The Spread of Academic Invention: A Nationwide Case Study on French Data (1995-2012)

Nicolas Carayol and Elodie Carpentier *

March 9, 2021

Abstract

Although numerous public policies have been introduced to incentivize scholars and researchers employed in universities and public laboratories to generate and transfer inventions, the extent and drivers of any spread in patenting behavior within the academic community have not yet been fully documented. We propose a nationwide empirical investigation of patented academic inventions in France over nearly two decades which offers a number of new insights. Firstly, the direct contribution of academia to the nation's flow of patented inventions is revised upwards, up to eleven percent of all patented inventions. Secondly, we show that patenting behavior is more pervasive in the academic community than expected with one in five professors or researchers having invented at least one patent in nearly all fields of hard and life sciences. Thirdly, even if academic patenting was strong before the 1999 reform favoring technology transfer, the propensity of professors and researchers to invent has significantly increased over the subsequent period. Fourthly, cohort and age effects cannot fully explain such behavioral change whereas local diffusion, in particular within labs, is found to be an important driver of the spread of academic patenting.

JEL Classification Codes: C81; I23; O31; O33; O34.

Keywords: university; technology transfer; academic patenting; disambiguation; peer effects.

*Nicolas Carayol: GREThA, CNRS and Department of Economics, University of Bordeaux, Pessac 33608, FRANCE. Elodie Carpentier: GREThA, CNRS and Department of Economics, University of Bordeaux, Pessac 33608, FRANCE. Emails: nicolas.carayol@u-bordeaux.fr and elodie.carpentier@u-bordeaux.fr. We would like to thank the ANR (UTTO Project, no ANR-15-CE26-0005) for its financial support and the Fulbright Commission for its support. N. Carayol also thanks the Bordeaux IdEx (no ANR-10- IDEX-03-02) for its support.

1 Introduction

Economists have long hypothesized that the ever-increasing stock of scientific knowledge has a huge impact on innovation and the pace of economic growth (Nelson, 1959; Arrow, 1962; Romer, 1990; Jones, 1995). Focusing on the direct contribution of academia to innovation, many pieces of public policy have been introduced around the world to encourage scholars to generate inventions and support their transfer to society. In this paper, we document the extent to which professors and researchers engaged in academic patenting in France over nearly two decades (1995-2012) and explore the factors leading them to do so.

We define academic patents as being those patents invented by professors and researchers employed in universities and research institutes. The nice feature of such a definition is that it is independent of the patent assignee and thus immune to the transfer strategies of the professors and universities. The flip side is that this definition is much more demanding in terms of data collection and data treatments. Rosters of professors and researchers need be collected and matched with patent inventors. The main originality of our approach with respect to previous attempts[1] is that we systematize the matching and filtering procedures on a large scale thanks to machine learning techniques that avoid time-consuming, painstaking checking procedures performed by humans. Our method requires a reliable benchmark though, to ensure that false positives and negatives are fully controlled and limited. This approach makes it possible to consider i) large lists of professors and researchers which become comparable to the reference population, and ii) sufficiently large time windows. We apply this method to France, the seventh country in the world in terms of GDP, sixth for scientific articles, and fourth for patents granted in this period of time.

Over the eighteen-year period under scrutiny, academic patents are found to account for more than 11% of all patents invented in the country. This is well above prior estimates and therefore provides a new insight into the real direct contribution of science to technological inventions. As our data provide interesting covariates on professors and researchers extending beyond those who might have a patent, we are able to characterize their involvement in technology transfer. We find that more than one in five professors and researchers is an inventor (excluding social sciences and humanities). This statement applies to nearly all fields in the hard and life sciences, meaning that academic patenting is not specific to a particular field of science. Obviously, faculty members do not operate in some "ivory tower" and are much more directly involved in technological invention than is often assumed.

[1] Meyer (2003); Lissoni et al. (2008); Thursby, Fuller and Thursby (2009).

Is this a recent phenomenon potentially entirely due to increasing incentives to patent and commercialize academic research? We find that academic patents already accounted for more than 9% of all patents invented in the country prior to the introduction of the first piece of legislation encouraging technology transfer –the 1999 Innovation Act.[2] This contradicts the idea of a very low pre-reform level of academic technology transfer, although it does not imply that nothing has changed in the more recent period. On the contrary, we find that faculty members' propensity to invent increased by 75% between 1995 and 2012. To control for a potential trend affecting patenting behavior (improvements in communication technology or instrumentation, for instance), we use non-academic patents as a reference point and show that academic inventors increase their propensity to invent significantly more than non-academic inventors over the same period.

What are the drivers behind the spread of patenting behavior in academia? We consider two series of factors: individual attributes on the one hand, and social and cultural influence on the other. Controlling for a large number of potential confounding factors, such as university, age, gender, status, field and year effects, we find that more recent cohorts are not more likely to engage in patenting. Age plays positively on academic patenting at the individual level, a result which is reminiscent of previous findings on smaller datasets (Carayol, 2007; Stephan et al., 2007) and consistent with the idea that incentives to invent are less susceptible to decrease over the life-cycle than traditional incentives to publish. However, the propensity to patent of professors is shown to increase significantly over time, controlling for age and cohort effects.

To further understand how patenting behavior spreads through the academic community, we explore the influence of local social factors. The organizational culture at the individual university level has been emphasized as key to the willingness of faculty members to engage in entrepreneurship (Grimaldi et al., 2011). Other studies have highlighted the importance of norms, role models and peer effects in the research group in explaining faculty engagement in technology transfer (Louis et al., 1989; Bercovitz and Feldman, 2008; Krabel and Schacht, 2014). We proxy community involvement towards patenting using per capita invention rates in the previous years, at the university level (excluding the focal lab staff) and within the laboratory (excluding the focal person). Controlling for year, university and individual fixed effects, we find that local diffusion within the lab plays a decisive role. One additional yearly

[2] The Innovation Act, or Allegre Law, voted in 1999, introduced the possibility for universities to open a technology transfer office called SAIC. These structures were in charge of the management of research contracts, patenting and licensing activities and the commercialization of the outcomes of professors and researchers activities.

patent invented by the average colleague in the lab in previous years raises the expected number of patents by a factor of four.

The rest of the paper is organized as follows. Section 2 reviews the literature on the extent of academic patenting and its drivers. Section 3 exposes data collection and the methodology. Section 4 proposes descriptive statistics on academic patenting in France. Section 5 assesses how the propensity to invent varied over the period in academia, as compared to non-academic inventor profiles. Section 6 discusses the factors explaining the spread of patenting behavior within the academic community. Section 7 concludes.

2 Literature on academic patenting

2.1 The extent of academic patenting

At least since 1980, with the passage of the Bayh-Dole Act in the US[3] attracting policy makers' and researchers' attention to this phenomenon, the number of patents invented by professors and researchers has been increasing over time in most advanced economies (Henderson, Jaffe and Trajtenberg, 1998; Mowery and Ziedonis, 2002).

The early 2000's data on university patenting pointed to an underperformance of UK and European universities as compared to their US counterparts, given the nations' investment in fundamental research (Geuna and Nesta, 2006). Based on this prior and the belief that the growth in US university patenting was driven by the Bayh-Dole Act, several European countries voted similar reforms (Mowery and Sampat, 2005).

