Via Po, 53 – 10124 Torino (Italy)
Tel. (+39) 011 6702704 - Fax (+39) 011 6702762
URL: http//www.de.unito.it

WORKING PAPER SERIES

THE PRODUCTION OF SCIENTIFIC KNOWLEDGE IN ITALY: EVIDENCE IN THEORETICAL, APPLIED AND TECHNICAL SCIENCES

Pier Paolo Patrucco

Dipartimento di Economia "S. Cognetti de Martiis"

Laboratorio di Economia dell'Innovazione "Franco Momigliano"

Working paper No. 01/2006

Università di Torino

THE PRODUCTION OF SCIENTIFIC KNOWLEDGE IN ITALY: EVIDENCE IN THEORETICAL, APPLIED AND TECHNICAL SCIENCES[1]

Pier Paolo Patrucco
Robert Schuman Centre for Advanced Studies (RSCAS)
European University Institute
Via delle Fontanelle 20
50016 San Domenico di Fiesole (FI)

and

Laboratorio di Economia dell'Innovazione "Franco Momigliano"
Dipartimento di Economia
Università di Torino
Via Po 53 – 10124 Torino
Tel. +39 011 6702767
Fax. +39 011 6702762
pierpaolo.patrucco@unito.it

ABSTRACT. The paper presents preliminary empirical evidence on the production of scientific knowledge in Italy, in theoretical sciences (physics), applied sciences (chemistry) and technical sciences (engineering and petrology). It elaborates on an original dataset of publications and citations for 2,673 Italian researchers, distributed across 61 universities, covering the years between 1990 and 2004. According to a well-established tradition of studies in the economics of science, the results show that individual distribution is quite asymmetric, with very few researchers accounting for a great amount of scientific output. More interestingly, the paper also shows that there are important differences in terms of asymmetric distribution when the different disciplines, universities and academic positions of the researchers are compared. These differences open the way to interpretation in terms of two main factors. Firstly, the various disciplines can be characterised by specific knowledge bases, learning practices, organisation of scientific labour, and communication norms. Secondly, specific weaknesses in the hiring, incentive and monitoring schemes at discipline and university level can explain different degrees of asymmetry. Both these factors have important implications for a research agenda on the governance of science. Finally, the paper shows that, at the aggregate level, scientific productivity benefit from a concentration of R&D expenditures only to a minor extent, and subject to decreasing returns. The effect of externalities stemming from R&D investments is limited. The scope of the concentration of R&D resources can therefore be questioned.

---

[1] The comments of Cristiano Antonelli, Aldo Geuna, Laurel Smith-Doerr and Rikard Stankiewicz are gratefully acknowledged, as well as the financial support of the European Commission through a Jean Monnet Fellowship at the Robert Schuman Centre for Advanced Studies (RSCAS) of the European University Institute in Fiesole and the support of Cristina Panozzo and Stefania Alifredi for data collection. The work has been made possible also by the funding of the European Union Directorate for Research, within the context of the Integrated Project EURODITE (Regional Trajectories to the Knowledge Economy: A Dynamic Model) Contract nr° 006187 (CIT3), in progress at the Fondazione Rosselli, as well as the funding of the Fondazione CRT through the Lagrange Project "Un'Analisi Comparata delle Performance Tecnologiche del Nord Ovest Sabaudo nel Lungo Periodo e nel Contesto nelle Regioni Italiane: Gli Anni 1980-2001", carried out at the Fondazione Rosselli.

1. INTRODUCTION

The fact that the production of scientific knowledge is very asymmetric and skewed across both individuals and institutions is one of the few consolidated results in the economics and sociology of science. The distribution of scientific production across individuals is by no means normal, but it is well described by a Pareto distribution, where the largest proportion of output is accounted for by a relatively small number of very productive researchers (Lotka, 1926; Katz, 1999; Merton, 1968). At the institutional level, scientific production and productivity are often skewed according to the role played by 'stars', rather than, for instance, by the scale of the research institutions (Zucker et al., 1998).

In this context, the paper elaborates upon an original dataset of individual publications and citations of 2,673 Italian researchers in theoretical, applied and technical sciences, distributed across 61 universities. The object of this paper is to provide empirical evidence on the characteristics of the individual distribution of scientific production according to the different disciplines, universities and academic positions of the single researchers.

More specifically, the paper compares the individual distribution of scientific production across disciplines, universities and positions, in order to: 1) show to what extent scientific production is asymmetric or not; 2) investigate whether the various disciplines, universities and individual positions are characterised by different levels of asymmetry, or whether, on the contrary, the variance across disciplines, universities and positions is limited; 3) reveal the differences in terms of a comparison between the proportion of very productive researchers and that compared to the proportion of less productive ones.

In this sense, the novelty of the paper lies in distinguishing the individual distribution of scientific production in terms of disciplines, universities and academic positions. In cases where the differences in terms of asymmetric distribution of scientific output across disciplines, universities and positions are great, these differences can be explained on the basis of two main factors. Firstly, discipline specificities can matter, especially in terms of idiosyncratic characteristics of the knowledge base, learning practices, models of organisation of scientific work, and communication patterns that distinguish the various scientific fields. Secondly, differences can be explained in terms of more or less effective (or more or less weak) hiring, incentive and monitoring schemes at discipline and university level: the more asymmetric the distribution of scientific knowledge, the weaker those schemes. The implications in terms of a research agenda on the governance of science may be relevant to both these factors. In particular, specific governance models should be identified according to the specificities shown by the various disciplines and academic positions.

Moreover, the analysis also compares (at the regional level) the amount of R&D expenditure with the number of publications, investigating to what extent scientific activity may benefit from the concentration of R&D resources in a given area.

The paper is organised as follows. Section 2 presents the empirical dataset and the methodology on which the paper is based. Section 3 shows the empirical findings on the characteristics of the distribution of scientific production, comparing the different distributions of scientific production by disciplines, universities and academic positions. It also correlates, at the regional level, scientific output in terms of publications with the amount of R&D investment. Section 4 puts forward an interpretative framework for providing a preliminary explanation of the empirical evidence (this explanation can also be read as a research agenda on the implications for the governance of science). The conclusions summarise the main results and place them in the perspective of future research.

2. DATASET AND METHODOLOGY

The paper elaborates upon an original database made of 2,673 Italian researchers (assistant, associate and full professors), distributed across 61 universities, active in the fields of chemistry (Physical chemistry; General and inorganic chemistry; Organic chemistry), engineering (Metallurgy; Material engineering; Electronics measurement), earth sciences (Petrology) and physics (Theoretical physics). These scientific fields were chosen in the first place because in these fields, with the exception of physics, Italy's scientific impact is higher than the European average (CRUI, 2002); in this regard, they can be thought of as 'best practices' or 'scientific champions' in Italian science. In the second place, they are a very good instance of the traditional classification between theoretical science (Physics), applied science (Chemistry), and technical or technology-oriented science (Petrology and Engineering)[2], characterised by codified, articulable and tacit knowledge bases respectively[3]. Here it may be interesting to compare the different fields, examining their different characteristics, and if possible identifying some specificities, on the basis of the knowledge base underlying each of the various fields.

This database was implemented using data from the Italian Ministry of University, Research and Technology (MIUR), and the 2,673 researchers included in the database represent the universe of the researchers in those fields. For each researcher, the database provides information on their position [1) tenured assistant professors; 2) assistant professors without tenure; 3) tenured associate professors; 4) associate professors without tenure; 5) tenured full professors; 6) full professors without tenure][4], the institution in which they are employed (university level), and the region in which the university is located.

This dataset was subsequently integrated with the number of publications and citations received by each researcher on ISI journals. In order to analyse scientific production in the various fields, two indicators for output are used, namely the number of publications cited by articles published in ISI journals, and the number of citations received by these publications. The number of publications and citations refers to the period 1990-2004, and is based on the Science Citation Index elaborated by the ISI[5].

---

[2] This classification has been confirmed by colloquia with researchers in the fields. However – to anticipate one of the results of the paper – our simple descriptive statistics on the distribution of publications will show a relatively different picture, in which physics and chemistry are more similar to each other and quite distinct from the technological fields (i.e., engineering and petrology). In this respect, the idea of considering chemistry an applied science, which is indeed common also among scientists in the field, should be discussed in future research, with an in-depth analysis of the nature of the different disciplines.

[3] For this classification, see for instance the work by Cowan, David and Foray (2000).

[4] In the Italian academic system, academic positions are 'allocated' by local competition. Once you win the local competition, you are given the position of, for instance, associate professor, *but* without tenure. After 3 years, you go through a process of local and national evaluation of your teaching and publishing activity during that period. If the response of this evaluation is favourable, you get tenure. In principle, there is quite a strong difference, in terms of power and resources, between assistant professors on the one hand, and associate and full professors on the other.

[5] We must point out that the database on publications does not take multiple authorship into account. The importance of multiple authorship may vary, depending on the scientific fields, and may explain some of the

By means of, the simple count of publications and citations per researcher and of the cumulative percentages of publications, citations and researchers, the empirical analysis will reveal the characteristics of scientific production in the various scientific fields, universities and academic positions. Moreover, it will show whether there are differences across fields, universities and academic positions, in terms of asymmetry in the distribution and proportion of the more productive researchers, when compared with the proportion of the less productive ones.

Table 1 shows the distribution of researchers, publications and citations across the different scientific fields.

Here the broadest fields are those of chemistry, which represent 25%, 23% and 18.5% of the population of researchers respectively. The engineering sectors are comparatively the narrowest, covering 8%, 4% and 3% of the total researchers. These relative weights are, quite clearly, reflected also in the shares of publications and citations. More interestingly, when we consider simple productivity measures such as the number of publications per researcher, we find that researchers in physics are the most productive, with 62 publications per researcher. Finally, as regards the citation impact (number of citations/number of publications), again chemistry, and in particular the field of general and inorganic chemistry, is the one with the highest impact (6.79).

Table 1. Distribution of researchers, publications and citations across the various scientific fields

Scientific field	**No. res**	**% res**	**No. publ**	**% publ**	**No. cit**	**% cit**	**Publ/ res**	**Cit/ res**	**Cit/publ (citation impact)**
Metallurgy	88	3.29	2,365	1.68	9,874	1.17	26.88	112.20	4.18
Material engineering	220	8.23	7,375	5.24	31,680	3.74	33.52	144.00	4.30
Electronic measurement	108	4.04	1,890	1.34	4,949	0.58	17.50	45.82	2.62
Petrology	117	4.38	3,889	2.76	18,860	2.23	33.24	161.20	4.85
Physical chemistry	493	18.44	27,715	19.70	157,915	18.63	56.22	320.31	5.70
General and inorganic chemistry	621	23.23	37,598	26.73	255,292	30.12	60.54	411.10	6.79
Organic chemistry	670	25.07	37,871	26.92	234,397	27.66	56.52	349.85	6.19
Theoretical physics	356	13.32	21,956	15.61	134,494	15.87	61.67	377.79	6.13
TOTAL	2,673	100.00	140,659	100.00	847,461	100.00	52.62	317.04	6.02

differences in scientific production across fields. One of the next steps of this research will be the analysis of multiple authorship at least for a sample of researchers across disciplines.

