

Does Research Organization Influence Academic Production?

Laboratory Level Evidence from a Large European University

Nicolas Carayol ^{φκ} and Mireille Matt ^φ

*^φ BETA, UMR CNRS 7522, Université Louis Pasteur
61 Avenue de la Forêt-Noire 67085 Strasbourg*

First Version: June 2003

^κ Corresponding author. Tel: +33(0)390242103. Fax: +33(0).3902420710;

carayol@cournot.u-strasbg.fr.

Abstract: The paper analyses the scientific research production at the laboratory level. The evidence on which the study is based describes precisely the research activity over a decade of more than eighty labs belonging to Louis Pasteur University, a large and well ranked European research university. We analyse the research organization of the labs via correlations of various inputs focusing on the characteristics of the research personnel and examine its effects on two outcomes: publications and patents. Moreover, the paper proposes a typology of laboratories, which enables us to present different styles of research organization and their performances. We show how an appropriate combination of inputs may be strongly associated to a high publication and inventing productivity. Some of the main results concern the shares in which to associate full time researchers and university professors in academic labs, the role of interdisciplinarity, the correlation between patents and publications, and the importance of non permanent researchers.

Key Words: Economics of science, Academic Research Laboratory, Organization, Typology.

JEL Classification: L31, 031, 032, 034, 038

1 Introduction

The individual researcher is usually the standard level for analysing academic research production. Most studies concentrate on the effects of individual determinants of academic productivity such as age (Zuckerman and Merton, 1972; Diamond, 1986; Stephan and Levin, 1991), cohort (Weiss and Lillard, 1982), training (Garcia-Romero and Modrego, 2000) and gender (Stephan, 1998). Nevertheless, Stephan (1996) recently highlighted that such studies have a weak ability to explain research productivity, due to the collective nature of research. Thus, some scholars suggest that further investigation of academic research production should take into account some collective level of organization, such as - specifically in the institutional European context - the laboratory level (Dasgupta and David, 1994; Stephan and Levin, 1997). Such a consideration is congruent with even more recent analysis, evidencing that the quality of other researchers belonging to the laboratory is a crucial variable for explaining individual productivity¹ (Mairesse and Turner, 2002). Nevertheless, it is one thing to assess collective effects on individual productivity and it is another to change the level of analysis as the results of Adams and Griliches (1998) demonstrate². The results at a collective level may be quite different from the individual researcher especially due to important externalities among researchers within labs: critical knowledge spillovers, reputation, sharing of equipment instrumentation and facilities, complementarities between different types of researchers, or even between different research agendas.

While most of the literature focus on the determinants of individual productivity or of the productivity of universities or groups of universities, few contributions analyse the academic production at the laboratory or team level. Such a concern brings into the scope papers more interested in organization and managerial issues, which tend to show that the laboratory is indeed a critical level of organization. Crow and Bozeman (1988) show that laboratory funding structure is strongly correlated with the nature of the research realized, Joly and Mangematin (1996) illustrated various types of contracting and laboratory strategies and Laredo and Mustar (2000) developed a technique for stressing 'laboratories activity profiles' according to five dimensions (production of certified knowledge, education and training, research and innovation, involvement in the construction of public goods, and participation in public debates). Arora, David and Gambardella (1998) studying academic research at the laboratory (or team) level showed that (even if relying on cross section data) there are

¹ These results are grounded in initial evidencing of the Matthew effect (accumulative advantages in science) by Merton (1968), first discussed in economics by David (1994). Carayol (2003a) introduced a model of such dynamically biased competition grounded on the difference the inherent productivity of research positions in different universities.

² They showed that university departments exhibit decreasing returns to scale at the university departments level of analysis while the whole system of universities exhibits constant returns to scale.

decreasing returns of funding on “quality adjusted” publication while the most reputed teams reveal an elasticity of scientific performance with respect to funding which approaches unity. Bonaccorsi and Daraio (2002) showed that the distributions of the average age of scientists in the labs and the size of the labs are correlated. In three out of six domains, they found that the size of the labs (i.e. the number of researchers) is negatively correlated with productivity.

Even if these studies improve considerably our understanding of academic knowledge production, there are still many unknowns to reach a robust industrial organization theory of academic research production. Such an aim certainly lies beyond the scope of this paper, the purpose of which is to improve our understanding of how laboratories differ in organizational terms, in outcomes and how organization is related to productivity. By organization we mean the composition of research laboratories in terms of labour force (full-time researchers, university professors, PhD and post-doc students, non researcher personnel) and its features (average age of the labour force, size of the team, discipline etc.). We use both publications and patents to characterize the outcome of the research.

A first originality of our contribution is to analyse the way inputs are correlated inside laboratories. For instance we question whether permanent and non-permanent researchers are complements or substitutes, how non-researchers, junior-like positions and females are allocated to different labs and the degree of inter-disciplinarity. The second novelty comes from the simultaneous use of patents and publications as outcomes of research (while the literature often focuses on one or the other). Thus we are able to study the relation between various inputs and various outputs. Moreover, we enrich the output-output relation by analysing the link between patents and the nature of publications (co-authored with foreign institutions or with firms).

If an empirical study such as ours is still lacking (to the best of our knowledge), it is probably due to the unavailability of appropriate data. The constitution of an original and unique database allows us to begin to tackle such issues. The data concern the research activity of more than eighty labs belonging to Louis Pasteur University (ULP) of Strasbourg, covering more than a decade. ULP is a large and well-ranked European research university, counting nearly two thousand permanent researchers and nearly hundred laboratories belonging to various scientific disciplines. In the paper we are privileging qualitative statistical methods in accordance with our objective, which consists in offering a first deep investigation of the organization of academic research production at the laboratory level. The last outcome of the study is a typology of five coherent classes of labs that we describe precisely.

The remaining part of the paper is organized as follows. The next section presents a literature survey on academic research production and gives the reader some examples of the expected added value of the laboratory level of analysis. The third section provides information on the data and some descriptive statistics. In the fourth section, we develop a systematic correlation study: between the variables characterizing the structure of the research laboratories (input indicators), between variables

describing the outcomes, and finally between input and outcome indicators. The fifth section is dedicated to the typology. The last section discusses the main results obtained and concludes.

2 Literature survey

The objective of this part is not to present a detailed review of the broad literature developed by sociologists and economists of science. We have selected among the various topics specific questions related to our paper. We underline results comparable with our findings, even if the levels of analysis or the methodologies differ. Lastly we present some examples illustrating how the laboratory level of analysis may contribute to the present state of our knowledge on academic research production.

2.1 Teaching vs. research positions and research productivity

The sociology of science questions whether research and teaching in academia are complementary or competitive activities. Some authors consider them as joint activities in the sense that one reinforces the other. Others regard them as “conflicting roles with different expectations and obligations” (Fox 1992, p. 293). Fox (1992) used a survey based on a sample of social science faculty to analyse how research and teaching activities influence the publication productivity of social scientists. The study covered four social sciences (economics, political science, psychology and sociology) in BA-, MA-, and PhD- degree granting departments in the US. She showed that faculty members with high publication productivity exhibit strong interest in, commitment of time to and orientation to research. They are not strongly involved in both activities, but favour research activities. Her results suggest that the more productive researchers spend fewer hours teaching, preparing for courses and consider teaching as less important than research. Her findings tend to prove that research and teaching are conflicting actions.

2.2 The research context and publication productivity

Empirical studies in the USA found that researchers at prestigious university departments are more productive and are more cited than their colleagues in lower-ranked universities (Cole and Cole 1973). A debate about the causality between productivity and department prestige appears: is there an effect of the department on the scientist's productivity or does productivity influence the prestige of the position obtained? These two hypotheses labelled by Allison and Long (1990) as the “selection hypothesis” and the “departmental effects hypothesis” are not incompatible and could be mutually reinforcing, but a need to assess their relative importance exists.