But counting university patents hardly reflects the true contribution of academia to innovation in Europe (Geuna and Nesta, 2006). Considering university-owned patents leads to bias downward the estimations of academic patenting because of cultural, regulatory and managerial differences regarding the attribution of intellectual property rights of professors and researchers employed in universities and public research institutes. Whereas US universities may strongly enforce the ownership of their inventions, their European counterparts often left the rights to their professors (the so-called "professor's privilege" in German and Nordic countries), or simply did not have the resources to manage it, and thus inventions were owned by the private sector or the individual inventor herself (Lissoni et al., 2008; Thursby, Fuller and Thursby, 2009). International comparisons become meaningless simply because

[3]The Bayh-Dole Act allowed university to retain the intellectual property rights over inventions stemming from federally funded research.

data are contaminated by the choice of a transfer strategy. Carayol and Sterzi (2021) even show that promising technologies are likely to be cherry-picked and transferred without the involvement of the university TTO.

In such a context, an appropriate methodology for measuring the contribution of academia to invention needs to find and count in those patents invented by professors employed at universities and not assigned to their institution.Following this methodology, in what is probably the most comprehensive national study, Meyer (2003) found that in Finland, for the years 1986-2000, 8% of national patents stem from academia. Another great advantage of such an approach is that it allows the analyst to appreciate the engagement of professors and researchers in patenting. For instance, Lissoni (2012) find that for samples of academics in six European countries, 2 to 5% of professors have filed a patent. The Research Value Program provided a similar figure for the US, with about 5% (Bozeman and Gaughan, 2007). From their literature review, Perkmann et al. (2013) conclude that the proportion of academics involved in patenting ranges from 5% to 40%, depending on the sampling strategy used.

Prior figures are based on samples of limited size, often of a cross-sectional nature, and most of the time cover a rather short time-span. In this paper, we introduce a methodology (Section 3) to build large datasets of academic patenting. We use it to set up a longitudinal dataset covering all scientific fields at a national level. Among other benefits, this allows us to extend the above mentioned lines of investigation in a more systematic way.

2.2 The drivers of academic patenting

Even though the extent of academic engagement was not yet fully appreciated, considerable concern was expressed in the 90's that the move towards commercialization in the university community may be coming at the expense of the production of fundamental knowledge (Dasgupta and David, 1994; Stephan and Levin, 1996). A stream of work investigated the nature of the relationship between patenting and publishing (Stephan et al., 2007; Azoulay, Ding and Stuart, 2007; Carayol, 2007; Czarnitzki, Glänzel and Hussinger, 2007), and concluded that both activities are complement rather than substitutes.

Once this concern was essentially dismissed, scholars became interested in the drivers of academic patenting. In the US, the positive trend in academic patenting was first attributed to the Bayh-Dole Act, although Mowery et al. (2001) suggest that the growth in federal financial support for basic biomedical research and the increased patentability in this field may better explain this positive trend than the Act itself. Conversely and on a similar level, Ejermo and Toivanen (2018) and Hvide and Jones (2018) find that the end of the Professor's Privilege had a negative impact on academics' propensity to invent.

At the individual level, the type of research performed, the degree of collaboration with the private sector as well as relevance of intellectual property rights protection varies greatly among scientific fields, making the discipline another important driver to consider (Stephan et al., 2007; Carayol, 2007). Tenure was also found to be relevant in the US context (Azoulay, Ding and Stuart, 2007; Stephan et al., 2007). The literature also identifies age as a key characteristic in explaining academic productivity. The life-cycle effect indicates that a scientist productivity grows up to a certain (biological) age before decreasing towards the end of the career. It is usually observed in scientists' publishing patterns, but also in their patenting activity. Carayol (2007) finds that patenting increases with age whereas Stephan et al. (2007) finds little evidence of such an effect. Those results need to be taken with caution as the cross-sectional nature of the data prevented these studies from accounting for cohort effects. Different generations of scientists may have a different productivity pattern because of the varying contexts in which they were trained and are working (Stephan, 2010).

Based on longitudinal data, Azoulay, Ding and Stuart (2007) and Thursby and Thursby (2007) find that, once controlling for cohorts, patenting decreases over the life-cycle. However, while the authors of the former paper argue that newer cohorts are more likely to patent than are earlier cohorts, the latter finds opposite results. We contribute to this debate by studying the age and cohort effects on our longitudinal dataset covering all scientific fields and almost all universities at a national level.

Finally, the literature on academic patenting has emphasized the role of the local culture within the university site and peer effects in embracing a research style that considers innovation and entrepreneurial attitudes (Grimaldi et al., 2011). The entrepreneurial culture in some university campuses (such as MIT, Stanford or the University of Wisconsin at Madison) is often highlighted as critical. The literature also suggests that besides the campus atmosphere, professors influence each other in their immediate work environment. Based on a survey of US life science faculty, Louis et al. (1989) first highlight the importance of "local group norms" in predicting active involvement in commercialization. Krabel and Schacht (2014) highlight the influence of Max Plank research institute leaders in disclosing inventions. Considering peer effects in faculty engagement in technology transfer activities, Bercovitz and Feldman (2008) show that faculty members of two US medical schools were more likely to disclose inventions when their peers did so in the previous year.

We investigate the influence of university recent involvement in patenting as well as similar effects within research labs.

3 Identifying academic inventions

In this section, we first discuss the different approaches to identifying academic inventions. We next present our data sources, before exposing our filtering methodology to merge faculty lists with inventors. Lastly, we show how we can estimate the number of academic inventions.

3.1 Academic inventions in the literature

Previous studies of inventions produced in academia have relied on a variety of definitions of what an academic patent is, and on associated data collection methodologies. Scholars initially assumed that academic patents were patents assigned to universities and government labs (Mowery et al., 2001; Mowery and Ziedonis, 2002), but this approach had the drawback of ignoring all patents invented by university personnel but which were not assigned to the university, for whatever reason. Many academic institutions traditionally did not manage their intellectual property rights and thus often did not retain the rights to the inventions their staff were involved in, whether intentionally or unintentionally. To avoid this issue, reference must be to the inventor field rather than to the assignee.

Several strategies can be deployed for the inventor information to find university research personnel. Some papers use the title "Prof. Dr." that may be mentioned in the inventor field (Czarnitzki et al., 2016), although this is barely feasible outside Germany. When national statistical institutes provide precise employee data, authors merge them with inventors (Ejermo and Toivanen, 2018). Another way is to merge authors of scientific publications with inventors (Stephane and Martinez, 2014).

Our approach started out with information on the research staff of universities. Several previous studies have used such lists (Meyer, 2003; Balconi, Breschi and Lissoni, 2004; Iversen, Gulbrandsen and Klitkou, 2007; Lissoni et al., 2008; Thursby, Fuller and Thursby, 2009; Hvide and Jones, 2018). The difficulties in this approach are i) collecting large research staff lists over a sufficiently long period of time and ii) performing a reliable and systematic merge of those persons with inventors. In this paper we use large lists of professors and researchers and develop a filtering procedure which simultaneously avoids performing time-consuming manual checking and controls for merge quality.

3.2 Data sources

Our data come from the French Ministry for Higher Education and Research. We know, each year, the full name, gender, age, field of science, employing university and status of all the professors employed in France. Besides, we know precisely in which laboratory each person works. This is important because the French research system, as in most continental European countries, is organized in research laboratories (see Carayol and Matt, 2004; Azagra-Caro, Llerena and Carayol, 2006). These laboratories are the elementary units structuring research activity in nearly all higher education and research institutions. In France, labs host both professors employed by the universities (or higher education schools) and researchers employed by research institutes. They vary in size from a few professors and researchers to several hundreds. Lab information is available as labs provide their staff lists when surveyed by the ministry every four to five years. As most labs were surveyed this way at least twice over the period, we can observe movements in time. At the end of this task, we had nearly 52,000 faculty members and researchers affiliated to 234 universities and research institutes and more than two thousand laboratories.[4]

Patent data were extracted from the "EPO Worldwide Patent Statistical Database" (PATSTAT Autumn 2017 Edition). We restricted the data to all patents filed at EPO or INPI (the French national office) for which at least one inventor had a home address in France. We obtained 428,000 French-invented patents from 1995 to 2012.