Table 2. Distribution of researchers, publications and citations across the Italian regions

Region	No. res	% res	No. publ	% publ	No. cit	% cit	Publ/ res	Cit/ res	Cit/publ (citation impact)
EMILIA-ROMAGNA	376	14.07	22,167	15.76	149,955	17.69	58.95	398.82	6.76
LOMBARDY	346	12.94	18,634	13.25	116,200	13.71	53.86	335.84	6.24
TUSCANY	248	9.28	15,513	11.03	100,052	11.81	62.55	403.44	6.45
LATIUM	256	9.58	13,229	9.41	77,500	9.14	51.68	302.73	5.86
CAMPANIA	232	8.68	11,050	7.86	63,558	7.50	47.63	273.96	5.75
VENETO	182	6.81	9,483	6.74	55,365	6.53	52.10	304.20	5.84
PIEDMONT	**152**	**5.69**	**9,143**	**6.50**	**55,815**	**6.59**	**60.15**	**367.20**	**6.10**
SICILY	207	7.74	7,748	5.51	45,357	5.35	37.43	219.12	5.85
APULIA	119	4.45	5,460	3.88	25,164	2.97	45.88	211.46	4.61
FRIULI	89	3.33	5,436	3.86	34,765	4.10	61.08	390.62	6.40
UMBRIA	75	2.81	4,823	3.43	25,439	3.00	64.31	339.19	5.27
SARDINIA	97	3.63	4,351	3.09	25,739	3.04	44.86	265.35	5.92
LIGURIA	80	2.99	3,621	2.57	16,236	1.92	45.26	202.95	4.48
MARCHE	63	2.36	3,396	2.41	16,752	1.98	53.90	265.90	4.93
CALABRIA	54	2.02	2,116	1.50	11,115	1.31	39.19	205.83	5.25
TRENTINO-A.A.	30	1.12	1,709	1.21	10,775	1.27	56.97	359.17	6.30
ABRUZZO	32	1.20	1,656	1.18	10,505	1.24	51.75	328.28	6.34
BASILICATA	28	1.05	887	0.63	5,771	0.68	31.68	206.11	6.51
MOLISE	7	0.26	237	0.17	1,398	0.16	33.86	199.71	5.90
TOTAL	2,673	100.00	140,659	100.00	847,461	100.00	52.62	317.04	6.02

Table 2 shows the distribution of researchers, publications and citations by region. Emilia-Romagna is the region with the largest proportion of researchers (14%), publications (15.76%) and citations (17.69%), and also with the highest citation impact (6.76). in terms of scientific productivity, that is to say the ratio of publications to researchers. Tuscany and Piedmont are, among the largest regions, the most productive ones, with 62.55 and 60 publications per researcher respectively.

Table 3 shows some descriptive statistics concerning publications and citations at the individual level, aggregated at the university and regional level.

Table 3. Descriptive statistics on scientific production

Variable	Publications Individual Tot. No.	Publications University Tot. No.	Publications Region Tot. No.	Citations Individual Tot. No.	Citations University Tot. No.	Citations Region Tot. No.
Total	140,659	140,659	140,659	847,461	847,461	847,461
No. observations	2,673	61	19	2,673	61	19
min	0	4	237	0	10	1,398
max	752	11,259	22,167	5,791	73,525	149,955
mean	53	2,306	7,403	317	13,893	44,603
st. dev	60	2,644	6,262	453	16,894	41,207
mean/st. dev	1.14	1.15	0.85	1.43	1.22	0.92

Each researcher published an average of 53 works, receiving 317 citations. However, there is a wide range of variations in individual publications and citations, since the most productive researcher published 752 works and received 5,791 citations, while the least productive one had zero publications and citations. Given this wide range, it is obvious that the standard deviation values are very high for both publications and citations. In both cases, the standard deviation is higher than the mean values, preliminarily showing that the distribution of both publications and citations is very dispersed and asymmetric. The next section will deal precisely with the characteristics of this individual distribution, showing to what extent scientific production is asymmetric, and ascertaining whether there are any differences in this asymmetry when the individual distributions are compared by disciplines, universities and academic positions.

3. EMPIRICAL EVIDENCE ON SCIENTIFIC PRODUCTION

The high variability and asymmetry of scientific production is quite clear in Figure 1, which shows, at the aggregate level, the distribution of the total number of publications and citations across researchers. Figure 1 shows that the distribution of both publications and citations is much more similar to a Pareto distribution than to a normal one, with very few researchers who are very productive, and a long queue of relatively less productive ones.

Figure 1. Individual distribution of publications and citations

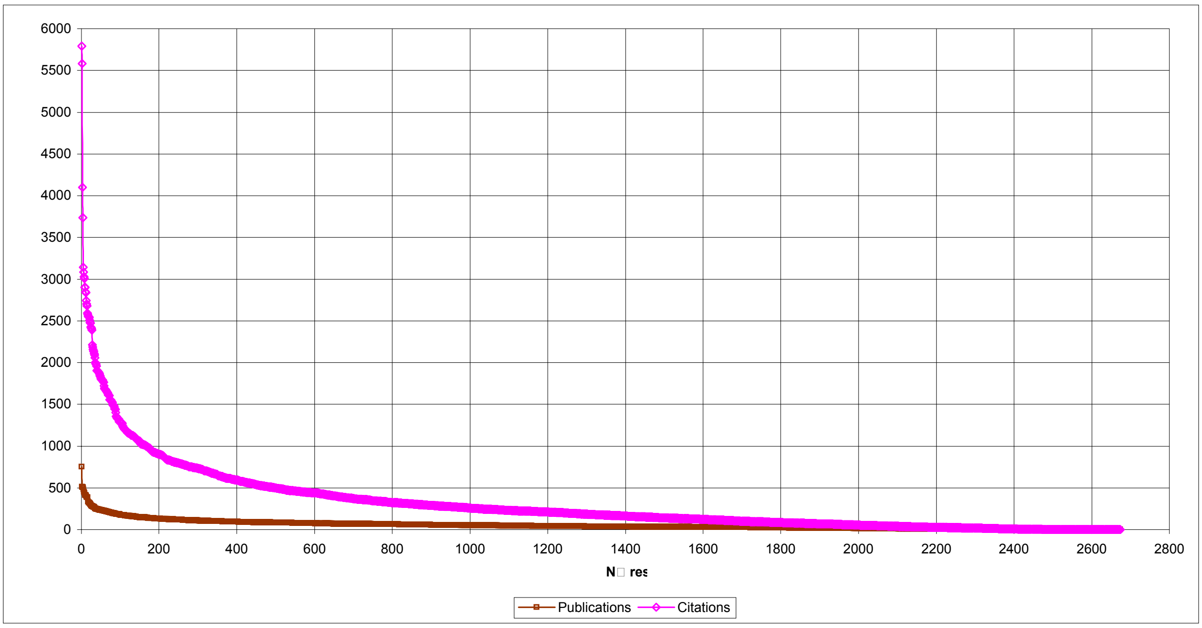


Figure 2. Individual distribution of publications and citations (cumulated)