Long and McGinnis (1981) based their study on a population of biochemists who obtained their doctorates in four specific years. They defined six organizational contexts by distinguishing between academic and non-academic sectors on the one hand, and contexts, which do or do not encourage publication on the other. The probability of being employed in a specific context is not strongly influenced by the publication productivity or the citations. And, once employed in a specific

environment, individual productivity soon conforms to the characteristics of that context. Long (1978) showed similar results in Biochemistry within the academic system only. Allison and Long (1990) analysed 179 job changes by scientists from one academic organization to another. They considered four disciplines: Physics, Chemistry, Mathematics, and Biology. In their study, scientists moving to more prestigious places increased their rates of publication and of citation; those moving to less prestigious institutions showed substantial decreases in productivity. Their results tend to underline that the departmental effects hypothesis is more important than the selection hypothesis.

Mairesse and Turner (2002) analysed the impact of laboratories characteristics³ on the individual productivity of CNRS physicists. They showed that if the stock of publications of the lab increased by 10%, the individual researcher of this lab would publish in journals receiving 0.26 more citations on average. The scientist would also publish 0.6 papers more per year on average. Thus, the productivity of the research labs and the quality of their publications influence positively the productivity of their researchers and the quality of their papers.

2.3 Age and research productivity

Some economists (Diamond 1986, Levin and Stephan 1991, Weiss and Lillard 1982) focused on the publishing activity of scientists in life-cycle models, trying to explain the link between age and scientific production. Diamond (1986) studied the publishing activity of mathematicians at Berkeley and showed that it declines continuously with age. Weiss and Lillard (1982) analysed the publications for 1000 Israeli scientists throughout their careers. The average annual number of publications first tends to increase, and then to decrease. Levin and Stephan considered six disciplines. Except for particle physicists, in the other fields publishing activity first increases, reaches a peak around mid-career and declines. Mairesse and Turner 2002 underlined similar life-cycle effects explaining the productivity of a sample of French CNRS physicists. Stephan (1996) emphasized that the explanatory power of these models is rather low. This would suggest that if an age-publication relationship exists, other important factors probably influence publication behaviours.

Bonaccorsi and Daraio (2002) analysed a large database related to the research activity of the Italian National Research Council (CNR). The minimum average age in labs is 30.4 and the maximum average is 48.56. They found that the distributions of scientist ages and of the total personnel are bimodal. This result suggests that entry of persons follows a waveform dynamics, due to expansion and contraction periods of recruitment. They obtained a negative relationship between productivity indicators and the average age of researchers. However, the average age of senior scientists and research directors is not significantly related to productivity. They refuted the hypothesis of a life-

³ They take into account the size of the laboratory (number of researchers), the productivity of the lab, the quality of the publications of the lab (impact factors of the journals) and the degree of international openness (share of papers written with foreign co-authors).

cycle effect and assumed that the average age of a laboratory reflects its attractiveness and scientific vitality. They implicitly had in mind the existence of a virtual circle: higher prestige induces greater resource availability for young researcher positions and thus increases the attractiveness.

Wallmark (1998) analysed the patent activity at Chalmers University of Technology (Göteborg, Sweden) and underlined that most patents are applied for by researchers between 30 and 35, but also by *emeriti*.

2.4 Size and research productivity

Is the production of scientific knowledge subject to increasing, diminishing or constant scale returns? For increasing returns, if all inputs were increased by a factor t , the output would increase by more than t . One of the main problems of measuring the nature of returns of scientific production lies in the possibility to assess all inputs used. In other words, the omission of one or more inputs could unduly lead to revealing decreasing returns. Thus, the results of this kind of analysis should be taken cautiously.

Adams and Griliches (1996) used data on 50 universities and 5 scientific fields (Biology, Chemistry, Mathematics, Medicine and Physics). They found diminishing returns to the individual university R&D, for number of papers and for the total citations. These diminishing returns at the individual university level are counterbalanced by spillovers from other universities, institutions and countries. They changed the nature of the input and replaced R&D by scientists and engineers. The coefficients on papers and citations increased, thus indicating measurement problems linked to the choice of input. In a later work (2000), they considered the scientific productivity of US universities in 8 scientific fields. They showed that at aggregate level, the research production follows constant returns to scale as at the individual university level, diminishing returns prevail. These differences are explained by more important measurement errors at the individual level and by the existence of spillovers, which may only be captured by an aggregate analysis.

For instance, at the laboratory level, Bonaccorsi and Daraio (2002) found that size is never positively correlated with productivity; on the contrary in three domains (Chemistry, Environment and Engineering) size and productivity are negatively linked. In almost all fields the most productive labs are the smallest, and the least productive ones may be large or small. They also underlined the correlation between size, institute age and average age of researchers. Mairesse and Turner (2002) stressed that the influence of the size of the laboratory on the individual production of researchers and on the quality of their papers is significant but small. The relation is negative for the average number of papers written by scientists and slightly positive for the citations.

Concerning the production of patents, the results seem to be more contrasted. Wallmark (1998) and Henderson *et al.* (1998) showed that the largest research organizations apply for more patents. Payne and Siow (1999) explained that an increase of \$1 million in federal research funding to a

university results, on average, in 11 more articles, 0.2 more patents and \$411000 more faculty salaries. Foltz *et al.* (2000) also found a positive and significant effect of federal funding on patents. Coupé (2001) used patent counts and patent citations as outputs, different R&D lags, types of universities, technology classes, and time effects and concluded to the presence of constant or diminishing returns to scale.

2.5 Patent and publication

Stephan *et al.* (2002) used data from the 1995 Survey of Doctorate Recipients and analysed the patent activity of a sample of doctoral scientists and engineers, focusing on the relationship between patenting and publishing at the individual level both in academia and industry. Their main question was whether publications and patents were complements or substitutes. They first underlined that for the whole sample, less than 20% of scientists applied for at least one patent while 70% published at least one paper in the period covered by the analysis. For academia, these percentages are 10% for patenting and 83,3% for publications. The probability for one scientist to apply for at least one patent is significantly related to whether or not this scientist has published at least one paper. The number of patent applications is positively and significantly linked to the number of papers published. Patenting and publishing activities seem thus to be strongly related at the individual level.

2.6 Expected effects at the laboratory level of analysis

The point of departure of our study is the focus on the laboratory level of analysis. Our expectation is that it sheds new light on the issues mentioned above and may allow us to explore new ones. Let us briefly present some examples, which may illustrate some of the expected added values the lab level analysis may have in this respect.

Concerning the impact of teaching and research vs. full time research positions on scientific production, the analysis at the laboratory level allows us to tackle the issue of a right proportion of both types. This issue is far from being trivial because for instance a few full time researchers may enhance the connectivity of a few university professors to front line research. On the other hand, having university professors in a lab may enhance the productivity of full time researchers just by driving PhD students to them. This raises new questions in terms of access to non-permanent research human resource (PhDs and post-docs) and its impact on productivity, which are ignored by the specialized literature.

Researchers age is also an interesting example of a variable that may induce collective effects. In a laboratory, seniors or experienced researchers may increase the productivity of juniors thanks to collective work or simply due to informal contacts. Conversely, junior researchers may stimulate the productivity of older ones, known to have fewer incentives in their late careers. Thus the question does not any more concern the productivity trajectory over the life cycle but concerns gathering researchers of different ages who may be in some respects complementary.

The institutions to which agents belong may also influence the individual productivity just because of the informational externality. For instance in France the prestigious labs are those affiliated to a public research organization such as CNRS or INSERM. Such an affiliation is operated by a peer review process taking into account the productivity of the lab. Affiliated labs benefit from increased monetary funds for research facilities, research positions and PhD or Post-doc grants. In other words, affiliated labs should constitute a favourable environment in terms of facilities and intellectual abilities, encouraging and facilitating scientific production. We should thus find that high institutional recognition should influence the production of the lab positively.