3.3 Filtering academic patents

Professors or researchers' first and last names are first matched with those of the inventor (using exact and fuzzy matching techniques to allow limited variation in spelling). This returns almost 91,000 patents and 148,000 professor-inventor pairs on a given patent that remain to be filtered out. We use a statistical model to estimate the probability that each match is correct. Our filtering process takes four stages.

In the first stage, we estimate a logit model on a set of validated and unvalidated couples. Such a benchmark was already used in Carayol et al. (2019) and was constituted on the basis of experts (mainly professionals of technology transfer employed in the universities) identifying professors as potential inventors. The benchmark is made up of almost 1,200

[4] Carayol and Lanoë (2017) used a very similar dataset to estimate the impact of project-based funding.

professor-inventor pairs.[5] Explanatory variables include Jaccard similarity between names, the inventor name frequency (in log), the distance between the patent technological classification and the professor's scientific disciplines as defined by Magerman et al. (2017) (in log), as well as dummies signaling consistency between the professor's age and the patent application year and between the assignee's name and the professor's employing institution. Regressions are performed per patent office as Hausman tests showed that logit coefficients are significantly different across offices. Results for each patent office are presented in Table A1 in the Appendix.

The second step uses the estimated coefficients to predict the probability that potential matches are correct or incorrect over the whole reference population.

In the third step, we consider various thresholds of the probability of accepting or rejecting matches. Let $TP(p)$ denote the number of true positives in the benchmark for a given threshold probability value p, $FP(p)$ is the number of false positives, and $FN(p)$ the number of false negatives. We compute precision as

$$P(p) = \frac{TP(p)}{FP(p) + TP(p)}, \tag{1}$$

and recall as

$$R(p) = \frac{TP(p)}{TP(p) + FN(p)}. \tag{2}$$

Precision and recall vary in opposite directions with threshold p. We thus calculate a synthetic indicator taking both into account:

$$F_\beta(p) = (1+\beta^2) \times \frac{P(p) \times R(p)}{\beta^2 \times P(p) + R(p)}, \tag{3}$$

with β, a strictly positive parameter weighting precision and recall. If $\beta < 1$, precision gets a lower weight than recall, whereas the reverse holds when $\beta > 1$. As we do not want our results to be sensitive to a particular value of β, all our statistics are computed for $\beta = 0.5$, $\beta = 1$ and $\beta = 2$. In Figure A1 in the Appendix, we display the computed values of those indicators for the different threshold probability values p.

[5] This benchmark includes 31 pairs for the USPTO, 236 for the EPO and 945 for the INPI. There were too few USPTO pairs for the algorithm to perform well on this subset so we removed them from the data.

The fourth and last stage consist in finding the optimal p threshold value, for each β and patent office i. That is, we want to find

$$p^*_{\beta,i} = \arg\max_p \left\{ F_{\beta,i}(p) \right\}, \tag{4}$$

for all β and i, with $F_{\beta,i}(p)$ the indicator defined in Equation 3, but calculated using the patents of office i only. Given that we consider two offices and three different values of β, we end up with a series of six optimal threshold values to be calculated. Optimal thresholds presented in Table A2 are significantly different for each considered office. Table A2 also indicates the precision and recall values for each pair (β, i). We compute these indicators on the benchmark pooling patents from each office. As expected, recall increases with β whereas precision decreases with β. EPO and INPI patents have very good recall rates (above 0.93) when $\beta = 2$. INPI patents have a satisfactory recall rate (above 80%) when β equals 1. Recall and precision rates for EPO patents are found to be simultaneously satisfactory when β equals 1 or 0.5.

For an external assessment of the quality of our data filtering, we use a more limited sample of faculty inventions which have been identified via a combination of web searches, emails and phone calls.[6] Having created faculty-inventor-patent tables in both datasets and excluded homonyms born in the same year, those tables are merged on prof name, first name, birth year and patent identifier code. This essentially leads us to restrict data to the benchmark, obtaining 1,016 faculty-inventor-patent combinations that were present in both datasets, involving 461 distinct scientists inventing 787 distinct patents. Interestingly, the filtering assessment maximizing F_β when $\beta = 2$ on the external dataset led to a 80.7% precision rate and a 83.2% recall. Filtering when $\beta = 1$ or 0.5 led to significant but limited gains in terms of precision at the price of a larger decrease in terms of recall.

This assessment was partially consistent with our own benchmarking exercise on EPO patents: both led to satisfactory recall rates when $\beta = 2$. They diverged slightly, however, with respect to precision. As it is better in principle to rely on external sources to appreciate the quality of a parametrization optimized on a given training set, we will use EPO patents preferentially in the rest of the paper, when applications at another office is not necessary. This renders comparison with other studies easier and rules out issues concerning institutional differences among patent offices. We will also restrict our sample to EPO patents validated according to $\beta = 2$. This sample performs well on both benchmarks in terms of recall. As our benchmark suggests that a $\beta = 1$ would improve precision significantly, we

[6]We thank F. Lissoni for giving us access to those data (Lissoni et al., 2008).

performed robustness checks of all our results with this parametrization on EPO patent applications. They are available upon request from the authors, as are robustness checks on INPI applications, patent families and the total number of patents at the two offices (EPO and INPI).

3.4 Estimating the number of academic patents

By definition, a patent is academic if at least one of its inventors is an academic staff member. In our framework, this translates as a patent is academic if at least one of its professor-inventor pairs (if any) has a probability of being a correct match above threshold $p^*_{\beta,i}$. Let $N^1_{\beta,i}$ be the set of these validated patents, the cardinal of that set is $n^1_{\beta,i}$ and the number of candidate but non-validated patents is $n^0_{\beta,i}$. Assuming that $n^1_{\beta,i}$ reflects the expected number of academic patents would be slightly misleading, as some patents counted in the underlying set ($N^1_{\beta,i}$) are considered as such (because of false positives) while some patents in the complement set ($N^0_{\beta,i}$) are also misallocated (because of false negatives). We can however use our own estimations of errors in both directions to correct those numbers and obtain a consistent estimation of the number of academic patents as follows:

$$\hat{x}_{i,\beta} = n^1_{\beta,i} \times \frac{TP(p^*_{\beta,i})}{FP(p^*_{\beta,i}) + TP(p^*_{\beta,i})} + n^0_{\beta,i} \times \frac{FN\left(p^*_{\beta,i}\right)}{FN\left(p^*_{\beta,i}\right) + TN\left(p^*_{\beta,i}\right)}, \tag{5}$$

for all $\beta \in \{0.5,\ 1,\ 2\}$, $i \in \{\text{EPO, INPI}\}$. We multiply the number of already validated academic patents by precision rate (true positives among positives), and add this number to the number of rejected patents multiplied by the rate of false negatives among negatives. This leads to Table A3 in the Appendix.