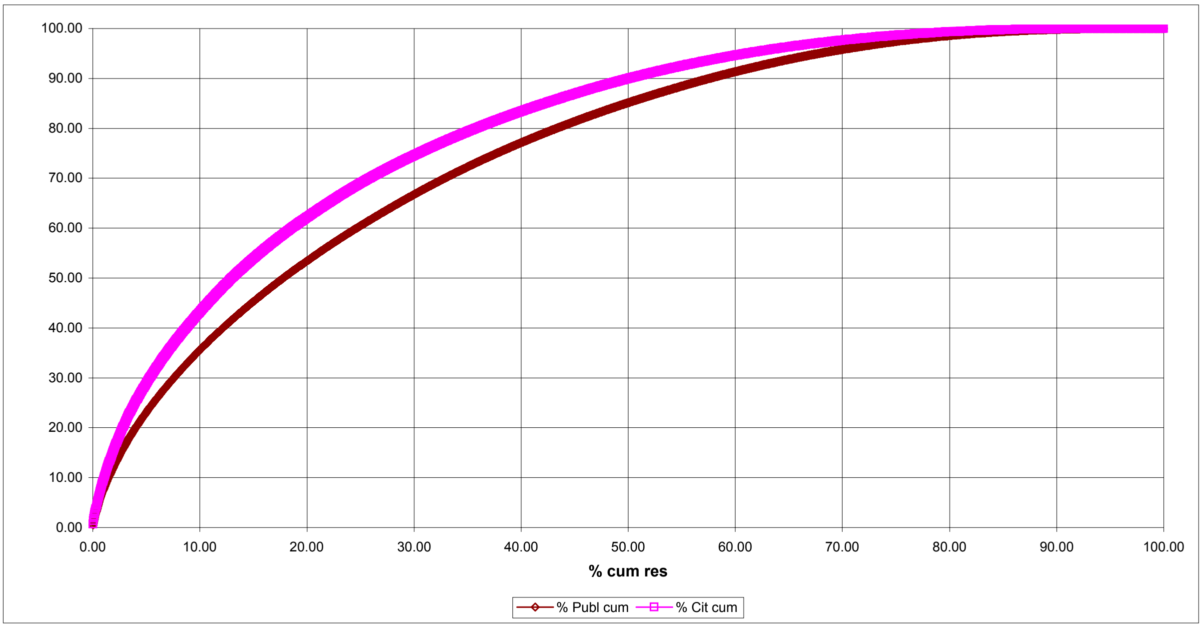


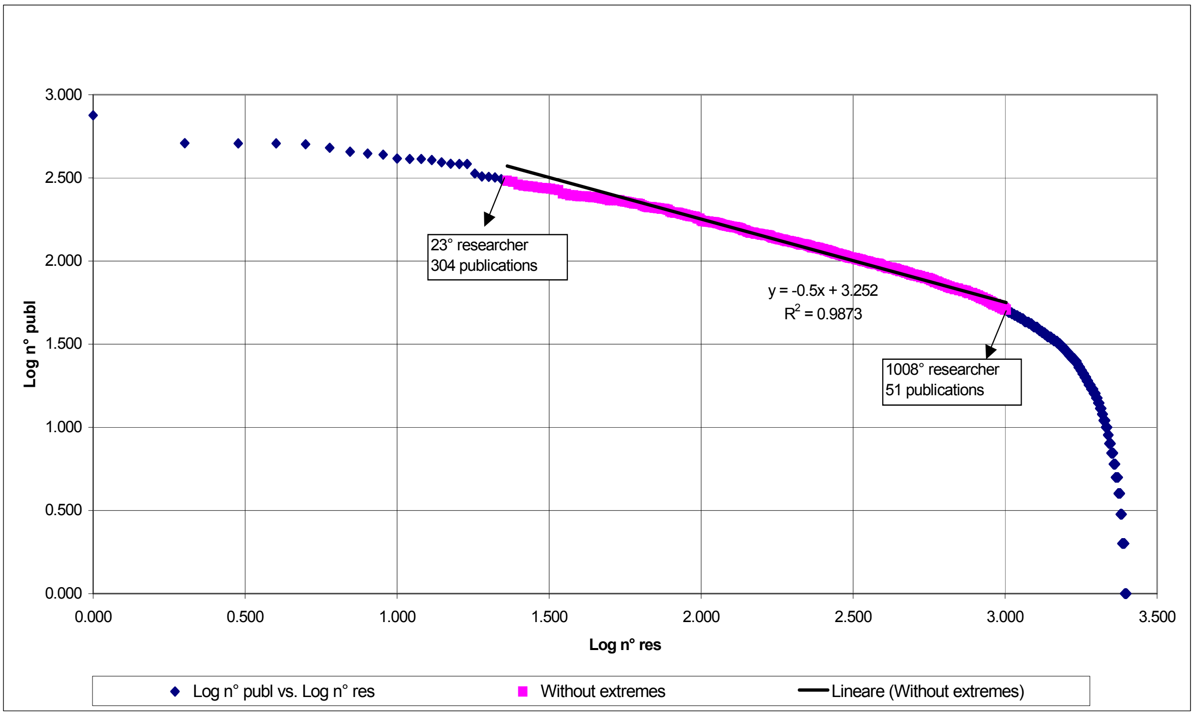
Figure 2 shows the distribution of cumulative publications and citations relating to the cumulative number of researchers. We can see that 62.33% of total citations and 54% of publications are concentrated on 20% of researchers, confirming the uneven distribution of both publications and citations across researchers.

The notion of an uneven distribution of scientific production is very familiar to scholars of economics of science. It has been shown that individual publications, as well as citations, are distributed very asymmetrically across the aggregate population of Italian researchers. Scientific production can be represented quite well by Lotka's law (Lotka, 1926), whereby, within a homogeneous group of researchers, the number of publications of a given researcher can be represented as a rapidly decreasing function of the number of researchers, as shown by the following equation:

$$\text{No. PUBL}_i = k/i^{0.5} \qquad (1)$$

where *k* is the number of publications of the most productive researcher, and *i* is the progressive number of the researcher in a list by decreasing number of publications[6]. Figure 3 shows the logarithmic distribution of the number of publications according to the logarithm of the researchers. If the data were fitted into Lotka's law, we would have a linear distribution with inclination -0.5.

Figure 3. Scientific production and Lotka's law



It is clear that the distribution is not linear, in that it decreases less than proportionally before a given threshold (log i = 1.342) and it starts decreasing more than proportionally after a given threshold (namely, after log i = 3.000). The distribution could be considered linear with an inclination of –0.5 only for researchers between the 23rd and the 1008th position.

---

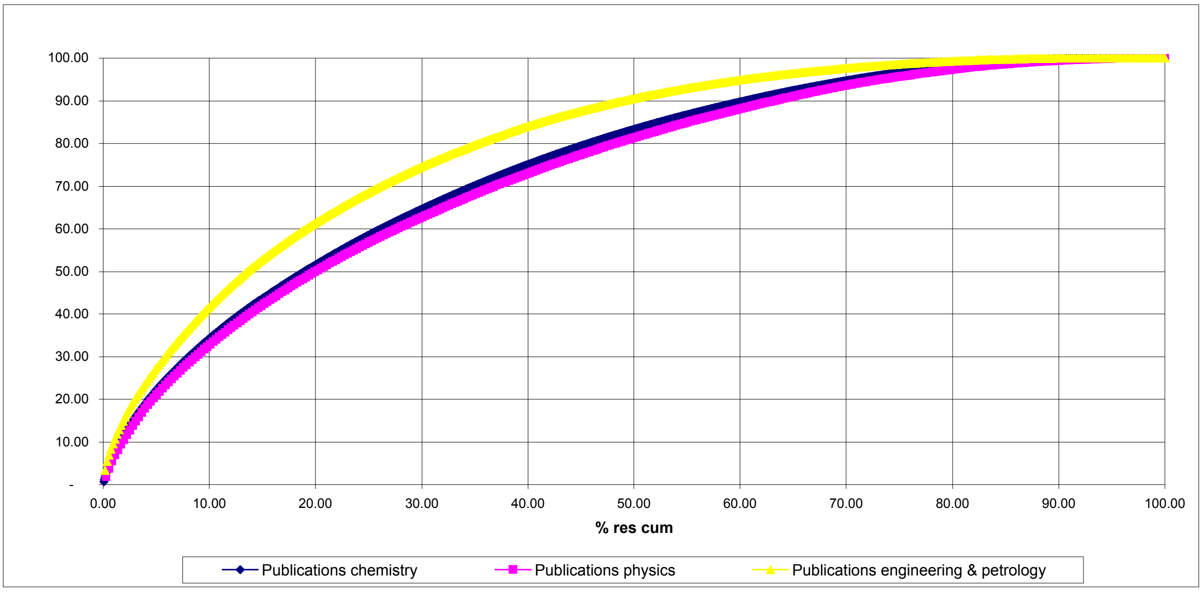
[6] Lotka's law describes the frequency of publication by authors in a given field. It states that the number of authors making n contributions is about $1/n^2$ of those making one; and the proportion of all contributors who make a single contribution is about 60%. This means that out of all the authors in a given field, 60% will have only one publication, 15% will have two publications, 7% will have three publications, and so on. According to Lotka's law of scientific productivity, only 6 % of the authors in a field will produce more than 10 articles. Lotka's law, when applied to large bodies of literature over a fairly long period of time, can be generally accurate, but not statistically exact.

In this context, a deeper analysis that distinguishes between disciplines, universities and the academic positions of the researchers may be useful in allowing us to find out whether at these levels the asymmetric distribution contains major differences and specificities. The next subsection will compare the distribution of individual scientific production in terms of disciplines, universities and positions.

3.1. Comparison of individual scientific production by discipline, university and academic position

In terms of a comparison across disciplines (Figure 4), a simple analysis of the distribution of cumulated publications shows, in the first place, that the distributions in chemistry and physics are much more similar to each other (actually almost overlapping) than the distributions in engineering and petrology. Since physics is much more a theoretical science, and chemistry, engineering and petrology are more applied and technical disciplines, we might have expected different results; namely a more similar distribution between chemistry and engineering.

Figure 4. Individual distribution of publications by discipline



In the second place, this comparison also shows that in engineering and petrology the distribution is more asymmetric than in physics and chemistry. When we compare the proportions of the most productive researchers (or ‘stars’) (i.e., the researchers who account for the first 10% of publications) with the proportion of the least productive researchers (i.e., those accounting for the last 10% of publications) across the three disciplines, we can detect some sectoral specificities.

On the one hand, the share of the ‘stars’ is very similar and very small across the three fields. In fact, 1% of the researchers and 2% of the researchers account for the first 10% of publications in engineering, and in physics and chemistry, respectively. Even using more relaxed criteria for identifying very productive researchers, i.e. considering those researchers who account for the first 20% of publications, the picture does not change very much, about 3% and 5% of the researchers accounting for the first 20% of publications in engineering, and chemistry and physics, respectively. On the other hand, the share of the least productive researchers is characterised by a certain degree of variance. If in chemistry and physics this share is quite similar (37% in physics and about 40% in chemistry), in engineering and petrology more than half (51%) of the researchers account for the last 10% of publications. Sectoral specificities in turn account for a variance of about 14% of poorly productive researchers. These sectoral specificities may be interpreted as differences in the knowledge base characterising the different disciplines, and therefore in the organisation of scientific work and production of knowledge.

The differences in the distribution of individual publications are found to be even higher in a comparison between various universities (Figure 5).

Figure 5. Individual distribution of publications by university

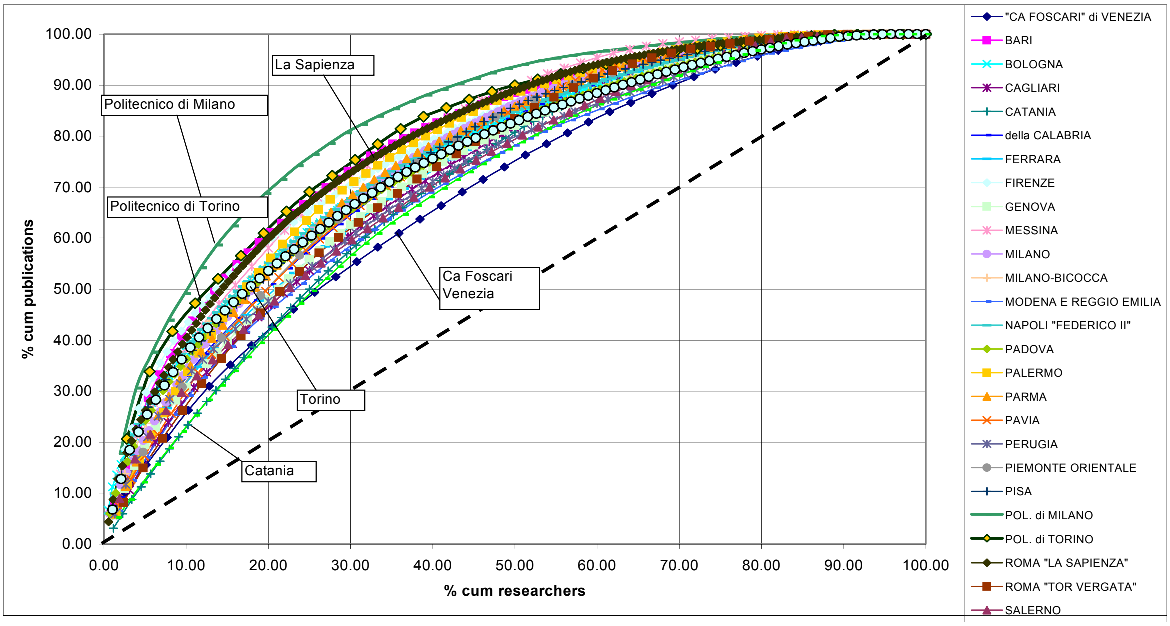


Figure 5 shows that, in general terms, the universities in which scientific production is most asymmetric are the two polytechnic schools in Milan, and the one where it is relatively least asymmetric is Turin.. This may be partly due to the fact that the two polytechnic schools are specialised in engineering, which, as Figure 4 showed, is more asymmetric than chemistry and physics. In this sense, there may be a discipline bias that explains why scientific production in polytechnic schools is more asymmetric. Among the 'generalist' universities, Roma La Sapienza is the one characterised by the highest level of asymmetry in scientific production. On the contrary, the universities in which scientific production is least asymmetric are the University of Catania and the University of Ca' Foscari in Venice. Between these upper and lower boundaries, there is a broad range of variance, particularly in the share of the least productive researchers.