A broader issue concerns the question of how inputs correlate in labs and how it affects research productivity. We may also test the relation between the outputs of the research production process: do patents and publications correlate or exclude each other at the laboratory level? Here again the laboratory level has an added value. The question is no more related to the allocation of the time of one scholar between research projects leading to patents and more fundamental ones. It becomes a question of allocation of resources within the labs, where scholars may be specialized on different but complementary research agendas, which may or may not have a potential for patenting. It may also be an issue of managing research agendas having different time line⁴.

3 The data

The objective of this part is to present the general context of our statistical analysis. We will briefly describe the different research activities of Louis Pasteur University and its scientific reputation. Then, we will explain the way we collected the data, present their characteristics and richness and provide some results of the distribution of labs.

3.1 Louis Pasteur University

The data concern the research activity of one single university, namely Louis Pasteur University (ULP) of Strasbourg. This university is quite large and diversified. Seventeen separate institutional components (i.e. engineering schools, teaching and research units, and various institutes) are located in six campuses in the Strasbourg area in which around 18000 students are enrolled. Research and teaching activities cover a wide range of subjects in the fields of medical sciences as well as Mathematics, Computer Science, Physics, Chemistry, Life Sciences, Geology, Geophysics, Astronomy, Engineering Sciences. Human and social sciences are also present in the specific fields of Economics, Management, Geography, Psychology and Educational Sciences.

⁴ For some empirical evidence on the management of research agendas in laboratories and how this influences their strategies for collaboration with firms one may see Carayol (2003b) and references therein.

ULP has an old tradition of fundamental research and a long-term standing of scientific excellence. Its researchers have received numerous national and international scientific prizes, including the Nobel Prize for Chemistry awarded to Jean-Marie Lehn in 1987 for his contribution to the field of Supra-molecular Chemistry⁵. Overall, ULP is one of the largest French universities in terms of research. The Third European Report on Science & Technology Indicators 2003 ranks first among French universities in terms of Impact and 11th among European universities. Such research capacity is enhanced by a close-knit with the major national research bodies such as the National Centre for Scientific Research (CNRS) and the National Institute for Health and Medical Research (INSERM) present in the Strasbourg area.

3.2 Origin and characteristics of the variables

The presentation of the variables depends mainly on the method used to collect them. We collected data related to the personnel and the laboratories from internal administrative sources. Information related to publications came from the ISI Web of Science and the French Patent Office provided patent data. We will present the characteristics of our variables and the way labs are distributed according to these variables.

3.2.1 Laboratories and personnel

We collected the variables from administrative reports completed for the 1996 contractual affiliation round. Such a round occurs every four years. All laboratories (and also Faculties and Institutes) have to produce a standardized document, which is usually divided into two distinct parts: i) a précis of the past four years and, ii) a project for the next four ones. The data concern the period from 1993 to 2000, which may be separated into two four-year sub-periods: 1993-1996 and 1997-2000. These documents are evaluated through standard peer review procedures conducted by both the Ministry of Research and Education and funding agencies such as the CNRS and INSERM whose support is expected. We recorded their decisions concerning the affiliation. The affiliation to CNRS and INSERM means increased funding for research facilities and positions. The affiliation operates through a peer review process mostly taking into account scientific production arguments and constitutes clearly a signal of the labs scientific excellence.

⁵ Also Ferdinand Braun was awarded the Nobel prize in 1909 in physics for his radio telephone (together with Marconi) Other Nobel price winners spent some time in their careers at Strasbourg University: Wilhem Roetgen, Alphonse Laveran, Hermann Staudinger, André Lwoff and Louis Néel. Louis Pasteur began his research career as an assistant professor at Strasbourg. René Thom who was awarded the Field Medal for his well know contribution to ‘catastrophe theory’ stayed at Strasbourg University for his PhD and to get his first tenured Professorship. Eleven of the university’s researchers are members of the Institut Universitaire de France (IUF) and eleven are members of the Academy of Science.

We collected the standard information mentioned in those documents. We gathered many typical information about the personnel of the labs, specifying the number of individuals in each detailed category of personnel and individual information on permanent researchers including the name, the sex, the age, the status (teaching position vs. full-time researcher), the grade ('senior-like' position stands for full-professorship and for Director of Research and 'junior-like' position for Assistant Professors and Associate Researchers).

We recorded 83 distinct laboratories in 1996. We have reliable and complete information for all but two of them about for which we miss the complete characterization of their permanent researchers. Thus for such variables as grade, age, sex of permanent researchers, a sub-population of 81 labs will be considered. Among the 83 labs, 43 are funded by the CNRS⁶, 9 by INSERM, while 31 are supported by the Ministry of Research and the resources of ULP. Among the 1460 permanent researchers, 760 are full-time researchers directly paid by the CNRS and INSERM and 700 are university scholars. On average, the permanent researchers are 51.5 year old. Among them, 360 are females (24,6%). 57% of permanent researchers occupy junior-like positions while the others benefit from senior-like positions. We also find some 1930 non-permanent researchers: 1230 PhD students and 710 post-docs. Lastly some 1120 non-researchers (administrative staff and technicians) and 410 visitors are recorded.

ULP laboratories are rather small in terms of permanent researchers (61 labs, i.e. 74%, have less than 20 permanent researchers), except 8, which declare more than 41 scientists. The distribution for non-permanent researchers exhibits the same characteristics: 70% of labs have less than 20 non-permanent researchers and 21% more than 41. For a large majority of labs (53, i.e. 65%), more than 50% of permanent researchers occupy junior-like positions. The average age of permanent scientists is above 50 in 70% of the labs, 5 labs have an average age between 40 and 45. Women represent less than 25% of the permanent researchers in 53 laboratories (65%). Only 4 laboratories are constituted by a majority of female permanent researchers.

We also got the main scientific disciplines of the labs according to the specific categories defined by the University. These are the following: Astronomy (1 lab), Biotechnology (12 labs), Chemistry (9 labs), Genetics and Cellular and Molecular Biology (9 labs), Geography (2 labs), Mathematics (1 lab), Mechanics (1 lab), Medicine (22 labs), Odontology (1 lab), Neurosciences (7 labs), Condensed Matter Physics and Chemistry (4 labs), Subatomic Physics (2 labs), Earth Sciences (2 labs), Information Sciences and Technology (3 labs), Humanities and Social Sciences (8 labs).

⁶ Two types of association with the CNRS exist: UMR (Unité Mixte de Recherche) and UPR (Unité Propre de Recherche). The latter is more closely supported by the CNRS. Nevertheless, these labs may be supported by the university and/or may host university researchers.

To complement such information, we used data we had on permanent researchers, in order to collect information on their disciplinary affiliation at the most detailed level possible as indicated by the institutions to which they are affiliated (University National Council, CNRS, INSERM). Such classifications do not perfectly match at the sub-discipline level we are interested in. Nevertheless, thanks to a normalization grid produced by the OST (Observatoire des Sciences et Techniques) specifically for the French system, we were able to allocate nearly all permanent researchers to 50 different sub-disciplines according to a unique nomenclature selected as the reference (namely the one of the National University Council) (sub-discipline level normalized through the OST nomenclature; cf. OST, 2003). We found that 11 laboratories are mono-disciplinary, 15 have 2 disciplines, and 21 have 3 disciplines. The mode of the distribution is 4 disciplines in 22 labs. The maximum number of disciplines is 8 and only 2 laboratories are concerned.

3.2.2 *The outcomes*

Our database also integrated information about the various outcomes of research production, namely publications and patents. For each permanent researcher we collected his/her published articles (using SCI, SSCI and Arts and Humanities ISI databases). We found more than 26000 occurrences over the 1993-2000 period. This amount includes some double counting as quite a few ULP researchers may have co-authored papers. By dividing each occurrence by the number of co-authors we obtain the effective (normalized) scientific contribution of each author⁷. The total scientific performance is 6040. The median number of co-authors is five. We differentiate between two types of co-authorship. A co-publication is “international” if at least one co-author belongs to a non-French institution. Some 10400 occurrences (i.e. 40%) exhibit international co-authorship. We also screened co-publications in which at least one of the co-authors belongs to a firm. 1200 publication occurrences with industrial partners exist and represent 4.6% of all publication occurrences.