We make another correction to those numbers because the data cover only universities and public research organizations recognized by the Ministry for Higher Education and Research. Some higher education or research institutions are funded by other ministries (defense, industry, agriculture) and not at all by Ministry for Higher Education and Research. Table 1 reports the new calculations. The gain from this correction is significant in our case. This tells us that our estimation of academic patenting will still be an underestimation as we are missing all the patents invented by professors and researchers who are not in our list or whose assignee is not an academic institution.

We can see in the table that the numbers obtained with different weightings of precision and recall (different values of β) actually provide very similar numbers, ranging from 44,759 to 45,637 academic patents over the period. The fact that those numbers are very close is reassuring in that the estimations are largely unaffected by the weightings of precision and recall.

Table 1: Expected number of academic patents for several β values (from 1995 to 2012) – Alternative method.

Office	$\hat{x}'_2$		$\hat{x}'_1$		$\hat{x}'_{0.5}$		All French-invented patents
EPO	19786	(11.1%)	21034	(11.8%)	21039	(11.8%)	177286
INPI	24973	(10%)	24604	(9.8%)	24273	(9.6%)	250605
Total	44759	(10.5%)	45637	(10.6%)	45312	(10.6%)	427891

Notes:
– For $i = 1, 2, 0.5$ we have $\hat{x}'_i = \hat{x}_i$ + all patents that did not match on names and are owned by French universities and public research organization (exclusively or in joint ownership with companies).
– This table displays fractional counts.
– The shares of academic patents - by office and overall - over all patents invented in France are placed into parentheses.

4 Academic patenting in France

In this section, we provide descriptive statistics on academic patenting activity in France. Firstly, we discuss the strength and specialization of academic patenting with respect to overall patenting in the country. Secondly, we describe the strength of patenting activity in the academic community.

4.1 Strength and specialization of academic patents

Table 1 shows that academic patents represented 11.1% of all patented inventions generated in France between 1995 and 2012[7]. Although it should be remembered that this is a floor value (as some academic patents owned exclusively by the private sector are still missing for the reasons mentioned above), it is way above previous estimates which reported that 3.4% of EPO patents from 1995 to 2001 stemmed from academia (Lissoni et al., 2008). Restricting our analysis to a similar period (1995-2002), we estimate the share of academic inventions in

[7]Recall that for the rest of the paper, we only interpret statistics and results for EPO patents validated according to $\beta = 2$ (see Subsection 3.3).

France to be as much as 9.3%[8]. Academic inventions were thus much greater than previously estimated by a factor more than 2.5.

Interestingly, even before the introduction of the Innovation Act in 1999, universities, higher education schools and public research institutes already contributed 9.1% of all patents invented in the country.[9] This number is much larger than expected and this is important in that policy reforms aiming to develop the university ownership model were introduced in France and in other European countries (Geuna and Nesta, 2006; Verspagen, 2006) on the prior that technology transfer was weak in return for the investment made by the nation in fundamental research. It would therefore appear that this prior had no empirical foundations.

Let us now consider the technological specialization of academia, as compared to the country. The first two columns of Table 2 give the number of academic inventions broken down by technology fields. The third and fourth columns provide the same information for all French-invented patents. The fifth column is the absolute specialization of academia in the different technology fields, whereas the sixth column displays the relative specialization (sometimes called "revealed technological advantage") of academia as compared to national invention. Academia is up to 51% less specialized in fixed constructions and 45% more specialized in chemistry and metallurgy than France. To a lesser extent, academia is more specialized than the country in the physics and electricity fields (respectively 13% and 5% more specialized). In all other technology classes, academia shows a technological disadvantage.

Later in this article, we investigate trends in academic patenting and their drivers. A potential confounding factor among these drivers is the specialization of academic patenting. For instance, a positive variation in academic patenting may be due to academia being specialized, or even increasing its specialization, in technological fields that are growing more rapidly in general.[10] Table A4 in the Appendix provides descriptive statistics based on the finer-grained classification in 35 technological sub-classes. The first column gives the RTA of academia in each sub-class. The second presents the compound annual growth rate (CAGR) of the revealed technological advantage of each class over the period 1995-2012. The third column (Growth ratio) gives the ratio between the annual growth rate of patents in that sub-class relative to the growth rate of patents in France. The last column (Share) displays the proportion of patents that fall in the corresponding class. Technological sub-classes are

[8] This calculation was performed separately from Table 1 and thus does not appear on it.

[9] Separate calculation based on EPO patents for the period 1995-1999. This share is similar on patents from other offices.

[10] Using a similar argument, Mowery et al. (2001) suggested that the growth in federal financial support for basic biomedical research and the increased patentability in this field may explain the positive trend in US academic patenting (rather than the Bayh Dole Act).

Table 2: Distribution of French patents and academic patents, by technology class (1995-2012).

Technology class	Academic patents # (A)	%	All patents # (B)	%	A/B	RTA
Human necessities	3,358	(17.2%)	30,458	(17.2%)	11%	1
Performing operations; transporting	1,936	(10%)	28,674	(16.2%)	6.8%	0.61
Chemistry; metallurgy	4,039	(20.7%)	25,367	(14.3%)	15.8%	1.45
Textiles; paper	210	(1.1%)	2,231	(1.2%)	9.3%	0.86
Fixed constructions	353	(1.7%)	6,610	(3.7%)	5.3%	0.49
Mechanical engineering; lighting; etc.	1,493	(7.6%)	17,428	(9.8%)	8.6%	0.77
Physics	4,626	(23.7%)	36,813	(20.7%)	12.6%	1.13
Electricity	3,432	(17.6%)	29,533	(16.7%)	11.6%	1.05
Total	19,449	(100%)	177,114	(100%)	11%	1

Notes:
- For a technology class i, the revealed technological advantage is RTA$= \frac{A_i}{B_i} \times \frac{\sum_i B_i}{\sum_i A_i}$.
- A patent may belong to more than one technology class so we use fractional counts.

ordered in decreasing order relative to the third column, which basically tells us how dynamic the sub-class was in France over the period. The fields that are growing faster than the national average (ratio greater than 1) are listed above the intermediate horizontal line. Of those 14 fast-growing sub-fields, academia has a technological specialization in only 6 of them (RTA greater than 1). Of those 6 sub-fields, academia is reinforcing its specialization in only 3 of them (positive CAGR). It is true that academia is strongly specialized in "Micro-structural and nano technology" (RTA of 5.4) which is also the most dynamic sub-field (growth ratio of 5.5), but this sub-class gathers only 0.2% of all patents. We can thus conclude that the technological specialization of academia may not explain a positive variation in overall academic patenting.

4.2 Who is patenting in academia?

In the previous subsection, we looked at the importance of academia with respect to all national inventions. Let us now reverse the viewpoint to investigate how important patenting is for academia, and who is participating. Since the goal here is to characterize the professors who invented at least once and the analysis does not relate to patent characteristics, we consider patents from all patent offices. To appreciate to what extent professors and researchers are concerned by invention, we calculate the share of inventors among professors and researchers[11]. Table 3 displays this information by scientific discipline (social sciences and humanities are not considered here). There are two important and somewhat surprising insights that can be drawn from this table.