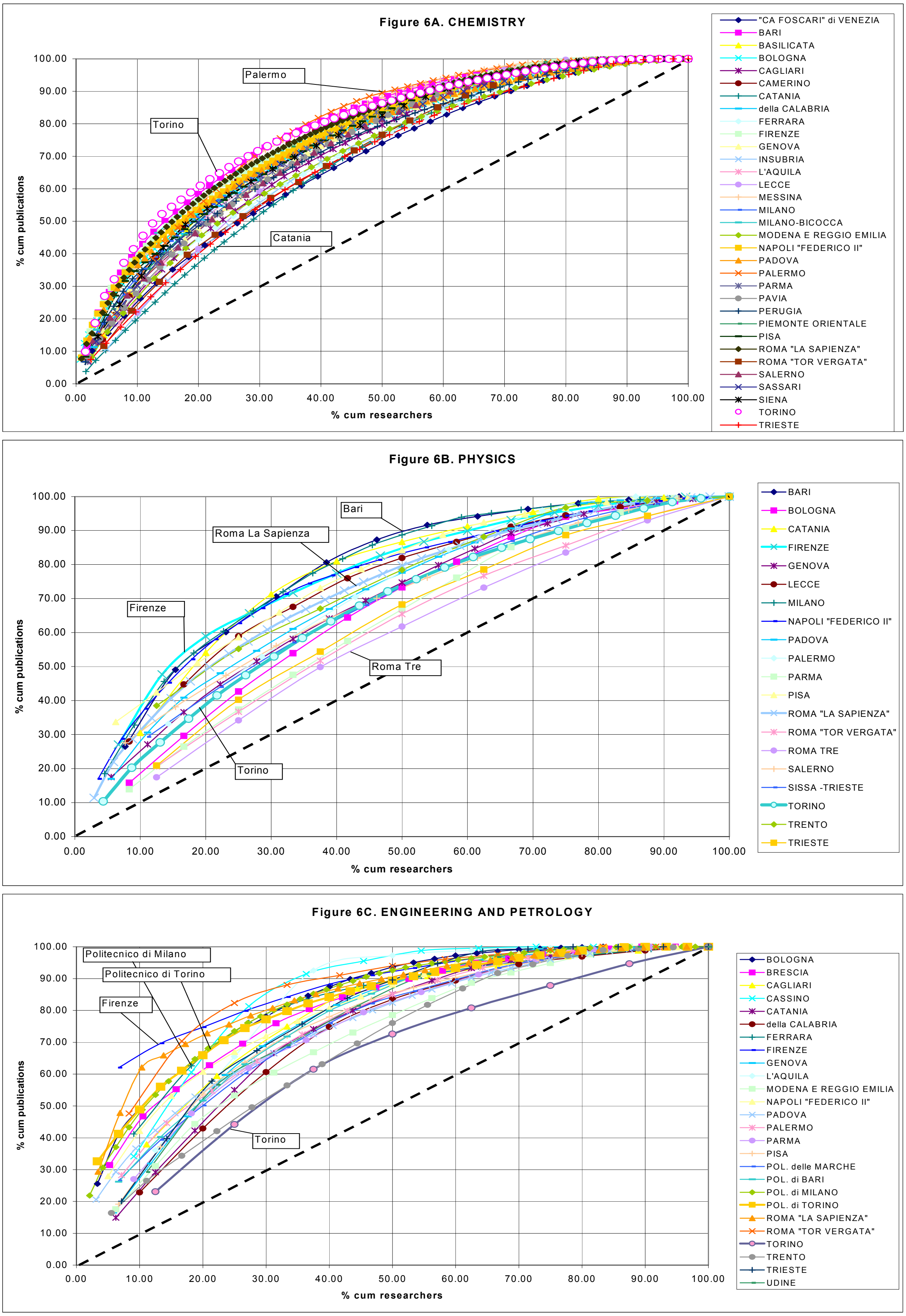
When we consider the 'stars' and, more generally, the very productive researchers, we find that the differences between universities are, as in the comparison by disciplines, relatively limited. At the 'lower bound' (University of Catania), 4.5% of the researchers account for 10% of the first 10% of publications, and 9% of researchers account for the first 20% of publications. At the ‘upper bound’ (Polytechnic School of Milan), 1% and 4% of the researchers account for the first 10% and 20% of publications respectively. The differences in very productive researchers among universities are in the magnitude of 4 to 5%. When we consider the different shares of the least productive researchers, on the other hand, the differences between universities are much higher, because at the university where scientific production is least asymmetric (Ca’ Foscari of Venice), 30% of the researchers account for the last 10% of publications, while at the Polytechnic School of Milan, 57% of the researchers are in the last 10% of publications. Therefore the maximum share of the researchers who are lagging behind is 57%, while the minimum is 31%: differences across universities account for a variance of almost 30% of the lower levels of scientific production.

The combination of the comparison by discipline with that by university (Figure 6) confirms that significant differences in the distribution of scientific production exist at both discipline and university levels, and that both sectoral and university specificities may explain the variance between more and less asymmetric scientific productions. For instance, the distribution of scientific publications at the University of Turin is more asymmetric for chemistry, less asymmetric for engineering and petrology, while for physics it is somewhere in the middle of the range of variance across universities.

Moreover, and perhaps more importantly, the comparison by university and discipline confirms the fact that the degree of variance across distributions is high both for ‘stars’ and very productive researchers, and for less productive researchers. The variance, however, is higher for the less productive researchers than for ‘stars’ and very productive researchers. The shares of ‘stars’ and very productive researchers are different across disciplines and universities, and the share of researchers lagging behind varies even more. This may raise a problem of governance (in terms of hiring, incentive and monitoring schemes) that is relevant at discipline and university level, and that is most important when we consider the risk of ‘marginalisation’ faced by increasingly large shares of less-productive researchers.

The comparison of individual publications by academic position of the single researcher may be useful in this context, allowing us to make out to what extent asymmetric distributions of scientific production are the result of weaknesses in hiring, incentive or monitoring schemes.

Figure 6. Individual distribution of publications by university and discipline



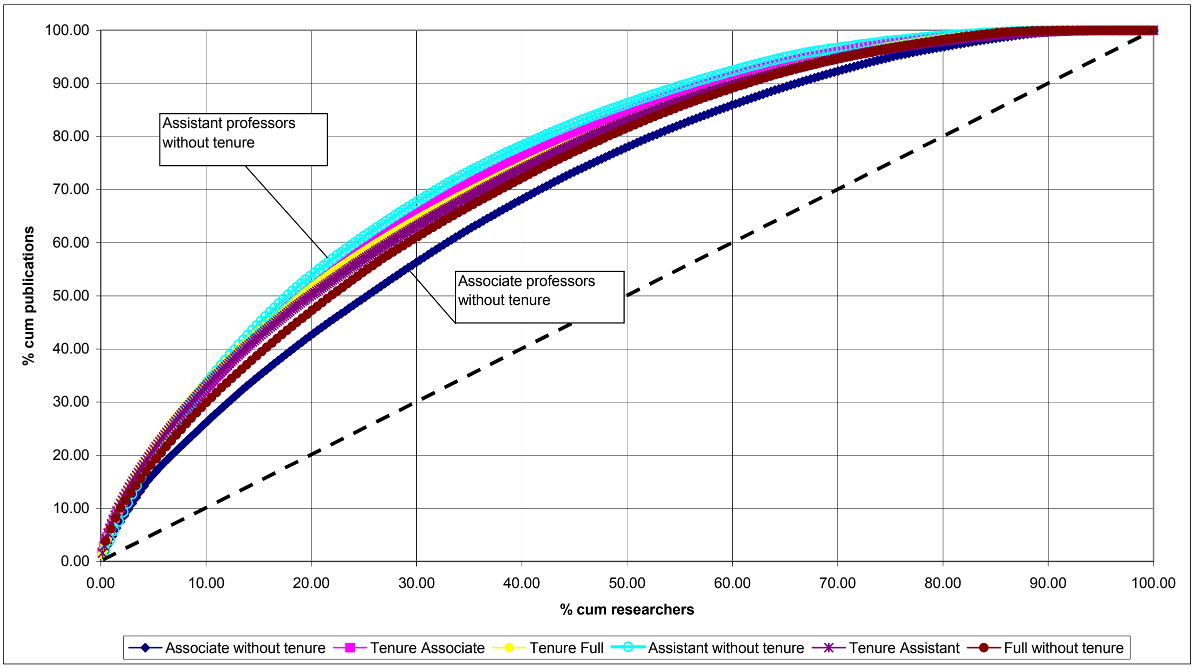
The relative frequency of the various positions (Table 4) by disciplines shows that whereas in the aggregate sample the different positions are represented quite proportionally, the relative frequency of the different positions varies by field. While assistant professors are generally a smaller category across disciplines, full professors are about 40% and 37% of all the researchers in physics and petrology respectively, and in engineering the larger group is represented by associate professors (37%). In chemistry the distribution of researchers across positions seems more balanced, each category representing about 1/3 of the total researchers.

Table 4. Relative frequency (%) of positions by discipline

Position	Disciplines				*Total*
	Chemistry	**Physics**	**Engineering**	**Petrology**	
Associate professors	34.87	30.34	37.26	31.62	*34.49*
Full professors	34.02	39.04	32.69	36.75	*34.61*
Assistant professors	31.11	30.62	30.05	31.62	*30.90*
Total	*100.00*	*100.00*	*100.00*	*100.00*	*100.00*

Figure 7 compares the distributions of individual publications on the basis of the academic position of the single researchers, distinguishing between researchers with and without tenure and between the following positions: 1) tenured assistant professors; 2) assistant professors without tenure; 3) tenured associate professors; 4) associate professors without tenure; 5) tenured full professors; 6) full professors without tenure.

Figure 7. Distribution of individual publications by academic position



The comparison by academic position confirms that individual publications are generally asymmetric, and also shows that there do exist some differences between positions (see also Table 5). To begin with, scientific production is less asymmetric when associate professors without tenure are considered, while assistant professors without tenure are characterised by a more asymmetric distribution. The share of 'stars' or very productive researchers shows in fact that 3% of associate professors without tenure account for the first 10% of publications, and 7% of associate professors without tenure account for the first 20% of publications. When other positions are considered, about 2% and 4% of researchers account for the first 10% and 20% of publications respectively.

More interestingly, when we examine the share of researchers who are lagging behind, we find that 33% of associate professors are in this condition, against about 45% and 40% of assistant professors without tenure and full professors (with or without tenure) respectively. If the relatively more asymmetric distribution across assistant professors can be explained by an age effect (for younger researchers, publishing may be more difficult, particularly as regards ISI journals), we can argue that the incentive mechanism, i.e. reaching tenure and eventually full professorship, is working better for associate professors without tenure. Considering that the share of tenured associate professors lagging behind is 42%, and that arguably there is no significant difference in terms of age between associate professors with or without tenure, it becomes possible to propose the hypothesis of a relative effectiveness of the incentive mechanism for associate professors without tenure.

