18*** (0.036)	0.16*** (0.041)	0.18*** (0.034)
Co-authors zeros	0.056 (0.074)	0.098 (0.084)	0.22*** (0.071)
Quantity before promotion	0.43*** (0.027)	0.48*** (0.030)	0.21*** (0.027)
Quality before promotion	0.34*** (0.028)	0.38*** (0.032)	0.13*** (0.019)
Zero before promotion	0.68*** (0.063)	0.72*** (0.073)	0.16*** (0.042)
Promotion before 1975	0.17** (0.070)	0 (0)	0 (0)
Zeros after promotion	-0.68*** (0.054)	-0.70*** (0.066)	-0.49*** (0.030)
Field 30 (<i>Milieux dilues et optique</i>)	0.013 (0.020)	-0.0070 (0.023)	0.070*** (0.017)
Italian co-authors	0.14*** (0.035)	0.11*** (0.041)	0.36*** (0.035)
US co-authors	0.26*** (0.029)	0.24*** (0.036)	0.23*** (0.030)
CNRS co-authors	0.25*** (0.023)	0.25*** (0.025)	0.31*** (0.022)
UMR co-authors	0.084*** (0.027)	0.069** (0.032)	0.093*** (0.025)
lambda		-0.053 (0.085)	
Constant	0.081 (0.10)	0.022 (0.18)	0.053 (0.092)
Observations	9018	6866	12057
R-squared	0.295	0.296	0.264

As explained in the previous section, the way chosen to correct for the individual unobservable features is to consider as fixed effect the average productivity until the year of nomination. This, as expected, has a positive effect: the higher the productivity before promotion, the higher the productivity after, both in terms of quality and quantity. Theoretically a mix of *Matthew Effect*, commitment of the scientist and ability in research supports this result. The only noteworthy and counter-intuitive exception is the negative impact for POs of quality of the scientific production before promotion on the number of publications after promotion.

Of particular interest to our research is the effect of “recruitment waves” on productivity. The sign and significance of variables *wave 1980* and *wave 1985* suggest that this is the case for MCFs in France, and RUs and PAs in Italy. The effect is stronger for quantity than for quality. However, the sign is opposite than expected for French PRs and Italian POs. Bearing in mind that we do not observe how many scientists entered as RUs or MCFs in 1980 and 1985 have then been promoted to higher positions (because they either left the university in the meanwhile, or because they have been further promoted, in the case of Italy, to PO), we propose the following interpretation of our results. Big recruitment waves in both countries fill up the lower positions in the academic system of less productive scientists, who do not make further progress in their careers, and generates the negative signs we observe for RUs and MCFs. Some statistics are useful to reinforce this interpretation: no less than 26.8% (166/618) of PAs and 14.28% (73/511) of RUs active in 2005 in Italy were recruited in 1980 (*law 382/1980*) and 9.9% (121/1212) of MCFs were recruited in 1985 in France (*1984 Higher Education Act*); this figures suggest that their influence on our estimates may be decisive.

Another point is about the affiliation of co-authors. Generally speaking having foreign co-authors or co-authors belonging to PROs have a positive impact on the two dimensions of productivity. One of the best ways to fuel productivity is to have had in the past at least one US co-author. Models give this results clearly and strongly, for every step of career. For Italian academics also having co-authors from CERN, INFN and CNRS has a big and positive impact. Less crucial are co-authorship with CNR, coming into sight the marginal role played by the Italian PRO if compared to the French CNRS. In France having a co-author from UMR have a positive impact on quantity and quality of scientific productivity but less pronounced than US and CNRS.

Interactions do not evidence any relevant group of cohort-period couples to explain productivity. Heckman procedure estimates a coefficient λ to correct for the problem of selection bias. The negative coefficient has to be interpreted as: the higher the chances on not being promoted the lower the productivity. Sometimes the coefficient is not significant and the two equations, selection and productivity, can be estimated separately without any kind of correction. The exceptional case is the λ estimated with positive sign in the model for quality of scientific production for PAs. The interpretation is quite striking: the higher the chances of not being promoted, the higher the average quality of the scientific production of scientists *i*.

Finally, we observe that controls for fields of research are also significant.