Firstly, the share of professors and researchers who have been involved in patenting (and thus in technology transfer activities) is significantly high, equal to 22.3%. This means that more than one professor or researcher in five invented at least one patent between 1995 and 2012. We would like to be sure we are not overestimating participation by not being conservative enough in the filtering procedure. Giving too much importance to recall may result in randomly accepting too many patents and therefore wrongly considering many professors and research as inventors. To check for this potential bias, we put more weight on precision and less on recall, using $\beta = 0.5$ at the filtering stage, and verified that it did not significantly alter the results (Table A5 in the Appendix). According to this specification, the share of professors and researchers who invented at least one patent equals 21.8%, which is still large and very close to the main result. When more weight is given to recall over precision ($\beta = 2$ in Table A6 in the Appendix), the share of inventors remains very close (24.5%). Overall, this shows that participation shares are affected by the filtering stage, but to a limited extent which does not modify the results qualitatively. Besides, note that the recorded share of inventors among professors is likely to be less than the share of professors and researchers who have ever invented a patent, as some of those who did not invent over our period may invent in the future or may have invented before 1995 (and are thus not considered here).

The second main insight is that the share of inventors in academia is high in almost all disciplines in the life and hard sciences. Professors and researchers in chemistry are the most active, with a share of one inventor for three professors and researchers, and the observed rates in fields such as physics and medicine are above 25%. Even in mathematics, more than one professor or researcher in five has been involved in a patent. The lowest rate is in universe science with a 13.2% participation rate and this can be explained by the very fundamental nature of research in that field.

A gender gap in academic patenting has been evidenced in several papers (Whittington and Smith-Doerr, 2005; Ding, Murray and Stuart, 2006; Frietsch et al., 2009). Our data show that 16% of the nearly 12,000 women in our dataset (again excluding human and social sciences) are patenting, which is 64% of the rate for men. Universe science is the most gender biased field with a rate of 40%, whereas chemistry and mathematics (respective rates of 70%

[11] A professor is counted as an inventor only once in a field, independently of its number of inventions or offices where her patents were applied for.

Table 3: Involvement of professors and researchers in academic patenting, by scientific discipline (1995 – 2012).

Scientific field	professors-inventors		All professors
Chemistry	2,364	(33.3%)	7,093
Applied Bio. Ecology	1,959	(23.1%)	8,469
Fundamental Biology	3,047	(24.1%)	12,639
Medicine	2,938	(25.8%)	11,409
Engineering Sciences	2,693	(24.8%)	10,862
Mathematics	1,586	(21.7%)	7,295
Physics	2,092	(25.2%)	8,309
Universe Science	445	(13.2%)	3,383
Total	7,692	(22.3%)	34,439

Notes:
– The scientific field is part of the employment data collected from the French Ministry of Higher Education and Research and follows the OST classification (Observatoire des Sciences et Techniques).
– 17,347 professors and researchers in Social Sciences and Humanities are not represented in this table. 754 of them have invented at least one patent over the period (4.3%). If these SSH inventors are included in the full sample (51,786 researchers), the global share of academic inventors goes down to 16.3%.
– 417 individuals have missing scientific field.

and 77%) are best at closing the apparent gender gap. This gender gap is smaller than in Whittington and Smith-Doerr (2005) who reported patenting among women scientists as representing about 40% of that for men in a random sample of 4,000 life science faculty members.

We now consider the distribution of academic invention over the population of professors and researchers and its trends over the period. There are $8,863$ academic inventors in our database, defined as those researchers who invented at least one patent over the period under study. Considering the professors and research who are active in each sub-period, we find that 3.9% of them are inventors in the 1995-1999 period, 7.9% in 2000-2006 and 10% in 2007-2012. This means that patenting is adopted increasingly widely within our population[12]. However, the most prolific inventors tend to maintain or even increase their role: the top 10% most prolific inventors invent 24, 28 and 30% of academic patents in the three periods respectively. The top 5% invent 14, 17 and 19% of academic patents in the three periods. At the same time, the 4 most prolific inventors among them represent a decreasing share over time: 1.11%, 0.56% and 0.67% respectively.[13] This means that although invention behavior tends to be

[12]This trend would result from a composition effect if the numerator (number of inventors) was increasing and the denominator (number of professors) was shrinking simultaneously. This is not the case here, since the denominator is always increasing or stable over the period (see Figure A2 in the Appendix).

[13]This corresponds to the C4 indicator. Similarly, the HHI was 6.61 for 1995-1999, 3.29 for 2000-2006 and 2.89 for 2007-2012.

spreading in academia, there are more and more prolific inventors and their role does not seem to be decreasing, but might even be slightly increasing.

5 The trend in propensity to invent in academia

In this section, we aim to appreciate how the probability of inventing varies among professors and researchers over the period. A simple representation of the number of academic patents invented over time may be misleading, as in fact the underlying population of professors and researchers that we consider is likely to be increasing over the period.[14]

5.1 The spread within academia

We create an unbalanced panel dataset using the repeated time observations presented above. When someone was observed several times, any variation in the data (a promotion for instance) was assumed to occur right in the middle between the two observations. The first entry date naturally determined entry in the panel. When a lab was surveyed several times and a staff member was not listed anymore there and not listed anywhere else, we assumed that they had exited two years after the last observation. Otherwise we assumed the individual was active until the last year considered.

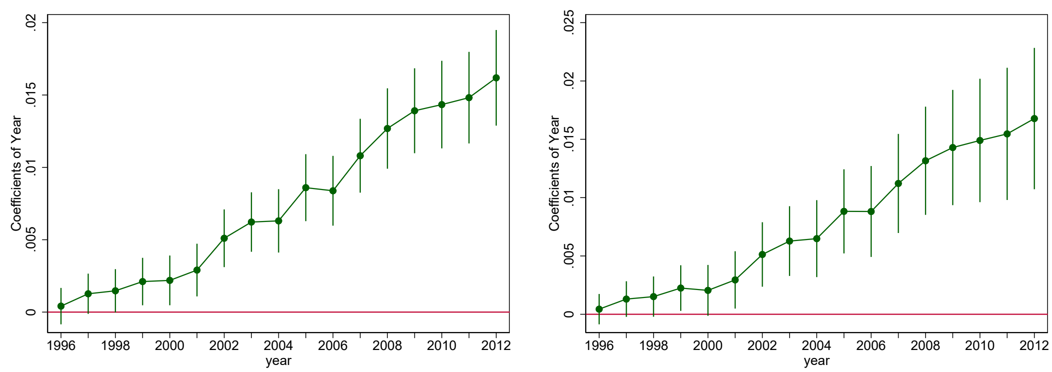
To control simultaneously for all time-invariant confounding factors (such as individual abilities or characteristics), we ran fixed effects regressions on the yearly number of inventions. The model is of the form:

$$y_{it} = \sum_{t=1995}^{t=2012} \alpha_t \, \mathrm{Year}_t + \sum_{j=1}^{j=35} \delta_j IPCj_{it} + \theta_i + \varepsilon_{it}, \tag{6}$$

where y_{it} is the outcome variable (number of EPO patent applications), Year_t is a year dummy, and θ_i is the individual fixed effect. $IPCj_{it}$ is a set of 35 dummy variables (one per technological field) controlling for the different dynamics in some fields between academia and the private sector. It equals to one if individual i invented at least one patent in the class j in year t. We are interested in estimating the coefficients of the year fixed effects α_t

[14] Figure A2 in the Appendix shows that the number of professors and researchers is increasing over time, meaning that the variation in the propensity to invent is difficult to appreciate just by computing yearly per capita ratios.

Figure 1: How the propensity to invent of academic staff varies over the period 1995-2012.



Note: The graphs present estimated coefficients of the year dummy mentioned in the horizontal axis –the α_t in Equations (6) and (7). In the left graph, the individual scientists fixed effects (the θ_i) are included whereas they are not included in the regressions leading to the right panel. Standard errors are clustered at the individual level.

for each year t. A positive trend in the estimated α_t would indicate an increasing propensity to patent over the years.