On the contrary, the relatively high share (40%) of full professors (both with or without tenure) lagging behind can be explained in terms of 'bad' incentive and monitoring schemes, i.e. schemes that are unable to ensure 'good' levels of scientific production across researchers. Here the emphasis may be placed on the monitoring schemes, rather than on the incentive schemes, at least in terms of advancement in scientific career. Full professors have reached the highest position in the academic carrier, and the governance problem might lie in the monitoring of their scientific activity.

Table 5. Comparison between 'stars' and less productive researchers by academic position

Position	**Stars**	***(%)***	**Lagging behind**	***(%)***
Tenured Full Professors	13	*1.82*	285	*40*
Full Professors without tenure	4	*1.9*	81	*39*
Associate Professors without tenure	8	*3*	91	*33*
Tenured Associate Professors	12	*1.8*	276	*42*
Assistant Professors without tenure	6	*2.5*	10	*45*
Tenured Assistant Professors	10	*1.69*	233	*40*

The comparison of the distribution of scientific production by academic position and discipline may be useful in showing to what extent there are sectoral specificities in the effectiveness of hiring, incentive and monitoring schemes (Figure 8). In all the three disciplines, the share of 'stars' does not vary much across positions and disciplines. About 1-2% of the researchers account for the first 10% of publications, irrespectively of their position and discipline, and only small differences arise when we consider the first 20% of publications, the share of very productive researchers varying between 4% and 6%. On the contrary, a higher variance characterises the proportions of researchers lagging behind, showing that the problem of 'marginalisation' is also specific to the academic position of the researchers, and requires specific governance mechanisms.

Moreover, at discipline level, scientific production in chemistry and physics is relatively less asymmetric across positions than in engineering. (But in physics the share of associate professors lagging behind is quite high, 40%, showing that here there may be a problem of incentives for those researchers. In chemistry, on the contrary, this share is the lowest across disciplines, 34%, and appears to be working better, or to be less weak, than in the other fields.)

Figure 8. Distribution of scientific production by position and discipline

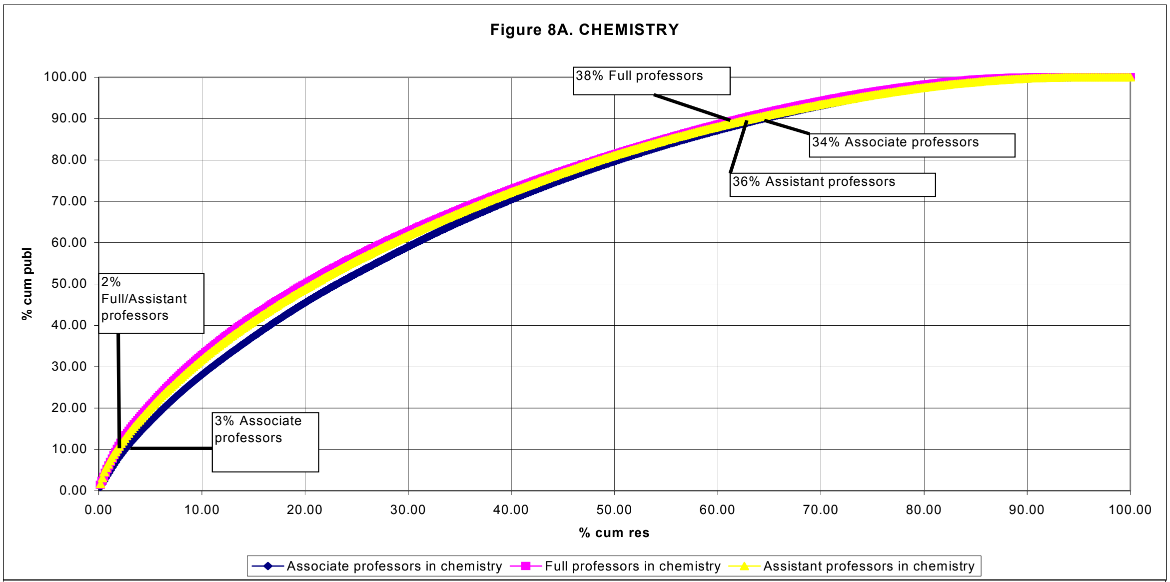


Figure 8A. CHEMISTRY

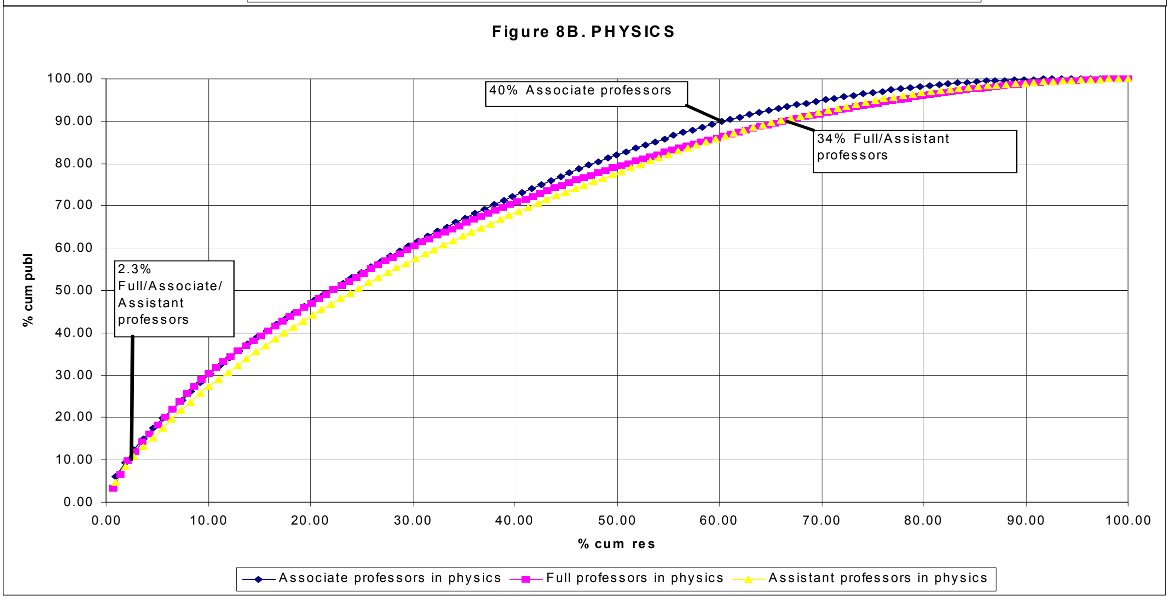


Figure 8B. PHYSICS

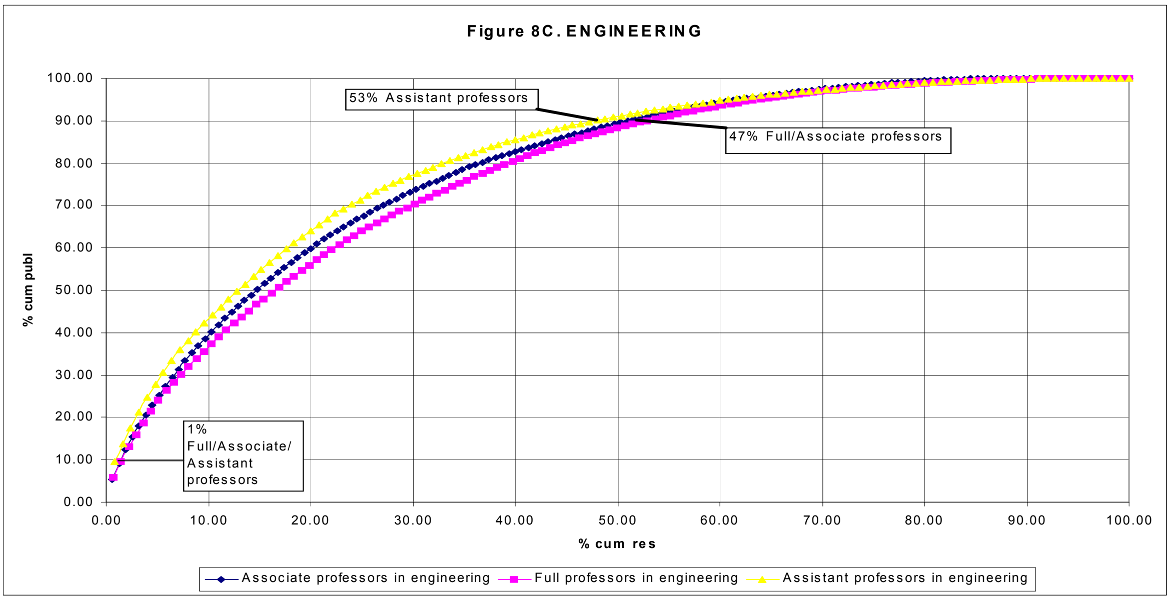


Figure 8C. ENGINEERING

The most interesting results seem to come from engineering. Not only is scientific production much more asymmetric in engineering than in chemistry and physics, but the magnitude of the problem of researchers lagging behind does not change much with position: 47% of full and associate professors and 53% of assistant professors are lagging behind and account for the last 10% of publications. There may therefore exist a structural problem of governance, specific to the characteristics of the discipline rather than to position. The organisation of scientific activity in an Italian university can be described as excessively individualistic: there is the risk of a high level of marginalisation among the researchers in the disciplines, such as engineering, that require models of governance more focused upon teamwork, collective research and joint projects. Sectoral characteristics matter, and require specific models of governance.

Finally, the comparison of the distribution of scientific production by university and position (Figure 9) confirms that university specificities matter when both the group of the 'stars' and very productive researchers, and that of the researchers lagging behind are considered. A high variance exists across universities and at all positions, showing that the quality of the hiring, incentive and monitoring schemes is also specific to the individual universities.