Concerning the total number of occurrences (with double counting), 70 laboratories have published between 1 and 400 papers. The distribution is rather uniform in this first interval. The distribution is much more dispersed for labs having published more than 401 papers. For instance, only 2 laboratories have published between 2001 and 3000 articles. Concerning the total scientific performance, 67 laboratories have published between 1 and 100, the distribution between these labs is heterogeneous and the mode is 20-40 publications in 18 laboratories. 3 labs have published between 300 and 500 papers. The distribution of the average performance by permanent researchers is unimodal. The mode of the distribution is an average of 2 to 4 articles per permanent researchers in 27 laboratories. 12 laboratories show an average of more than 8 papers. These results show that a small number of labs publish a large amount of papers and that a large number of labs publish a small

⁷ An author is necessarily a permanent researcher.

number of articles. At the individual level, many contributions show the same tendency, and more precisely Lotka (1926) found that 6% of publishing scientists produced half of all papers.

The behaviour in terms of co-publication is rather contrasted: in 5 laboratories the average number of co-authors is between 1 and 2 and in 5 other laboratories the average is above 8. The mode is 5 to 6 co-authors on average for 30 laboratories. The distribution of international co-authorship is heterogeneous. In 9 laboratories, more than 50% of the publications are international. In 62 laboratories, less than 37.5% of the papers signal a foreign co-author. The mode is between 25 to 37.5% of international publications in 28 laboratories. 10 laboratories never published with industry and at the other extreme one laboratory published between 25 to 30% of its papers with industrial partners. The mode is between 0.1 and 5% of publications with private partners in 38 laboratories.

We also looked for the French and European patents, which had been invented by at least one of the ULP permanent researchers. To do so we matched our permanent researchers table with the patent data provided by the French Patent Office (INPI) covering the same 1993-2000 period. We found 850 occurrences of French or European patents granted. After a cleaning process, which eliminated the extensions of French to European patents, we ended up with 463 patents invented by researchers from ULP over the eight years considered. Some 189 patents were granted in the first sub-period (1993-1996) and 274 in the second one (1997-2000) giving an increase rate of 45% between the two sub-periods.

The data exhibit a decreasing relation between the number of laboratories and the number of patents obtained by a laboratory. 38 laboratories have no patent, 31 laboratories have between 1 and 10 patents, 5 laboratories have between 11 and 20 patents and one lab obtained between 51 and 60 patents. The distribution of labs according to the number of patents per permanent researcher also shows a decreasing shape. In 26 laboratories, the number of patents per permanent researcher is between 0,01 and 0,5, and in 8 laboratories it is between 0,5 and 1. Finally, two labs have between 2 and 4 patents per permanent researcher.

4 Academic research organization at the laboratory level

In this section we turn toward studying correlations between the various variables concerning labs. These variables may be separated in two subsets: inputs and outputs. The inputs set should be interpreted broadly since under this label we associate all variables describing the way laboratories work and are organized.

The correlation studies are conducted in three steps. First input-input correlations describe the common and potential distinctive features of the organization of public research within laboratories. Second, input-output correlations provide information about how publications and patents are influenced by each type of input. Finally, we also study output-output correlations in order to have a first investigation about whether the patents and publications are complements or supplements, i.e.

whether the labs which publish more are also the ones that patent more and if the nature of the publications counts. Distinguishing between sub-periods may help us track causal relations.

4.1 General and specific features of laboratories structures

We mainly use the various characteristics of the personnel to analyse the organization of research laboratories. We provide results about the way permanent and non-permanent researchers are linked, about the allocation of non-researchers, junior-like positions and females in the different labs. We also present evidence of the link between disciplines, size and institutional recognition and finally about the distribution of disciplines inside labs.

4.1.1 The labour force structure

When analysing the structure of the personnel, one may wonder whether permanent and non-permanent researchers are complements or substitutes. To support the former hypothesis, we find that the number of permanent researchers and the total non-permanent researchers are significantly correlated (0.66).

At a more disaggregated level, the underlying association of permanent and non-permanent researchers within labs is more complex. PhD students are primarily correlated with professors (0.67). Nevertheless, they are also strongly correlated with CNRS researchers (0.60). This result does not mean that full-time researchers often supervise PhDs. It could be due to the fact that CNRS researchers are mainly located in the most recognized labs, which may provide more grants, thus attracting more PhDs. That observation seems to be supported by the fact that PhDs are negatively correlated with EA (Associated Teams) laboratories (-0.32). The latter are only supported by the Ministry of Research and involve only a few full-time permanent researchers. Thus, PhD students seem to be primarily allocated to laboratories in which university professors are present. This may be explained by the importance of personal contacts during the late stage of their graduate studies. Excellence also matters for the matching process of students to labs, but appears to be secondary.

On the contrary, post-docs seem to value only fame and excellence when choosing labs, especially foreign post-docs who are correlated with CNRS researchers (0.46) and high institutional recognition (0.23), especially in UPR-type of CNRS association (0.32). They are not correlated with university professors while at least a small but significant correlation would have been expected. This result could give some support to the 'attractiveness' hypothesis developed by Bonaccorsi and Daraio (2002) but for post-doc positions, which could be considered as pre-recruitment situation.

The non-researcher personnel are strongly correlated with CNRS scientists (0.70) and with non-permanent researchers (0.72). Their allocation seems to be also very dependent on the disciplinary environment. Non-researchers are correlated with Subatomic Physics (0.47) and Genetics and Cellular and Molecular Biology (0.23), which require (especially the former) heavy instrumentation and many

technicians and engineers. One may observe that the presence of non-researchers is not significantly correlated with any variable characterizing the institutional recognition of the labs. These observations tend to support the hypothesis that non-researchers are mainly allocated for practical needs (determined by disciplinary environments), and through a pure linear scale fashion not being driven by excellence or reputation considerations.

Average age is only correlated with Medicine (0.30). This probably indicates a specific long-lasting training in that domain. Moreover the share of junior-like positions among permanent researchers is not correlated with any other variables. These observations tend to indicate that the arrival of researchers in labs probably followed uniform global trends without being affected by any of our variables. For instance, neither the size of laboratories, nor their institutional recognition or their disciplines seem to have modified the arrival of permanent researchers in labs. Again the hypothesis of a linear and regular allocation⁸ of personnel in labs would hold for scientists. Such results also show that the laboratories housing older permanent researchers are not necessarily the most recognized (institutional recognition is not correlated with average age).

Concerning gender issues, correlation studies do not indicate a specific allocation of female permanent researchers. Nevertheless the share of women among permanent researchers is correlated with INSERM affiliated labs (0.27), which indicates high institutional recognition and human health concern.

4.1.2 *Disciplinary differences*

When looking at the disciplinary differences, we found that the laboratories in Medicine tend to be of a smaller size (negative correlation with all variables recording numbers of agents). These labs seem also to have a rather low institutional recognition. This may reflect a low quality of research but also explain their ability to get substitute funding from their simultaneous commitment within the University Hospital. In that respect, the laboratories specialized in the promising and rapidly growing field of Neurosciences are exactly the opposite: funding agencies support them.

The labs in Genetics and Cellular and Molecular Biology, which are also of a big size especially considering the number of post-docs, appear to be strongly supported by funding agencies. Subatomic Physics labs are also of a large size, but only in terms of CNRS permanent scientists and non-researchers (technicians), but not in terms of non-permanent researchers (PhDs or post-docs).