Some other factors may influence patenting but do not offer sufficient variation to be properly accounted for in a fixed effects framework. We thus estimate the following equation

$$y_{it} = \sum_{t=1995}^{t=2012} \alpha_t \, \text{Year}_t + \phi X_{it} + \sum_{j=1}^{j=35} \delta_j IPCj_{it} + \varepsilon_{it}, \tag{7}$$

where X_{it} stands for a vector of control variables, such as age and age squared, and a number of dummies for professional status, gender, field of science, university, cohort.

Figure 1 displays the estimated α_t coefficients in Equations (6) and (7) obtained via OLS regressions, allowing for many fixed effects and the clustering of standard errors. The left panel was obtained when we did not use individual fixed effects (Equation 6), whereas the right one includes them (Equation 7). The two regressions provide very similar results concerning the year fixed effects that we aim to estimate. We see that the coefficients of year fixed effects rise significantly over the period. In 2012, academic professors and researchers invented an average of 0.015 patents more than in 1995. As the average number of patents per capita in 1995 was 0.02, this means they actually increased their propensity of 75% over the period.

5.2 Comparison with non-academic inventors

The propensity to invent among academic scientists may be affected by yearly shocks affecting the economy. It could also be affected by changes in the productivity of all inventors (not just academic ones) increasing under the effects of improvements in communication technology or instrumentation, for instance. We thus need a reference point outside academia to compare the variation of the propensity to invent of academic profiles with respect to the variation observed for non-academic profiles.

We create a panel table of all French inventors from PATSTAT. The only individual identifier available in PATSTAT is "PSNID" (Magerman et al., 2009). This identifier is far from perfect but its flaws are not likely to alter the results qualitatively. The initial merge of inventor names with academic profiles presented above (see Subsection 3.3) is used to identify potential "academic" PSNIDs. A PSNID inventor profile is academic if at least one of its patents has been validated as academic and thus attached to an academic profile in our dataset of professors and researchers. Otherwise it is not academic. This clearly shows that the academic character of PSNIDs is dependent on the parametrization of the filtering stage (the choice of the β). All inventors are assumed to be in the dataset from year 1995 to 2012 so that we have a balanced panel dataset. Making the alternative assumption according to which inventors enter the dataset in the year of their first patent and leave it at the year after their last invention does not change the results qualitatively either.

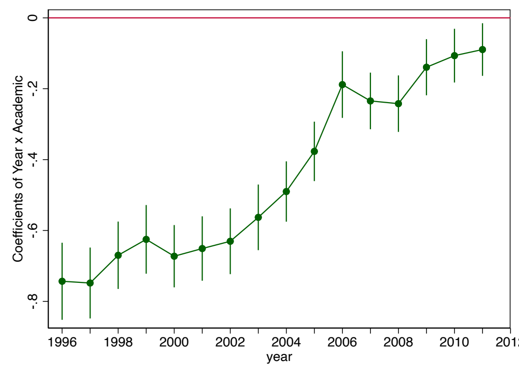
To control simultaneously for all time-invariant confounding factors (such as individual abilities or characteristics), we run fixed effects regressions on the yearly number of inventions. The model is of the form:

$$y_{it} = \sum_{t=1995}^{t=2012} \alpha_t \text{Year}_t + \sum_{t=1995}^{t=2012} \gamma_t \,\text{Academic}_i \times \text{Year}_t + \sum_{j=1}^{j=35} \delta_j IPCj_{it} + \theta_i + \varepsilon_{it}, \tag{8}$$

where y_{it} is the outcome variable (number of EPO patent applications), Year_t is a year dummy, and Academic_i is a dummy equal to one if the inventor profile is academic.

The Academic_i variable is not introduced directly into the regression as its effect is fully captured via the individual fixed effect θ_i. $IPCj_{it}$ is a set of 35 dummy variables (one per technological field) controlling for the different dynamics in some fields between academia and the private sector. It equals to one if individual i invented at least one patent in the class j in year t. The main goal of this model is to estimate the coefficient of the interaction term between the academic profile dummy and the time fixed effect, γ_t, for each year t. A positive trend observed on the γ_t would indicate that academic inventors increased their propensity

Figure 2: How the propensity to invent of academic inventors is affected relative to non-academic inventors over the period 1995-2012.



Note: The graph presents estimated coefficients and confidence intervals of the interaction term between the year dummy mentioned in the horizontal axis and the academic profile dummy (the γ_t in Equation (8)). Standard errors are clustered at the individual level.

to patent over the period at a higher rate than non-academic ones (or decreased at a lower rate).

Figure 2 displays the estimated γ_t coefficients in Equation (8) obtained via OLS, allowing us to control for many fixed effects. Profiles are declared here as academic using parameter $\beta = 2$ in the filtering stage, but again, using any of the other two values of β (1 and 0.5) leads to similar results. We can see in the figure that the estimated coefficients of the interaction terms between the year dummies and the academic profile dummy increase over the years. All coefficients are negative but tend to zero at the end of the period, suggesting that academic inventors were progressively closing the gap with non-academic inventors.

6 Factors in the spread of academic patenting

We have seen that patenting behavior increases over the period in the academic community. We now aim to unveil the drivers of that spread at the micro level. We subsequently consider two series of factors: individual attributes on the one hand, and social and cultural influence on the other hand.

6.1 Individual factors

The spread of patenting is first considered according to individual characteristics. Our main interest at this stage is to disentangle age from cohort factors in patenting, but we also consider other dimensions, such as professional status or gender. We estimate the following model:

$$y_{it} = \alpha_1 \text{Age}_{it} + \alpha_2 \text{Age}_{it}^2 + \alpha_3 \text{Cohort}_i + \alpha_4 \text{Status}_{it} + \alpha_5 \text{Gender}_i + \gamma X_{it} + \eta_t + \varepsilon_{it}, \qquad (9)$$

where η_t is the year fixed effect and y_{it} the number of EPO patent applications. Individual fixed effects are not introduced, so that time-invariant factors of academic patenting can now be considered. In particular age and cohort effects can now be estimated. The four cohorts are defined as follows: cohort 1 (the reference) groups professors born before 1950, cohort 2 groups the ones born in the 50's, cohort 3 in the 60's, and cohort 4 in the 70's or later. There are four professional statuses: associate professor, (full) professor, associate researcher, or (full) researcher. Vector X_{it} accounts for a number of controls such as scientific field and university dummies. The latter account for a number of other dimensions affecting patenting behavior, which may be correlated with the explanatory variables of interest. This model is close to the one presented in Equation (7), with some differences, but focuses on different explanatory variables.

Table 4 summarizes the regression results, again using linear regressions with many levels of fixed effects. Controlling for the cohorts, we find that age plays positively on patenting. This confirms previous research evidencing a life-cycle effect in patenting (Carayol, 2007; Stephan et al., 2007; Azoulay, Ding and Stuart, 2007; Thursby and Thursby, 2007).

When age is not included among regressors (column 1), the second, third and fourth cohort dummies are positively correlated with the outcome variable. However, when age is controlled for, cohort dummies are not significantly correlated with patenting anymore. This contrasts with Thursby and Thursby (2007) who find that more recent cohorts are in fact less likely to disclose inventions, controlling for tenure and age, or Azoulay, Ding and Stuart (2007) who evidenced the opposite. Professional status makes significant differences. Full professors invent almost twice as many patents per year as associate professors (the reference), junior researchers 67% more, and full researchers invent more than three times more often. Lastly, gender is also an important driver of patenting as women invented 50% fewer patents for equivalent years, ages, disciplines, cohorts, universities and statuses.