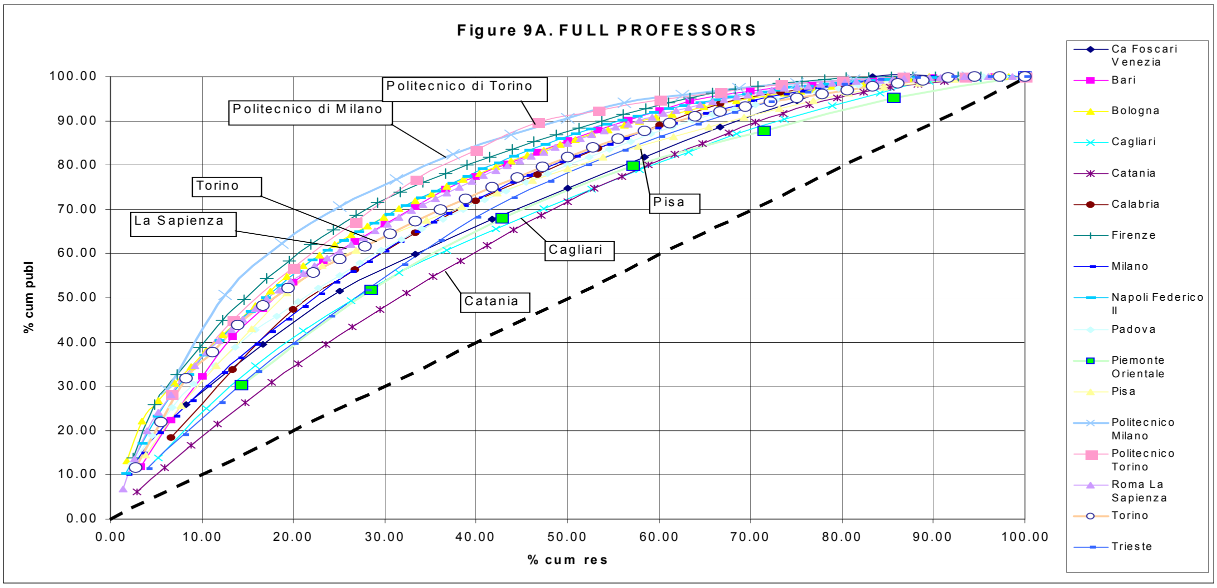
For instance, when we consider assistant professors lagging behind, at the lower bound (the university of Ca' Foscari in Venice), we find that only 20% of the researchers account for the last 10% of publications. On the opposite side, at the Polytechnic School of Milan, about 60% of assistant professors are in the last 10% of publications. Similarly, considering full professors at the University of Cagliari, we find that 26% of researchers account for the last 10% of publications, while at the Polytechnic School of Turin, 54% of full professors are lagging behind. Finally, 58% of associate professors at the Polytechnic School of Turin are lagging behind, against 30% of less productive researchers at the university of Ca' Foscari in Venice. In this context, the fact that the two polytechnic schools in Turin and Milan are characterised by a very asymmetric distribution may depend partly on discipline specificities, and, more precisely, on the highly asymmetric nature of scientific production in the engineering field.

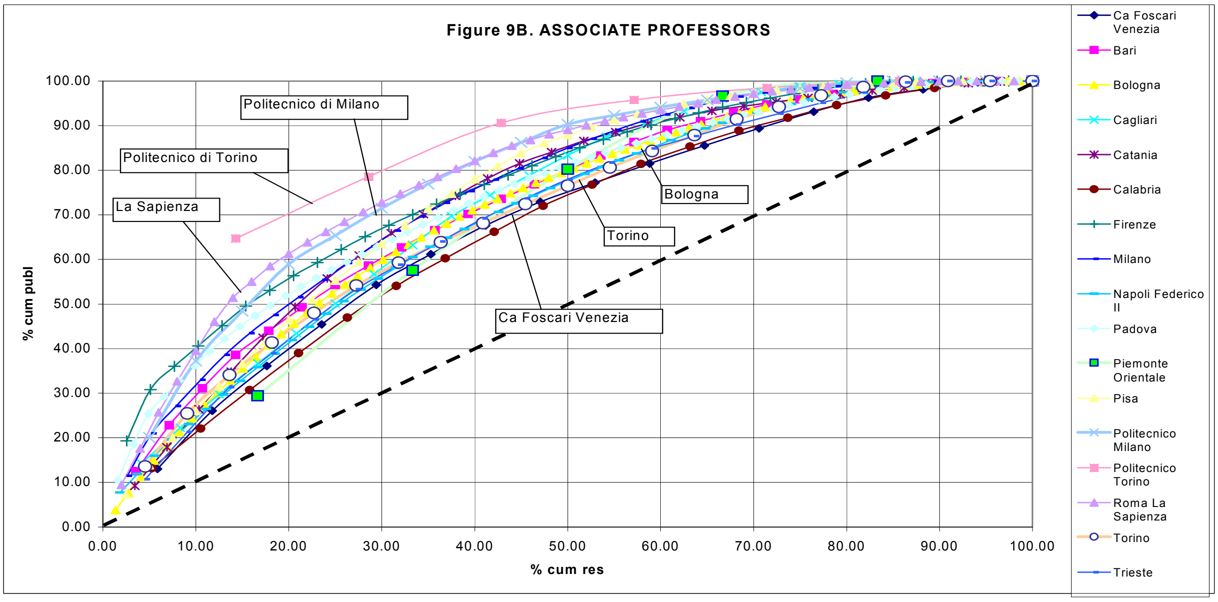
At university level, however, it is clear that the variance can be affected by at least two factors: 1) the scale of the university and the number of researchers by position, and 2) the number of publications generated by the most productive researcher in each position in the single institutions. In other words, a less asymmetric distribution may also be the result of a lower average quality of scientific research in the individual universities, allowing for positions. Future research should account for this effect as well.

In sum, the descriptive evidence of the distribution of scientific production by discipline, position and university shows that the asymmetry of publication is characterised by a relatively high level of variance, in relation to the scientific field, the academic position held by the researchers, and the university in which the researchers are active. These specificities may be relevant when we try to provide an interpretative framework that stresses the weaknesses of the models that govern scientific activity, especially in terms of hiring, incentive and monitoring schemes.

The next section will endeavour to put forward such an interpretative framework, which may also form the basis for a research agenda both on the characteristics of scientific production, and on organisation and its governance implications.

Figure 9. Distribution of publications by university and academic position





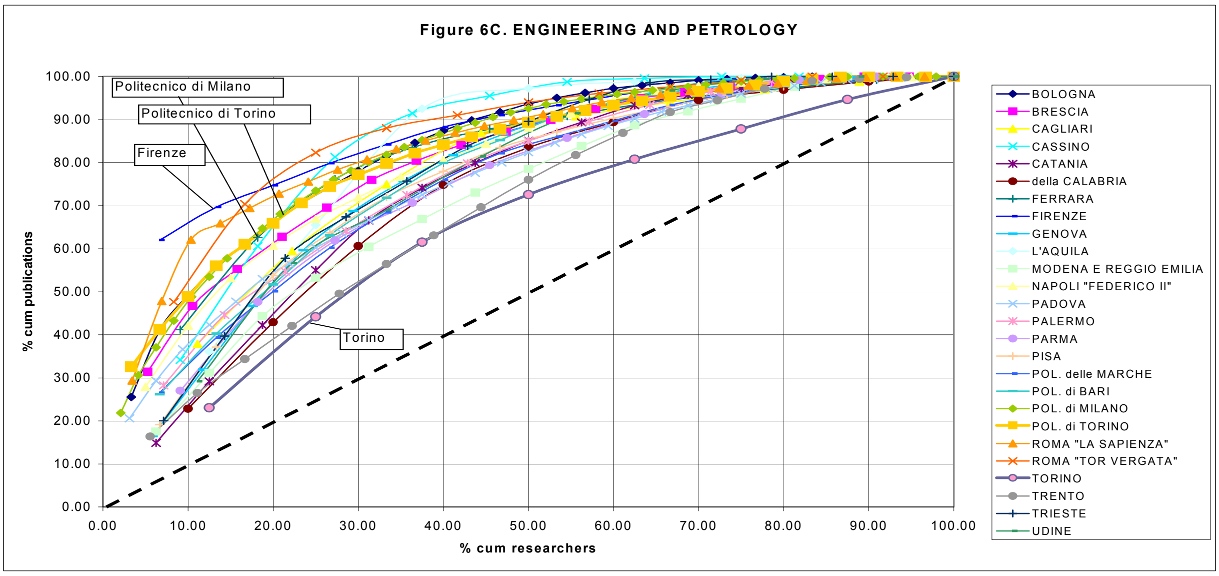


Table 6 compares the distribution of ‘stars’, very productive researchers and researchers lagging behind at aggregate, discipline, university and position levels, where ‘stars’ are defined as researchers who account for the first 10% of publications, very productive researchers as those who account for the first 20% of publications, and the least productive researchers are those who account for the last 10% of publications. The universities of Ca’ Foscari in Venice, La Sapienza in Rome, and Turin are taken as a ‘sample’ of the whole population, since they represent, respectively, the lower and upper bounds (i.e., the universities in which scientific production is relatively less and relatively more asymmetric), and a sort of average distribution, between the lower and upper boundaries.

The picture described in Table 6 synthesises the fact that, given the aggregate distribution, there exist some interesting differences between the specific levels (disciplines, universities, and positions) and the aggregate one, as well as within each level. While the share of ‘stars’ does not vary much (though at position level some differences do exist), and is very small across levels, the share of very productive researchers and, particularly, that of researchers who are lagging behind are characterised by higher levels of variance. This variance may suggest, in a preliminary way, that different disciplines, universities and positions imply different modes of scientific production, different ways of organising scientific labour, and therefore different governance models.

Table 6. ‘Stars’, very productive researchers and researchers lagging behind at aggregate, discipline, and position level (cumulated percentages)

	% publications (cumulated)		
% researchers (cumulated)	**First 10%**	**First 20%**	**Last 10%**
Total aggregate	1.0%	4.0%	42.5%
Engineering	1.0%	3.0%	51.0%
Physics	2.0%	4.8%	36.5%
Chemistry	2.0%	4.3%	39.0%
Ca Foscari Venezia	2.5%	7.5%	30.0%
Torino	2.0%	4.0%	37.0%
Roma La Sapienza	1.5%	3.4%	48.0%
Tenured full professors	1.8%	4.8%	40.0%
Full professors without tenure	1.9%	5.5%	39.0%
Associate professors without tenure	3.0%	7.0%	33.0%
Tenured associate professors	1.8%	5.0%	42.0%
Assistant professors without tenure	2.5%	5.0%	45.0%
Tenured assistant professors	1.7%	4.5%	40.0%

3.2. Scientific activity and R&D expenditures

In order to capture the effect of positive externalities stemming from R&D investments on scientific activity, this subsection correlates the amount of R&D expenditures at the regional level with scientific productivity (measured by publications per researcher) in each region (see Table 7).