4.1.3 *Inter-disciplinarity*

The number of sub-disciplines to which permanent researchers are associated is correlated with most of the variables related to size indicating an expected scale issue. The strongest correlation

⁸ Bonaccorsi and Daraio (2002) found that allocation of researchers followed a waveform dynamics.

coefficient is with the number of PhDs (0.48) potentially reflecting a reinforced attractiveness of inter-disciplinarity in the eyes of younger researchers. In support of that view, the entropy index appears to be correlated with the number of national post-docs (0,27) while it is not significantly correlated with other input variables. The inter-disciplinarity entropy index records on a unique scale how flat the distribution of the researchers over the whole set of sub-disciplines is: the lower the entropy coefficient, the higher the concentration of disciplines. Such a diversity (disciplinary entropy) index is defined as follows: $-\sum_{i \in j} ((n_{ij}/N_j)\ln(n_{ij}/N_j))$ with subscript i indicating the sub-discipline and subscript j referring to the laboratory considered, n_{ij} is the number of permanent researchers belonging to lab j who are associated to sub-discipline i, and N_j the total number of permanent researchers belonging to lab j. Looking at Table 5 (cf. Annex I), one may observe that some labs have twice the entropy index of others while both have the same number of sub-disciplines (e.g. 3 or 4).

4.2 Characteristics of the laboratories and their production of publications and patents

The aim of this part is to study the connections between the various outcomes of research and the characteristics of laboratories. We first examine the correlation between inputs and the publishing performance of labs. Then we focus on scientific articles co-authored with researchers belonging to foreign institutions or to private firms. Next we look at the correlation between inputs and patents. Finally we concentrate on the link between disciplines and the types of outcome produced.

4.2.1 *The publication performance*

One main and probably not so surprising observation concerns the productivity of university professors. The correlation between publication performance per permanent researcher and university professor is negative (-0.30). Moreover, university professors are negatively correlated with the share of authors among permanent researchers (-0.29) while there is no significant correlation between university professors and both the variance and the kurtosis of performance among authors. Thus apparently, professors who publish may be quite as productive as CNRS researchers while other professors totally gave up publishing. This difference may be explained by the fact that some faculties favour research activities and others dedicate more efforts toward administration or teaching activities. This result could thus be justified by the idea developed by Fox (1992) that teaching and research may be considered as conflicting activities (cf. 2.1).

Publication performance is also strongly negatively correlated with 'junior' positions (-0.42). This result could be explained by a lack of experience of younger researchers. Nevertheless, scientific performance appears not to be significantly correlated with age. Several complementary explanations may be provided. Some junior scientists may probably be un-promoted researchers with lower research abilities or publication incentives. Or, this result may be partially influenced by the fact that a high share of junior researchers occupy Assistant Professor positions, combining teaching, research

and, sometimes, administrative activities. Concerning age our results do not corroborate those provided by the life-cycle literature (cf. 2.3)

Finally, the distribution of publication performance among authors within labs may be affected by the presence of non-permanent researchers: the variance of publication performance is correlated with foreign post-docs (0.33). Moreover PhDs and foreign post-docs are strongly correlated with the kurtosis of scientific performance distributions among authors (0.37 and 0.47). These results seem to indicate that the hosting of non-permanent researchers benefits unequally to the permanent researchers of the lab. Their efforts are probably allocated to the benefit of the most famous scientists or, in other words, post-docs are strongly attracted by labs with a few highly productive researchers. These explanations could in some ways be supported by observing that there is a high correlation between the kurtosis of the number of co-authors and the PhDs and foreign post-docs (0.41 and 0.44), indicating that the more numerous the non permanent researchers the skewer the distribution of co-authors among publications.

4.2.2 *International publications*

CNRS researchers have strong connections with the international scientific community: their correlation with the share of international collaborations among all publication occurrences is 0.34. A scale issue seems to appear since this share is correlated with the number of permanent researchers (0.31) and the number of non-researchers (0.31). Nevertheless, the latter result may be due to an indirect effect since there are more non-researchers in labs belonging to disciplines such as Subatomic Physics in which publications count impressive numbers of co-authors. The number of international publications is not correlated with university professors. This result should be compared with the 0.74 correlation coefficient one gets with CNRS researchers. This indicates that the publishing professors collaborate on a more local or national basis than other permanent researchers.

The variance of international collaborations among authors is strongly correlated with both CNRS researchers and non-researchers. The kurtosis of the distribution of international collaborations among authors is equally correlated with all size variables: this seems to indicate that when size increases, the long range collaborations are more concentrated on a lower share of permanent researchers.

4.2.3 *Publications co-authored with industry*

The share of publications written with at least one industrial partner is negatively correlated with both university professors (-0.30) and junior-like scientists (-0.26). They both seem to be more oriented toward the scientific community. As mentioned earlier, both categories are also the least productive ones. As a consequence, they may not be attractive in the eyes of potential industrial partners. This observation is congruent with the fact that the kurtosis of the distribution of industrial

collaborations among authors is correlated with university professors (0.78). The latter is also true concerning PhD candidates (0.57), who may ‘fuel’ such collaborations.

4.2.4 *Patent production*

CNRS researchers are the most important contributors (0.55) to patent production, while university professors are not significantly correlated with patent production. This may be due to historically grounded differences, which may be progressively reduced: university professors are positively correlated with an increase in patenting (0.24) and with a move from non-patenting to patenting (0.25) between the two sub-periods. A possible scale effect may exist, since we observe that CNRS researchers and PhDs are also correlated with the increase rate in patenting activity (0.49 and 0.40). Nevertheless this explanation may be controversial and the results may primarily express a progressive popularity of the patenting activity among the scientific community.

The quality of the labs measured by high institutional recognition is correlated with patents (0.25). The patenting productivity appears to be especially strong in Biotechnology and Genetics and Cellular and Molecular Biology both in terms of patents per permanent researcher and per researcher. It should be noted that high institutional recognition is also correlated with the number of publications (0.34) and the publication performance (0.42). Our results seem thus to follow the same pattern as the one developed in the literature (cf. 2.2): high recognition is positively correlated with publication and patent productivity.

The non-permanent researchers seem to be even more important for patenting than the permanent researchers (high correlation coefficient especially for post-docs 0.69). The French post-docs are correlated with the patenting productivity of permanent researchers (0.24). Permanent scientists may probably assign more applied (even if potentially less academically rewarding) problems to French post-docs. The latter may also be willing to select these more applied activities in order to acquire a first practical experience valuable on the private job market. Finally non-researchers are also correlated with patents (0.51).

4.2.5 *Disciplinary specificities*

Biotechnology is characterized by a strong patenting activity of both permanent (0.31) and non-permanent researchers; this field is also correlated with patents per researcher (0.36). In this field, patenting activity seems to have been supported by a specifically high level of collaboration with industry since Biotechnology is the only field to be correlated with a high share of publications co-authored with firms (0.30). Nevertheless, the patenting activity in this field, which seems to correspond to a long tradition, decreased surprisingly between the two sub-periods: Biotechnology is positively correlated with patenting on the 1993-1996 period (0.26) while it is negatively correlated with an increase in patenting between the two sub-periods (-0.29). The dynamics of patenting in that field is quite surprising and should be more carefully analysed.

The field of Genetics and Cellular and Molecular Biology is also correlated with both the number of patents per researcher and per permanent researcher (0.29 and 0.27). This field is moreover strongly correlated with the total number of patents. This result is probably due to size issues: labs of this field tend to be quite big. That may also account for the observation that this domain is correlated with the laboratories publication performance. Contrary to the Biotechnology domain, the patenting activity in the field is increasing (0.30). That increase could be grounded in scientific performance, which appears to be specifically important in the first period: the publication performance in the first period and the field are strongly correlated (0.24) while it is not significantly correlated with the publication performance in the last sub-period.

The field of Medicine appears to be the only one positively correlated with the publication performance per permanent researcher and per researcher (0.35 and 0.31). That result seems to be due to a quite flatter distribution of that performance among authors in the field since it is negatively correlated with the kurtosis of scientific performance distribution among authors (-0.29). Nevertheless, Medicine is negatively correlated with the share of international collaborations among publication occurrences (0.30).