Table 4: The individual factors of academic patenting.

	(1)	(2)	(3)	(4)	(5)
Age		0.003***	0.003***	0.002***	0.002***
		(0.000)	(0.000)	(0.000)	(0.000)
Age squared		-0.000***	-0.000***	-0.000***	-0.000***
		(0.000)	(0.000)	(0.000)	(0.000)
Cohort 2	0.006***		0.002*		0.002
	(0.001)		(0.001)		(0.001)
Cohort 3	0.007***		0.003		0.001
	(0.001)		(0.002)		(0.002)
Cohort 4	0.004***		0.003		0.003
	(0.001)		(0.003)		(0.003)
Professor				0.011***	0.011***
				(0.001)	(0.001)
Associate Researcher				0.008***	0.008***
				(0.001)	(0.001)
Researcher				0.028***	0.028***
				(0.002)	(0.002)
Female	-0.014***	-0.014***	-0.014***	-0.013***	-0.013***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	827031	822032	822032	820068	820068
Adjusted R^2	0.011	0.012	0.012	0.013	0.013
Mean dep variable	0.021	0.021	0.021	0.021	0.021
F-statistics	169.756	268.484	136.962	164.378	113.066

Notes:
- Cohort 1 groups professors born before 1950, cohort 2 groups the ones born in the 50's, cohort 3 in the 60's, and cohort 4 in the 70's or later.
- The reference is cohort 1 for cohorts, Associate Professor for status, and male for gender.
- Standard errors into parentheses are clustered at the individual level.
- Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

6.2 Local diffusion effects

None of the individual factors examined above can fully account for the growing patenting behavior in the French academic community. The fact that the population under study is aging over the years plays in this direction, but age is already controlled for in Equation (7) and thus cannot explain the phenomenon. We thus now investigate the role of local culture within the university site and the lab.

We create two variables to capture the influence of these two layers of local social influence: variable UnivExp is the average number of patents per capita in the research community (the university site) in the previous three years ($[t-3; t-1]$), excluding all members of the focal person's lab. It proxies for the university culture towards academic entrepreneurship and patenting. The second variable LabExp is the same per capita average but considering the members of the lab only, excluding the focal person. Note that, as previous years are used to calculate some explanatory variables, observations from the first three years (1995-1997) are not considered.

We rely upon fixed effects regressions of the form

$$y_{it} = \alpha \text{UnivExp}_{it} + \beta \text{LabExp}_{it} + \phi X_{it} + \eta_t + \theta_i + \varepsilon_{it}, \tag{10}$$

where η_t is the year fixed effect. The error terms may be correlated for a given professor, a given research lab and a university, so we cluster standard errors with respect to these three identifiers (multi-way). We include time-varying controls via X_{it} such as the number of professors and researchers in the lab (LabSize) and in the university excluding those from the focal person's lab (UnivSize).

The fixed effect framework allows us to capture variation for a given professor or researcher. Therefore the estimated α and β are likely close to capturing peer effects at the university and laboratory levels respectively. However, professors are not assembled in labs and universities randomly, so correlated effects may still explain both past patenting variation of peers and contemporary variation in the patenting of the focal professor.

A number of other controls could not be introduced as they do not offer sufficient variation. We thus perform similar regressions as in Equation 10 but without individual fixed affects and introducing instead a number of supplementary control variables :

$$y_{it} = \alpha \text{UnivExp}_{it} + \beta \text{LabExp}_{it} + \phi X_{it} + \eta_t + \varepsilon_{it}, \tag{11}$$

where X_{it} stands for a vector of control variables, that includes LabSize and UnivSize as in Equation 10, but also the age, age squared, and a number of dummies for professional status, gender, field of science, and cohort.

Table 5 presents the fixed effect regression results (Equation 10). We find that although both variables are positively related to academic patenting, only LabExp coefficients are significant. When lab peers each produced one more patent per year on average in the previous years, the average faculty member invents almost 4 times more patents.[15] This supports the idea that academic patenting behavior is likely to increase in laboratories where

[15] As professors and researchers invented 0.023 patents per year on average, a coefficient of 0.065 means that productivity increases by a factor of 3.8.

such behavior has been pervasive recently. It underlines that academic patenting likely spreads locally, potentially thanks to local "peer effects".

Table A7 in the Appendix exposes the regressions results of the specification introduced in Equation 11. Results are very similar to the fixed effects for LabExp and also very similar to those presented in Table 4 regarding individual factors. The only significant difference is that the university recent experience UnivExp is now positive and significant.

Table 5: The social factors of academic patenting.

	(1)	(2)	(3)	(4)	(5)	(6)
UnivExp	0.037		0.034	0.028		0.025
	(0.050)		(0.048)	(0.048)		(0.046)
LabExp		0.068***	0.067***		0.066***	0.065***
		(0.020)	(0.020)		(0.020)	(0.020)
UnivSize				0.000	0.000	0.000
				(0.000)	(0.000)	(0.000)
LabSize				0.000	0.000	0.000
				(0.000)	(0.000)	(0.000)
Observations	658387	661617	657634	658387	661617	657634
Adjusted R^2	0.131	0.131	0.131	0.131	0.131	0.131
Mean dep variable	0.023	0.023	0.023	0.023	0.023	0.023
F statistic	34.639	34.064	22.776	17.837	17.474	14.086

Notes:
- All regressions control for individual fixed effects.
- Standard errors into parentheses are clustered at the individual, laboratory and university levels.
- Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

7 Conclusion

Since a few decades, academic patenting has been a growing phenomenon in most advanced economies worldwide, and the subject of a longstanding stream of investigation. Policy makers as well as the academic community have raised various concerns about this new practice in academia. National governments voted several reforms in order to shape universities' and professors' involvement in the phenomenon, but they often lacked reliable and consistent scientific information. Most prior literature relied on interviews, small samples of a at best few hundred professors, from one scientific field or a single university and generally covered a short period of time.

In this article, we attempt to fill this gap by investigating academic patenting including all scientific fields and several thousands of professors and researchers. We develop a methodology to appreciate the importance of, and trends in, academic patenting in France over nearly two decades. This methodology improves on existing ones as it avoids time-consuming human checking and proves reliable when trained on a benchmark set of only a few thousand professor-inventor pairs. The method is thus tractable to document patenting behavior in large datasets of academic staff and over sufficiently long periods.

We estimate that, among the 428,000 patents filed at the EPO and INPI and invented in France over the years 1995-2012, more than 44,000 stemmed from academia. The involvement of professors in technology transfer is found to be larger than expected, with one professor or researcher out of five having invented at least one patent, and widespread across most fields of the hard and life sciences (social sciences and humanities being excluded).

Even if academic patenting is strong before reforms favoring technology transfer were passed, professors and researchers are increasingly likely to invent after such reforms. An aging population of professors or simple cohort effects can not explain this phenomenon. Our results rather indicate that local diffusion, in particular within laboratories, plays a key role in the diffusion of academic patenting. A given professor is up to four times more likely to invent when her colleagues in the laboratory are more involved in patenting in the recent years (controlling for individual fixed effects). Though we cannot identify the extent of their exact influence, this suggests that peer effects may actually play a central role in fostering technology transfer.

References

Arrow, K. 1962. Economic welfare and the allocation of resources for invention. In *The rate and direction of inventive activity: Economic and social factors*. Princeton University Press pp. 609–626.