Table 7. Publications and R&D expenditures (average value 1990-2001; .000 Euros; constant prices 1995) in Italian regions

Region	N□ publications	R&D tot	R&D private	R&D public
EMILIA-ROMAGNA	22,167.00	711,752.58	379,880.17	331,872.67
LOMBARDY	18,634.00	2,329,485.25	1,808,760.58	520,724.58
TUSCANY	15,513.00	558,687.58	206,042.25	352,645.25
LATIUM	13,229.00	1,904,128.17	606,956.08	1,297,171.75
CAMPANIA	11,050.00	500,306.83	191,737.92	308,569.00
VENETO	9,483.00	423,185.33	213,965.83	209,219.25
PIEDMONT	9,143.00	1,567,384.17	1,367,001.75	200,382.25
SICILY	7,748.00	311,484.67	59,695.00	251,789.50
APULIA	5,460.00	205,285.08	68,893.08	136,392.25
FRIULI	5,436.00	222,353.50	119,898.08	102,455.42
UMBRIA	4,823.00	84,283.17	17,613.50	66,669.92
SARDINIA	4,351.00	117,888.08	17,461.33	100,426.75
LIGURIA	3,621.00	332,267.25	166,293.00	165,974.17
MARCHE	3,396.00	96,425.67	31,421.08	65,004.17
CALABRIA	2,116.00	49,656.25	4,463.67	45,333.92
TRENTINO-ALTO ADIGE	1,709.00	75,325.58	31,990.83	43,334.50
ABRUZZO	1,656.00	150,082.92	84,504.83	65,578.00
BASILICATA	887.00	37,495.42	9,772.33	27,723.08
MOLISE	237.00	10,239.17	2,067.00	8,171.92

The results of the correlation (Figure 10a, b, c) show that the correlation is positive and significant when we consider total R&D, private R&D and public R&D.

However, scientific productivity increase less than proportionally with respect to R&D expenditure, since the exponent of the power law that represents the correlation is < 1 for total R&D, private R&D and public R&D (precisely, 0.7278, 0.5143 and 0.9052, respectively).

This shows that the role of positive externalities from R&D to science is limited and that the concentration of high levels of R&D expenditures generates benefits for scientific activity only to a minor extent. More precisely, positive effects are relevant only under a given threshold, that is, for smaller regions (in terms of both publications and R&D investments: the regions located in the bottom-left corner of the space). On the contrary, the cases of Piedmont and Latium (and even Lombardy), which are the three largest regions in terms of R&D expenditure, show that the output in terms of publications is considerably less than proportionate to R&D investment. We can identify only two 'good' models in which scientific production benefits from R&D investments, that is, in which scientific production is more than proportionate to R&D expenditures, namely Emilia-Romagna and Tuscany.

Figure 10a. Correlation between publications and total R&D expenditures

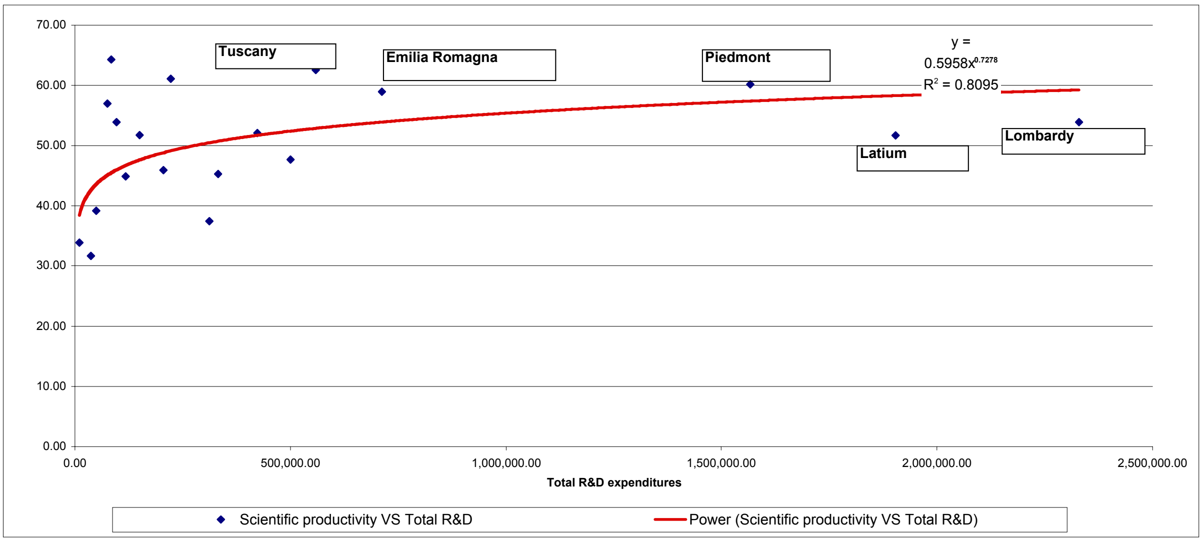


Figure 10b. Correlation between publications and private R&D expenditures

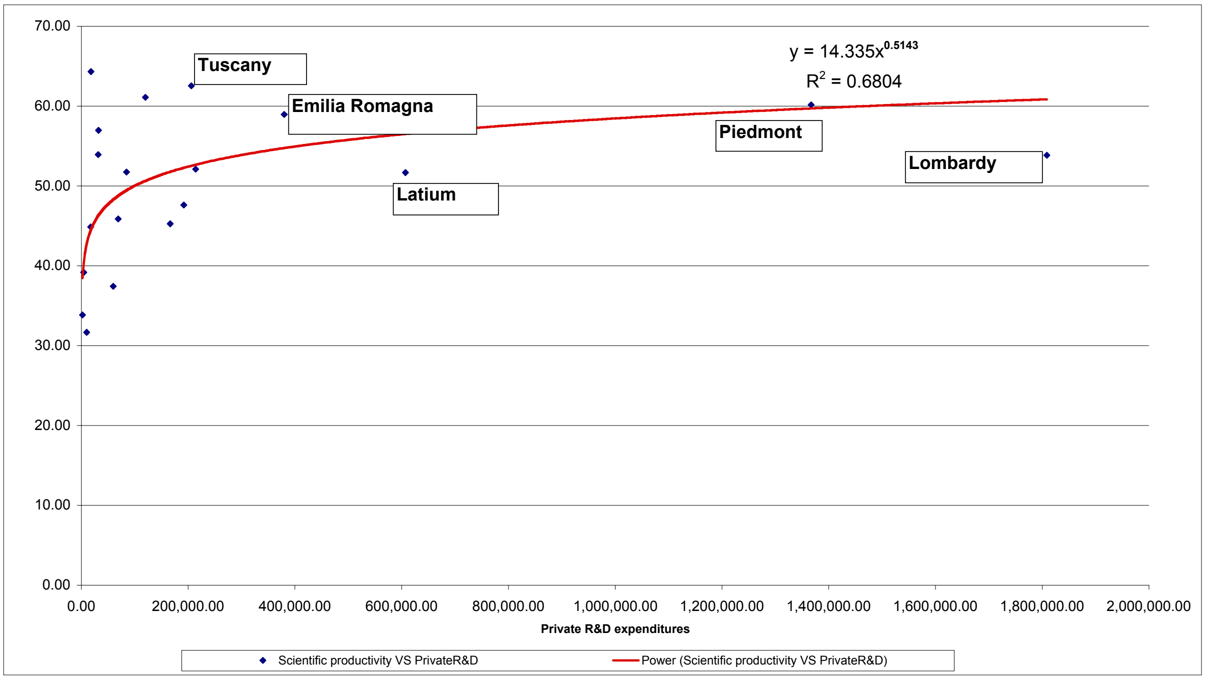
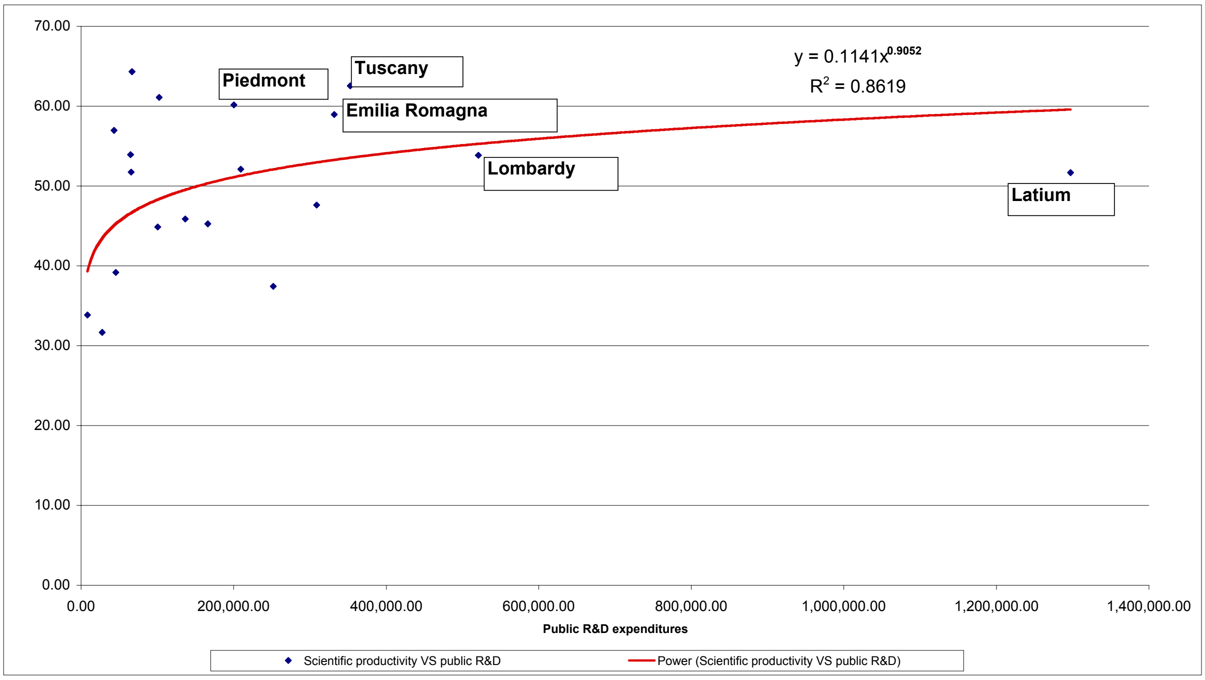


Figure 10c. Correlation between publications and public R&D expenditures



These results are particularly striking when we consider total R&D investments, and are reinforced when we consider private R&D expenditures (coefficient = 0.5143).

Our findings open the way for questioning the science policy that aims at the concentration of resources in given areas, in order to benefit from externalities from R&D to academic systems. They show that high concentrations of R&D expenditures do not promote a better performance in terms of scientific publications.

4. INTERPRETATIVE FRAMEWORK AND RESEARCH AGENDA

Our results, based on simple statistics, show that the extent to which the distribution of scientific production is more or less asymmetric largely varies according to discipline, university and academic position. How can these differences be explained?

In the first place, at discipline level, a new perspective may be useful for understanding the relationship between sectoral characteristics and scientific production, i.e. a perspective that places at the centre of the analysis the specific knowledge bases of the scientific fields, and relates this knowledge base to the specific pattern of learning, organisation of scientific work, and communication that characterises each scientific field. These patterns and norms, in relation to the characteristics of the knowledge base of the various scientific fields, emerge here as relevant.

The analysis developed by Michael Polanyi (1958 and 1966) and by Cowan, David and Foray (2000) on the different characteristics of knowledge may help us to understand 1) the way in which different knowledge bases require different learning and communication patterns, and 2) the way in which these learning and communication norms imply different forms of organisation of scientific labour.