The Neurosciences appear to be at an earlier stage, exhibiting still few or no patents: correlated with the “zero-to-zero patents” dummy (0.25). Nevertheless that field seems to have recently strongly interested firms since it is correlated with both the share of publications co-authored with firms (0.33) and its increase over the two sub-periods (0,35).

4.3 Complementarities or substitution between outcomes

We first look at the potential correlations between the publication performance and characterization of the set of publication occurrences. Next we turn toward the potential reinforcing or exclusion between publishing and patenting.

4.3.1 Performance and type of publications

The share of publications with foreign institutions among all publications is not correlated with the publication performance per permanent researcher apparently indicating that there is no complementarity between international and national publications: the labs which have more international collaborations are not the ones that publish more on a national basis. Nevertheless the share of international publications in the first sub-period is correlated with the publication performance (0.26) indicating probable scale or threshold effects in publication performance of the laboratory for getting access to international publication networks.

Industrial collaborations per permanent researcher are strongly correlated with publication performance per permanent researcher (0.52) indicating that the labs, which publish more, have also more research collaborations with firms. Moreover, it appears that the labs in which permanent

researchers are the most connected to the international scientific community are also the ones that collaborate more with firms: international publications per permanent researcher and industrial collaborations per permanent researcher are positively correlated (0.31). That may express the fact that internationally visible scientific labs tend to attract more firms.

4.3.2 *Publication and patenting performances*

The complementarity hypothesis between pure scientific productivity and invention appears to be strongly supported by the data against the exclusion hypothesis: the publication performance per permanent researcher is strongly correlated with the patents per permanent researcher (0.33). The result is also true when looking at the correlation with patents per researcher, even if less strongly (0.24). Since the increase in patenting is strongly correlated with the whole publication performance (0.37) there may be a scale effect here. Our results seem thus to confirm the findings of Stephan *et al* (2002) (cf. 2.5).

Patents per permanent researcher and patents per researcher are even more strongly correlated with collaborations with industry (0.44 and 0.37). The causality seems to go this way since the rate of increase in patenting between the two sub-periods is strongly correlated with the collaborations with industry (0.49). The intensity of publication collaborations with industry seems to be important for the intensity in patenting since the share of industrial collaborations among all publication occurrences of the lab are positively correlated with patents per researcher and permanent researcher (0.27 and 0.29).

Complementarity is also observed in the case of international collaborations: the rates of increase in patenting and international publication occurrences are correlated (0.29). Nevertheless the size issue may generate such results. The important observation is that the number of international collaboration occurrences is the only variable significantly correlated with the “zero to positive” patenting dummy (0.29). Such a result tends to indicate, in the second sub-period considered, that the most internationally connected labs may have shifted from a non-patenting and exclusive publishing behaviour to a patenting behaviour.

To support that observation and more generally, the recent patents seem to be more strongly correlated with publication performance than the ones of the first sub-period: patents in period 1993-1996 are correlated with publication performance of the two sub-periods with coefficient 0.55 and 0.54; while patents of sub-period 1997-2000 have the following correlation coefficients with the publication performances: 0.65 and 0.64. This may indicate that patents tend to be increasingly grounded in publication performance without any feedback decrease publication.

5 The typology of laboratories

As we have seen earlier, the labs differ according to both their input and their output structures. Moreover we have obtained many insights on how the variables correlate. We are now turning toward

a deeper analysis that would allow us to identify various ‘styles of research production’ at the lab level. There are obviously different solutions to associate inputs in labs that generate different amounts of the two types of outcomes. To do so we built a typology of laboratories following a standard methodology composed of a multi-correspondence analysis (MCA) followed by an ascendant hierarchical classification (AHC). The reader should refer to the Methodological Appendix in Annex II for a precise description.

Since our typology is intended to mix both considerations, the variables retained for the MCA are related to both. More precisely, the variables used for building the typology are the following: i) inputs: full time researchers, university professors, share of junior positions among permanent researchers, average age of permanent researchers, sub-disciplinary entropy of permanent researchers, national post-docs, foreign post-docs, non-researchers, institutional recognition; ii) outcomes: publication performance per permanent researcher, patents per permanent researcher, share of authors among permanent researchers, variance of publication performance among authors, share of industrial collaboration among publication occurrences and share of international collaborations among publication occurrences. These variables are transformed into qualitative variables with three modalities: low, medium and high. Only the shares of international and industrial collaborations among publications are transformed into dichotomised variables. Information on the variables selected are presented in Table 6 in Annex II. The typology is built for the 81 labs about which we have complete information on the above-mentioned variables of interest.

Following the MCA, we selected four axes, which collectively represented 39.13% of the inertia. Then, the AHC was realized based on the co-ordinates of the labs on these four axes. Five coherent classes of labs were selected because, in retaining these five types, the within-classes variance of the total variance was nearly 44.7%, which is usually admitted to be a good ratio. Then the centres of the classes obtained may be represented on the four axes space (cf. Figure 1 and Figure 2 in Annex II).

As Figure 1 shows, the first axis (contributing for 14.1% of the inertia) accounts for the size of labs and other variables, which are structurally associated with scale issues such as institutional recognition and share of international collaborations among publications which have been highlighted to be subject to scale effects. On this axis, the fifth class is mainly opposed mainly to Class 2 and to the other classes to a lesser extent. The second axis (10.3% of the inertia) opposes strongly research-intensive labs with the ones also concerned with teaching and PhD supervision. Along this axis Class 4 (and to a lesser extent Class 1) is opposed to Class 2 (and 5 to a lesser extent). Figure 2 gives the projections of the variables modalities and the classes on the last two axes. The third one (8% of the inertia) opposes highly publishing labs to the ones which are more median in that respect. According to this axis, classes 4 and 3 are especially distant. The last axis retained (6.5% of the inertia) illustrates the distinction between younger, more international more attractive toward post-docs and more open

toward industry to older, less attractive and less connected with firms. This is why we synthetically labelled this axis “openness intensity”.

We discuss the characteristics of the five classes below. Some of our arguments are based on descriptive statistics, which are not presented due to space constraints but can be produced upon request.

5.1 Class 1: The standard research intensive labs

Class 1 involves 22 labs. The laboratories of this class are of a rather small size and “research intensive” since they count on average a high share of full time researchers as compared to university professors. They host only few PhDs and few post-docs, which is quite surprising considering their research-intensive nature. The ratio of permanent researchers over all researchers is high (50% on average). The simultaneous presence of many non-researchers reflects a probable substitution between non-permanent researchers and non-researchers. While the labs have on average fewer sub-disciplines than other labs (probably due their small size), their entropy index is the highest. The average age of permanent researchers is higher than average. Moreover, the share of junior-positions lies below average, which may indicate on average lower abilities of these permanent researchers.

The scientific performance of these labs is higher than the average labs, with 5.96 papers written (in individual contribution) per permanent researcher. The papers are written on a less international basis on average even if permanent researchers have written on average as many internationally co-authored papers as average. The collaboration with industry is below the average. This may help explain that permanent researchers produce on average only 0.12 patent, (one-third of the average). 63% of the labs never patented over the period. The labs belong mainly to the Medicine field, most of the labs involved in Neurosciences, one in Chemistry and one in Bio-Pharmacy. The labs of this class are nearly equally shared between high and low recognition which implies that this class counts a third of the less recognized labs and half of the labs supported by INSERM. This observation led us to compare the characteristics of the highly recognized ones with the other ones. The two noticeable differences were that the recognized ones count more full time researchers, attract more non-permanent researchers and tend to publish substantially more. Nevertheless the two sub-groups are quite similar in all other respects, especially concerning the rather low inventing activity.

5.2 Class 2: The teaching labs in the fields of social and human sciences

This Class gathers 10 labs, among which we find the eight belonging to the Social Sciences, Humanities and Geography (the remaining two belong to Bio-Pharmacy and Biology). Their size is below average. Permanent researchers are mainly university professors (the share of full-time researchers is close to zero). They count many PhD students and nearly no post-docs. The share of females among permanent researchers is higher than the average (28% against 23%).