6. Conclusions

To be added later

References

- [1] A.J. Lotka. The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, 16(12):317–323, 1926.
- [2] J.M. Cattell. Statistics of American Psychologists. *American Journal of Psychology*, 14:310–328, 1903.
- [3] B. Godin. On the origins of bibliometrics. *Scientometrics*, 68(1):109–133, 2006.
- [4] B. Godin. From Eugenics to Scientometrics: Galton, Cattell, and Men of Science. *Social Studies of Science*, 37(5):691, 2007.
- [5] R.K. Merton. The Matthew Effect in Science. *Science*, 159(3810):56–63, 1968.
- [6] R.K. Merton. The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property. *Isis*, 79(4):606, 1988.
- [7] P.A. David. *Positive Feedbacks and Research Productivity in Science: Reopening Another Black Box*. MERIT, 1993.
- [8] L.H. Summers. Remarks at NBER Conference on Diversifying the Science & Engineering Workforce. Retrieved September, 8:2005, 2005.
- [9] E.S. Spelke. Sex Differences in Intrinsic Aptitude for Mathematics and Science?: A Critical Review. *AMERICAN PSYCHOLOGIST*, 60(9):950, 2005.
- [10] H. Etzkowitz, C. Kemelgor, and B. Uzzi. *Athena Unbound: The Advancement of Women in Science and Technology*. Cambridge University Press, 2000.
- [11] M.F. Fox. Gender, Hierarchy, and Science. *Handbook of the Sociology of Gender*, pages 441–457, 1999.
- [12] H.D. Graham and N. Diamond. *The Rise of American Research Universities: Elites and Challengers in the Postwar Era*. Johns Hopkins University Press, 1997.
- [13] S.G. Levin and P.E. Stephan. Research productivity over the life cycle: evidence for academic scientists. *American Economic Review*, 81(1):114–132, 1991.
- [14] J.S. Long and M.F. Fox. Scientific Careers: Universalism and Particularism. *Annual Reviews in Sociology*, 21(1):45–71, 1995.
- [15] J. Ben-David. *Centers of Learning: Britain, France, Germany, United States*. Transaction Publishers, 1992.
- [16] B.R. Clark. *The Research Foundations of Graduate Education: Germany, Britain, France, United States, Japan*. 1993.
- [17] R.G. EHRENBERG, H. KASPER, and D.I. REES. Faculty Turnover at American Colleges and Universities: Analysis of AAUP Data.
- [18] C. Musselin. *Le marché des universitaires: France, Allemagne, États-Unis*. Presses de Sciences Po, 2005.
- [19] R. Moscati. Italian university professors in transition. *Higher Education*, 41(1):103–129, 2001.
- [20] T. Chevallier. French academics: Between the professions and the civil service. *Higher Education*, 41(1):49–75, 2001.

- [21] A. Bonaccorsi and C. Daraio. Age effects in scientific productivity. *Scientometrics*, 58(1):49–90, 2003.
- [22] J. Heckman. Sample Selection Bias as a Specification Error. *Econometrica*, 47(1):153–161, 1979.
- [23] J.A. Dubin and D. Rivers. Selection Bias in Linear Regression, Logit and Probit Models. *Sociological Methods and Research*, 18(2):360–90, 1989.
- [24] W.H. Greene. *FIML Estimation of Sample Selection Models for Count Data*. New York University, Leonard N. Stern School of Business, 1997.
- [25] B.H. HALL, J. MAIRESSE, and L. TURNER. Identifying Age, Cohort and Period Effects in Scientific Research Productivity: Discussion and Illustration Using Simulated and Actual Data on French Physicists. *NBER Working Paper*, 2005.
- [26] J.S. Long. Rank Advancement in Academic Careers: Sex Differences and the Effects of Productivity. *American Sociological Review*, 58(5):703–22, 1993.
- [27] DJ Price and DD Beaver. Collaboration in an invisible college. *Am Psychol*, 1966:1011–8, 1911.
- [28] D. Crane. *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*. University of Chicago Press, 1972.
- [29] P.D. Allison and J.A. Stewart. Productivity differences among scientists: evidence for accumulative advantage. *American Sociological Review*, 39(4):596–606, 1974.
- [30] JR Cole and H. Zuckerman. Marriage, motherhood and research performance in science. *Sci Am*, 256(2):119–25, 1987.
- [31] M.F. Fox. Gender, Family Characteristics, and Publication Productivity among Scientists. *Social Studies of Science*, 35(1):131, 2005.