On the one hand, when more tacit and technology-oriented sciences are considered (such as engineering), teamwork can be important because of the different communication patterns required. In this case, knowledge is not public and does not rely on a purely scientific base, but is much more the result of the implicit accumulation of experience, routines, learning by doing, and learning by interacting. However, when knowledge is tacit, scientific norms, methodology and techniques are only locally shared within small communities working on the same technical problem through idiosyncratic learning and communication practices. Engineering can be described as a synthetic work, based on the interaction and integration of different pieces of knowledge in order to solve complex problems. The integration of these different portions of knowledge, and their translation into a common language, may therefore be difficult. In this context, face-to-face communication and learning by interacting are a most important practice in the organisation of research activities. Because this knowledge base is mainly personal, even when it is shared within a community, master-apprentice-like relations, vis-à-vis interactions, and learning on the task are the typical modes of learning and communication of that knowledge, hence of the definition of a new scientific output (Polanyi, 1958 and 1966). In this context, good monitoring and incentive schemes, especially when focused on promoting interactions and the formation of research groups and collective projects, may be more important than hiring schemes.

On the other hand, more codified sciences such as physics may be characterised by spatially widespread learning and communication patterns, and by a more individualistic organisation of scientific work. According to the analysis developed by Cowan, David and Foray (2000), when knowledge is mainly codified, there exists a specific “codebook” defining the methodological and technical specifications of a given scientific output. The members of the scientific community systematically and explicitly refer to that codebook when generating scientific outputs, which are codified to a great extent. In this case, the knowledge base underpinning the generation of science is public to a greater extent, and acquisition and distribution occur mostly by means of so-called ‘blueprints’. The codified character of their knowledge base allows scientists to potentially share the same language, and to learn scientific methodology and techniques widely, in part independently from the team and academic context in which they are working. In this perspective, the individual characteristics of the single researcher (such as creativity, technical competence, the ability to find important research questions, and persistence) may be more important than teamwork. Good hiring schemes may be more relevant than monitoring and incentive activities.

Located somewhere between tacit and codified knowledge, articulable knowledge entails some degree of codification. Nevertheless, the codebook that establishes the definitions of this knowledge (and the procedures for its implementation into a given scientific output) is not manifest even to the members of the community that ultimately employs the knowledge and develops the scientific output. Even if there does exist an explicit book of definitions and instructions, it is not explicitly consulted, and the contents of this codebook have been so fully internalised and appropriated within the community that they operate as implicit sources of technical knowledge and procedural rules. Articulation being social communication, the degree to which articulable knowledge is public or private, hence the extent to which it is shared within the social and scientific community, depends on the costs of access, transmission and absorption of the relevant technical specifications and implementation procedures (Cowan, David and Foray, 2000). This may be the case of chemistry, typically an applied science, where learning and communication patterns, and the importance of the role of teamwork may vary according to the specific field. From the governance viewpoint, an appropriate mix, in terms of hiring, monitoring and incentive schemes, should take these field specificities into account.

In the second place, considering the academic position, our evidence (particularly when combined with a consideration of discipline specificities). reveals the existence of differences across positions. An appropriate governance of science should take these specificities into account, implementing localised and specific mechanisms and schemes, according to the particular characteristics of the category of researchers and discipline considered. In this regard, governance mechanisms aiming to reduce the marginalisation of assistant professors (i.e. by cutting down the number of less productive assistant professors) should focus on appropriate hiring criteria and upgrading programs. At the same time, governance mechanisms aimed at ensuring a less asymmetric scientific production for associate and full professors should be focused on the implementation of appropriate incentives and monitoring procedures[7]. However, these governance models cannot take sectoral differences for granted. For instance, hiring, incentive and monitoring mechanisms aiming to ensure ‘good’ scientific production in engineering should focus upon the collective nature of scientific work in the field, rather than encourage an individualistic behaviour that may lead to a considerable marginalisation of the relatively less productive researchers, such as those observed in our evidence. Here an appropriate science policy might be based, for instance, on the implementation of a network of researchers, as well as on joint projects and multidisciplinary collaborations.

In the third place, the combination of sectoral and positional specificities may be relevant at university level. The high variance between the various universities in terms of asymmetry in the distribution of publications across disciplines and positions also demonstrates that local governance systems vary according to the scientific field and the position of the researcher. In some cases, governance may turn out to be weak, for instance when we consider the effectiveness of the hiring and upgrading mechanisms for assistant professors in chemistry; in other cases, for instance in physics, governance may turn out to be weaker in monitoring the full professors’ activity. In this context it is clear that the specialisation of the individual universities in given fields, but also in given positions (i.e., the relative share of full, associate and assistant professors in the various universities) does matter. At the same time, the level of scientific publications, particularly that of the most productive researchers, also matters in this context. In other words, the distribution of scientific production in a given university may be less asymmetric than in others simply because on average the researchers there publish less than researchers in other universities.

Finally, considering the relationship between R&D expenditure and scientific publications at the regional level, science policies supporting the concentration of a high amount of resources in a given area, in order to take advantage of the externalities between R&D and science, might be called into question by this research. Positive effects of R&D investments on scientific publications exist only under a certain threshold, and when we consider regions characterised by large R&D investments, we find that some of these benefits disappear. The

---

[7] Obviously monitoring and incentive issues are relevant also to young assistant professors, and upgrading should be relevant also to associate and full professors. What I wish to stress here is the importance of specific mechanisms that are crucial for the different academic positions, or at least relevant to the Italian academic system. Hiring a ‘bad’ assistant professor, or failing to provide the appropriate upgrading programmes for a ‘good’ young assistant professor, will lead to the progressive marginalisation of the researcher, who will become an inefficient member of the system. Thus hiring, upgrading and training are crucial for assistant professors. On the contrary, it is clear that the hiring issue is no longer relevant to associate and full professors.

system of incentives and the rationales for science may differ from those of the R&D system, and the next steps of this research should investigate this issue.

5. CONCLUSIONS

The simple and descriptive evidence presented in this paper confirms that the distribution of scientific production is highly asymmetric. It also shows that significant differences emerge in this distribution when scientific production is compared across disciplines, universities and academic positions. A considerable degree of variance across distributions appears when scientific production is compared at the sectoral, university and position levels. These differences can be interpreted in terms of sectoral, university and positional specificities, and can also be correlated with the ways in which scientific activities are organised and governed. In particular, sectoral specificities seem to be quite strong according to the ways in which researchers acquire and upgrade their competence, organise and implement their research activity, and are stimulated to publish by various governance schemes. For instance, high levels of marginalisation in technical sciences such as engineering and petrology, can be due to a lack of incentives to organise collective research activities and to implement teamwork and collaborative projects that are crucial for the discipline. Engineering is very much a synthetic science based on the integration of, and interaction between, the different skills required to solve complex technical problems. An individualistic organisation of work and individualistic incentives may be inappropriate, and may explain the very asymmetric nature of scientific production in the field. Finally, positional specificities also seem important for proposing appropriate governance schemes. Assistant, associate and full professors may be stimulated to achieve 'good' levels of scientific production by different means. When we consider the differences between individual distributions of scientific production across positions, we find that good hiring and upgrading mechanisms may matter less in the case of young assistant professors, while on the contrary effective incentive and monitoring schemes may be more important for associate and full professors.

A variety of localised models of science governance may turn out to be relevant when the differences in scientific production across disciplines, universities and positions are considered. Generalist governance models may be inappropriate to sustain high quality and less asymmetric levels of scientific production, since they do not take these differences into account. Models of governance must be localised according to discipline, university and position specificities. The articulation of these governance models may be very important in the understanding and implementation of science policy, and will be the subject of one of the next steps of this research project. Obviously, this articulation should be based on deeper evidence of the characteristics of scientific production at discipline, university and position levels. In particular, the following issues may be very important and should be taken into account in the next steps of the research.

To begin with, an analysis of the characteristics and organisation of science in the various disciplines may be quite important for explaining differences in scientific production, and deserves a more thorough treatment. For instance, the initial assumption that chemistry can be understood as an applied science, more similar to technical sciences such as engineering and petrology rather than to theoretical sciences such as physics, does not seem correct. The distributions of scientific production in chemistry and engineering are more similar to each other and quite distinct from those of the technical fields. In this respect, the nature and organisation of scientific activity in the different types of science needs to be qualified carefully, through a qualitative and historical analysis.

Moreover, the differences across disciplines in terms of the knowledge base (tacit, codified, or articulable) underlying the scientific fields needs to be integrated with an analysis of other factors and cognitive characteristics that may be important for understanding the different structure and organisation of scientific production in different disciplines. For instance, some of the differences between the disciplines may be due to differences in the definition of what is regarded as an important output. Engineers may view publication as less important than theoretical physicists do. In this perspective, ISI publications may be a less appropriate measure of scientific production in engineering than they are in the natural sciences.

Furthermore, engineers may also have different career aspirations, which could make them more mobile (less likely to stay in the academic environment) and more prone to external and consulting activity, thus devoting less time and smaller resources to 'true' scientific research and publications. Differences between disciplines in terms of the number of years individuals work in academia may exist, and may explain the differences in scientific productivity.

Finally, the next steps of this research should also account for:

1) The existence of scale effects, in terms of both the number of researchers and R&D expenditures, in scientific production at discipline and university level.
2) The effect of local specialisation (university level) in terms of sectoral scientific production.
3) The differences, in scientific production and in its asymmetry, across universities in terms both of the shares of the various types of researchers (e.g., are there universities in which assistant professors are overdimensioned with respect to associate and full professors? how does this relate to an asymmetric distribution of scientific production?), and of the value of publications (i.e., what is the relationship between different levels of scientific production and asymmetry in the distribution of publications?).

REFERENCES

Cowan, R., David, P. A. and Foray, D. (2000), The explicit economics of knowledge codification and tacitness, *Industrial and Corporate Change* 9 (2), 211-253.

CRUI – Conferenza dei Rettori delle Università Italiane (2002), La Ricerca Scientifica nelle Università Italiane: Una Prima Analisi della Banca Dati ISI, CRUI: Rome.

Dasgupta, P, and David, P. (1994), Towards a new economics of science, *Research Policy* 23 (59, 487-521.

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

Katz, J. S. (1999), The self-similar science, *Research Policy* 28 (5), 501-518.

Polanyi, M. (1958), *Personal Knowledge. Towards a Post-critical Philosophy*, London: Routledge & Kegan Paul.

Polanyi, M. (1966), *The Tacit Dimension*, London: Routledge & Kegan Paul..

Senker, J. (1995), Tacit knowledge and models of innovation, *Industrial and Corporate Change* 4, 143-151.

Zucker, L. G., Darby, M. R. and Brewer, M. B. (1998), Intellectual human capital and the birth of US biotechnology enterprises, *American Economic Review*, 88 (11), 290-306.