Eight of the labs are not supported by the funding agencies. The share of authors among permanent researchers is extremely low: on average only 45% of permanent researchers publish. Considering these authors, their average performance is only one-third of the average performance of authors in labs. Only few of the publications are written through international collaborations and the collaborations with co-authors outside academe represent only 1.6% of the publication occurrences.

5.3 Class 3: The non research intensive and industry oriented labs

Class 3 counts 15 labs with a size below average. They are “non research intensive” as we find mainly professors and only few full-time researchers. This explains the presence of many PhDs and few post-docs. Moreover, the share of permanent researchers is higher than average (54% against 47%). Permanent researchers are much younger than the average and the junior-like positions represent 70% of the permanent positions.

The ratio of authors among permanent researchers and the publication performance of authors is also below average: they publish 1.63 papers less than in the average labs. Nevertheless, the proportion of international collaborations is higher than average as well as the share of collaborations with industry (8,7% of the publications of the labs). But this did not favour very much patenting activity since their average inventive performance is still below the general average (only 0.3 patents per permanent researcher). These specificities are probably not related to their disciplinary peculiarities. The labs belong to various disciplines: five labs are active in Bio-Pharmacy, two in Biology, in Physics related fields, Engineering, Medicine and one in Chemistry and in Neurobiology. Nearly half of the labs are supported by the CNRS.

5.4 Class 4: The elite research intensive labs

The 12 labs of Class 4 are of a small size even if slightly bigger than the ones of Class 1 and smaller than the ones of Class 3. The share of full time researchers among permanent researchers is slightly below average. The share of juniors and the average age are both below the average of all labs. This tends to indicate a quite good individual recognition of the permanent researchers. The ratio of permanent researchers is the lowest of all classes (only 38%). This comes from numerous PhD students (nearly as many as in Class 3 but with much fewer professors), twice as many national post-docs and more than three times as many foreign post-docs than in Class 1. With the same number of sub-disciplines as average, the sub-disciplinary entropy is the highest (0,9) but never very high (max at 1,3).

These labs exhibit a strong cohesion: only 2% of the permanent researchers do not publish at all. Permanent researchers publish on average 2.8 papers more than the average, and authors publish 1.6 more paper more than the authors of Class 1 (the other research intensive class). While the share of international collaborations among publication occurrences is below average, the number of international collaborations is clearly the highest compared to all other classes. With 8.7% of the

publication occurrences being co-authored with industrial partners, the number of industrial collaborations per permanent researchers is more than twice the average. This probably contributes to explaining that most labs of this class produce patents (85%) and that they patent nearly twice as much as the average (10 patents invented) and nearly three times as much per permanent researcher (more than one patent per permanent researcher). These labs belong to different disciplines: five belong to Medicine, four belong to Bio-Pharmacy, two to Chemistry, and one to Biology.

5.5 Class 5: The large laboratories

This class covers 22 labs, which may be distinguished by their large size: they have 37 permanent researchers on average. Most of them benefit from an important support from CNRS, which allows them to accommodate more than 20 full-time researchers on average. This is the highest ratio of full-time researchers over permanent researchers among all classes. The ratio of permanent researchers is quite similar to the average, but these labs have fewer PhD students and essentially more foreign post-docs. They also count many non-researchers. Age, share of junior-like positions, share of females, entropy are similar to average.

One should notice that neither the size, nor the important ratio of full-time researchers allowed these labs to be more productive: their patenting and publication performances per permanent researcher are similar to the average (even if the patenting activity is increasing, 20% of the labs do not patent at all). The share of international collaborations is higher than average but that may due to a higher number of co-authors on average. The share of collaborations with industry among publication occurrences is below average. These labs belong to a very large range of disciplines: five in Chemistry, five in Biology, four in Physics related fields, two in Engineering, two in Earth sciences, one in Bio-Pharmacy, Astronomy, Mathematics, and Neurosciences.

6 Discussion and conclusion

This paper offers a first empirical investigation of an original set of data describing the research activity of a large European university. Such a data set allow us to analyse the organization of research at laboratory level. We study the labour force composition of labs, how these various inputs are correlated with the various outputs, namely publication performance, publication occurrences co-authored with researchers belonging either to foreign institutions or to firms, and patents. We also study the output structure of labs in order to stress possible complementarities or crowding out between outputs. Finally we construct a typology of labs, which identifies different types of research organization and their outcomes.

In the correlation studies we observe that permanent and non-permanent researchers are correlated and seem thus to be complements rather than substitutes. At a more disaggregated level, PhDs appeared to be more intensively associated to university professors following primarily a logic of personal contacts and only secondarily a logic of excellence. The latter applies primarily to post-

docs. We also find that the number of non-permanent researchers affects the distribution of outputs among the permanent researchers of the lab: foreign post-docs increased with the kurtosis and the variance of the average publication performance of permanent researchers. That may indicate both that post-docs go to labs in which there are one or a few “stars”, or symmetrically post-docs work preferentially with the most productive researchers, thus specifically enhancing their productivity.

This study underlines scale effects on international collaborations: the number of permanent researchers is correlated to their average number of publications co-authored with international partners. Such a result tends to highlight possible scale effects in the connection with high-level international collaborative networks. We also found a significant impact of national post-docs on the average number of patents invented by permanent researchers of the laboratory. That result could indicate that national post-docs are specifically employed for invention activities. It may come from both their weak autonomy in research agenda selection or from their involvement in a career path turned toward research in industry.

University professors tend to decrease all the scores in average outcomes. Concerning publication, this result seems to be induced by the presence of non-publishing university professors. We have no evidence of a distinctive behaviour between publishing professors and full time researchers. Institutional recognition, which constitutes a proxy for quality, is not correlated with any performance measure. It is only correlated with the share of international collaborations indicating that the former quality proxy is robustly linked across disciplines to collaborative access capacity to the international scientific community.

About the outcome structure, we found quite surprisingly that the share of international collaborations is not associated with a higher average publication performance: those who collaborate more intensively with international co-authors are not necessarily those who publish more. On the contrary, the average performance in international collaborations goes along with a high performance in terms of collaborations with industrial partners. One important result is that the intensity of patenting activity is correlated with all publication intensity measures (both per permanent researcher): strongly with the intensity of publications with industrial partners and weakly with international partners. Moreover, while it grew over the period covered (1993-2000), the patenting activity seems to become more intensively linked to and supported by publication performance.

We also built a typology of labs based on both input and output variables. We have constructed five classes of labs, which illustrate five styles of research organization and production. As expected, a class (Class 2) groups together all the labs belonging to the Social Sciences, Geography and Humanities and another (Class 5) covers the ‘big’ labs. Interesting new insights come from distinguishing labs belonging to classes 1, 3 and 4. Although they have quite comparable sizes, they exhibit different input structures as well as output performances. Class 3 is more oriented toward teaching and PhD supervision while having few post-docs and a below average publication

performance. Class 1 is the more “pure research oriented” with more full-time researchers than university professors among permanent researchers. These labs publish significantly more than the average but their patent performance is only one third of the average. This performance structure contrasts with the labs belonging to Class 4: while their permanent researchers publish even more than the ones of Class 1, they also patent over eight times more. This noticeable performance (best among all classes for the two outputs considered) is obtained by younger and promoted permanent researchers, equally allocated between full-time researchers and university professors, focused on various sub-fields (highest entropy index), allowing them to attract both many PhD students and many post-docs.

The literature review underlined several factors which might influence the productivity of individual researchers: potential conflict between teaching and research activities, department effect on productivity, life-cycle effects, decreasing returns to scale, potential conflict or complementarity between the two outputs (patents and publications). One important outcome of the typology consists in proving the existence of a virtuous combination of factors leading to both high publishing and high patenting performances. Indeed, some labs (Class 4) are by far the most productive in terms of publication performance and (even more) in terms of patents. These labs exhibit an interesting combination of inputs: permanent researchers are specialized in different sub-disciplines (highest interdisciplinary entropy), and equally split between professors and full time researchers, who attract many non permanent researchers.

The results obtained by both the correlation studies and the typology tend to show that an appropriate combination of different types of research personnel has a positive impact on the research productivity of the lab. Equal share of full time researchers and professors allows maintaining incentives for the latter to perform research. Moreover, the full time researchers increase notably the research performance of the lab and attract post-docs. University professors increase the number of PhD candidates. That result highlights the often-ignored influence of non-permanent researchers on publication and patenting performances. It is certainly one of the most important insights of our study to show that policy makers and managers of laboratories and universities should design strategies that explicitly take into account their ability to attract the growing population of non permanent researchers. In the emerging European research space, the attractiveness of labs for young researchers, who constitutes a low cost and highly motivated human resource, may well make the difference in terms of performance.

7 Acknowledgements

This work is part of a larger project on knowledge production at ULP developed by a team of researchers at BETA led by P. Llerena. We are grateful to him and to all other members of the team whose support was crucial at some points: Joaquin Azagra-Caro, Monique Flasaquier, Caroline Hussler, Rachel Levy, Thuc Uyen Nguyen Thi, Musa Topaloglu, Panos Vassiliadis, and Sandrine

Wolff. Acknowledgements should extend to the administrative departments of ULP, the technology transfer office (ULP-Industrie), and to the French patent office (INPI) for allowing us to get access to their databases. A. Bonaccorsi (Santa Anna School of Advanced Studies, Pisa) and F. Lissoni (CESPRI, Università Bocconi, Milano) should be acknowledged for their help in data collection. Some technical assistance was also provided by P. Roux. A preliminary version of this paper was presented at the 3rd EMAEE (European Meeting on Applied Evolutionary Economics) Conference, “The Knowledge-Based Economies - New Challenges in Methodology, Theory and Policy”, Augsburg, Germany, April 10 - 12, 2003.

8 References

Adams J., Griliches Z., 1996, “Measuring science: an exploration”, Proceedings of the National Academy of Science, vol. 93, pp. 12664-12670

Adams J., Griliches Z., 2000, “Research productivity in a system of universities”; in Encaoua D., Hall B., Laisney F., Mairesse J. (eds), *The Economics and Econometrics of Innovation*, Kluwer Academic Publishers, pp. 105-140.

Arora A., David P.A., Gambardella A., 1998, “Reputation and competence in publicly funded science: Estimating the effects on research group productivity”, *Annales d’Economie et de Statistique*, no. 49/50, pp. 163-198.

Allison P.D., Long J.S., 1990, “Departmental effects on scientific productivity”, *American Sociological Review*, vol. 55, pp. 469-478

Benzécri, J.-P., 1992, *Correspondence Analysis Handbook*, Marcel Dekker, New York.

Bonaccorsi A., Daraio C., 2002, “The organization of science. Size, agglomeration and age effects in scientific productivity”, NPRnet Conference, “Rethinking Science Policy”, SPRU, Brighton, march 21-23.

Carayol N., 2003a, “An economic theory of academic competition: Dynamic incentives and endogenous cumulative advantages”, Mimeo, BETA, University Louis Pasteur, Strasbourg.

Carayol N., 2003b, “Objectives, agreements and matching in science industry collaborations: Reassembling the pieces of the puzzle”, *Research Policy*, vol. 32, n°6, pp. 887-1148.

Cole S., Cole J., 1973, *Social stratification in Science*, University of Chicago Press, Chicago.

Coupé T., 2001, “Science is golden: academic R&D and university patents”, proceedings of the conference ‘Innovations and Intellectual Property: Economic and Managerial Perspectives’, Brussels, November 22-23.

Crow M., Bozeman B., 1987, "R&D laboratory classification and public policy: The effects of environmental context on laboratory behavior", *Research policy*, vol. 16, pp. 229-258.

David P.A., 1994, "Positive feedbacks and research productivity in science: Reopening another black box", in Granstrand O. (ed.), *Economics of Technology*, Elsevier, Amsterdam.

Dasgupta P., David P.A., 1994, "Toward a new economics of science", *Research Policy*, vol. 23, no. 5, pp. 487-521.

Diamond A.M., 1986, "The life-cycle research productivity of mathematicians and scientists", *The Journal of Gerontology*, vol. 41, no. 4, pp. 520-525.

Foltz J., Barham B., Kim K., 2000, "Universities and agricultural biotechnology patent production", *Agribusiness*, vol. 16, pp. 82-95.

Fox M.F., 1992, "Research, teaching and publication productivity: mutuality versus competition in academia", *Sociology of Education*, vol. 65, pp. 293-305.

Garcia-Romero A., Modrego A., 2002, "Research training in Spain", An assessment exercise", mimeo.

Greenacre, M., 1993, *Correspondence Analysis in Practice*, Academic Press, London.

Henderson, R., Jaffe, A.B., Trajtenberg, M., 1998, "Universities as a source of commercial technology: a detailed analysis of university patenting, 1965-1988", *Review of Economics and Statistics*, vol. 80, pp. 119-127.

Joly P.B., Mangematin V., 1996, "Profile of public laboratories, industrial partnerships and organisation of R&D: The dynamics of industrial relationships in a large research organization", *Research Policy*, vol. 25, pp. 901-922.

Laredo, P., Mustar P., 2000, "Laboratory activity profiles: An exploratory approach", *Scientometrics*, vol. 47, no. 3, pp. 515-539.

Levin S.G., Stephan P.E., 1991, "Research productivity over the life cycle: Evidence for academic scientists", *American Economic Review*, vol. 81, no. 1, pp. 114-132.

Long J.S., 1978, "Productivity and academic position in the scientific career", *American Sociological Review*, vol. 43, pp. 880-908.

Long J.S., McGinnis R., 1981, "Organizational context and scientific productivity", *American Sociological Review*, vol. 46, pp. 422-442.

Mairesse J., Turner L., 2002, "A Look at Individual Differences in Scientific Research Productivity: An econometric analysis of the publications of the French CNRS physicists in condensed matter (1980-1997)", presented at the conference 'Rethinking Science Policy: Analytical Frameworks for Evidence-Based Policy', SPRU, University of Sussex, Brighton, 21-23th March.

Merton R.K., 1968, "The Matthew effect in science", *Science*, vol. 159, no. 3810, pp. 56-63.

Payne A.A., Siow A., 1999, "Does federal funding increase university research output", IGPA Working Paper 74.

Stephan P., 1998, "Gender differences in the rewards to publishing in Academe: Science in the 70's", *SEX Roles*, vol. 38, 11-12.

Stephan P.E., 1996, "The economics of science", *Journal of Economic Literature*, vol. 34, pp. 1199-1235.

Stephan P.E., Gurmu S., Sumell A.J., Black G., 2002, "Individual patenting and publication activity: having one's cake and eating it too", NPRnet Conference, "Rethinking Science Policy", SPRU, Brighton, march 21-23.

Stephan, P.E., Levin S.G., 1997, "The critical importance of careers in collaborative scientific research", *Revue d'Economie Industrielle*, no. 79, pp. 45-61.

Wallmark, J.T., 1998, "Innovations and patents at universities: the case of Chalmers University of Technology", *Technovation*, vol.17, n°3, pp. 127-139.

Weiss Y., Lillard L.A., 1982, "Output variability, academic labor contracts, and waiting times for promotion", in R.G. Ehrenburg (ed.), *Research in Labor Economics*, vol. 5, pp. 157-188.

Zuckerman H.A., Merton R.K., 1972, "Age, aging, and age structure in science", in M.R. Riley, M. Johnson et A. Foner (eds.), *A sociology of age stratification: Aging and society*, vol. 3, reprinted in R.K. Merton, 1973, *The sociology of science*, N.W. Storer (ed.), University Chicago Press, Chicago, pp. 497-559.