Azagra-Caro, J., P. Llerena and N. Carayol. 2006. "Patent production at a European research university: Evidence at the laboratory level." *Journal of Technology Transfer* 31:257–268.

Azoulay, P., W. Ding and T. Stuart. 2007. "The determinants of faculty patenting behavior: Demographics or opportunities?" *Journal of economic behavior & organization* 63(4):599–623.

Balconi, M., S. Breschi and F. Lissoni. 2004. "Networks of inventors and the role of academia: an exploration of Italian patent data." *Research Policy* 33(1):127–145.

Bercovitz, J. and M. Feldman. 2008. "Academic entrepreneurs: Organizational change at the individual level." *Organization science* 19(1):69–89.

Bozeman, B. and M. Gaughan. 2007. "Impacts of grants and contracts on academic researchers' interactions with industry." *Research policy* 36(5):694–707.

Carayol, N. 2007. "Academic Incentives, Research Organization and Patenting at a Large French University." *Economics of Innovation and New Technology* 16:119–138.

Carayol, N., L. Berge, L. Cassi and P. Roux. 2019. "Unintended triadic closure in social networks: The strategic formation of research collaborations between French inventors." *Journal of Economic Behavior and Organization* 163:218–238.

Carayol, N. and M. Lanoë. 2017. "The impact of project funding in science: Lessons from the ANR experience." *mimeo GRETHA working Paper* .

Carayol, N. and M. Matt. 2004. "Does research organization influence academic production?: Laboratory level evidence from a large European university." *Research Policy* 33(8):1081–1102.

Carayol, N. and V. Sterzi. 2021. "The transfer and value of academic inventions when the TTO is one option." *Journal of Economics & Management Strategy* forthcoming.

Czarnitzki, D., T. Doherr, K. Hussinger, P. Schliessler and A. A. Toole. 2016. "Knowledge creates markets: The influence of entrepreneurial support and patent rights on academic entrepreneurship." *European Economic Review* 86:131–146.

Czarnitzki, D., W. Glänzel and K. Hussinger. 2007. "Patent and publication activities of German professors: An empirical assessment of their co-activity." *Research Evaluation* 16(4):311–319.

Dasgupta, P. and P. David. 1994. "Toward a new economics of science." *Research policy* 23(5):487–521.

Ding, W., F. Murray and T. Stuart. 2006. "Gender differences in patenting in the academic life sciences." *Science* 313(5787):665–667.

Ejermo, O. and H. Toivanen. 2018. "University invention and the abolishment of the professor's privilege in Finland." *Research Policy* 47(4):814–825.

Frietsch, R., I. Haller, M. Funken-Vrohlings and H. Grupp. 2009. "Gender-specific patterns in patenting and publishing." *Research Policy* 38(4):590 – 599.

Geuna, A. and L. Nesta. 2006. "University patenting and its effects on academic research: The emerging European evidence." *Research Policy* 35(6):790–807.

Grimaldi, R., M. Kenney, D. Siegel and M. Wright. 2011. "30 years after Bayh–Dole: Reassessing academic entrepreneurship." *Research Policy* 40(8):1045–1057.

Henderson, R., A. Jaffe and M. Trajtenberg. 1998. "Universities as a source of commercial technology: a detailed analysis of university patenting, 1965–1988." *Review of Economics and statistics* 80(1):119–127.

Hvide, H. and B. Jones. 2018. "University Innovation and the Professor's Privilege." *American Economic Review* 108(7):1860–98.

Iversen, E., M. Gulbrandsen and A. Klitkou. 2007. "A baseline for the impact of academic patenting legislation in Norway." *Scientometrics* 70:393–414.

Jones, C. 1995. "R&D-based models of economic growth." *Journal of political Economy* 103(4):759–784.

Krabel, S. and A. Schacht. 2014. "Follow the leader? How leadership behavior influences scientists' commercialization behavior (or not)." *Economics of Innovation and New Technology* 23(2):134–160.

Lissoni, F. 2012. "Academic patenting in Europe: An overview of recent research and new perspectives." *World Patent Information* 34(3):197–205.

Lissoni, F., P. Llerena, M. McKelvey and B. Sanditov. 2008. "Academic patenting in Europe: new evidence from the KEINS database." *Research Evaluation* 17(2):87–102.

Louis, K. S., D. Blumenthal, M. E. Gluck and M. A. Stoto. 1989. "Entrepreneurs in academe: An exploration of behaviors among life scientists." *Administrative Science Quarterly* pp. 110–131.

Magerman, T., J. Callaert, K. Debackere and B. Van Looy. 2017. "Science informing technology: A concordance scheme between scientific disciplines and technology domains (based on scientific non-patent references)." *Scientometrics* .

Magerman, T., J. Grouwels, X. Song and B. Van Looy. 2009. Data production methods for harmonized patent indicators: Patentee name harmonization. Technical report EUROSTAT Working Paper and Studies, Luxembourg.

Meyer, M. 2003. "Academic patents as an indicator of useful research? A new approach to measure academic inventiveness." *Research Evaluation* 12(1):17–27.

Mowery, D. and A. Ziedonis. 2002. "Academic patent quality and quantity before and after the Bayh–Dole act in the United States." *Research Policy* 31(3):399–418.

Mowery, D. and B. Sampat. 2005. "The Bayh-Dole Act of 1980 and University-Industry Technology Transfer: A Model for Other OECD Governments?" *Journal of Technology Transfer* 30:115–127.

Mowery, D., R. Nelson, B. Sampat and A. Ziedonis. 2001. "The growth of patenting and licensing by US universities: an assessment of the effects of the Bayh-Dole act of 1980." *Research policy* 30:99–119.

Nelson, R. 1959. "The simple economics of basic scientific research." *Journal of Political Economy* 67:297–306.

Perkmann, M., V. Tartari, M. McKelvey, E. Autio, A. Broström, P. D'Este, R. Fini, A. Geuna, R. Grimaldi, A. Hughes et al. 2013. "Academic engagement and commercialisation: A review of the literature on university–industry relations." *Research policy* 42(2):423–442.

Romer, P. M. 1990. "Endogenous technological change." *Journal of political Economy* 98(5, Part 2):S71–S102.

Stephan, P. 2010. "The economics of science." *Handbook of the Economics of Innovation* 1:217–273.

Stephan, P., S. Gurmu, A. Sumell and G. Black. 2007. "Who's Patenting in the University? Evidence from the Survey of Doctorate Recipients." *Economics of Innovation and New Technology* 16(2):71–99.

Stephan, P. and S. Levin. 1996. "Property rights and entrepreneurship in science." *Small Business Economics* 8(3):177–188.

Stephane, M. and C. Martinez. 2014. "Identifying author-inventors from Spain: methods and a first insight into results." *Scientometrics* 101:445–476.

Thursby, J., A. W. Fuller and M. Thursby. 2009. “US faculty patenting: Inside and outside the university.” *Research Policy* 38(1):14–25.

Thursby, J. and M. Thursby. 2007. Patterns of research and licensing activity of science and engineering faculty. In *Science and the University*, ed. Paula E. Stephan and Ronald G. Ehrenberg. Madison: University of Wisconsin Press.

Verspagen, B. 2006. “University research, intellectual property rights and European innovation systems.” *Journal of Economic surveys* 20(4):607–632.

Whittington, K. B. and L. Smith-Doerr. 2005. “Gender and commercial science: Women’s patenting in the life sciences.” *The Journal of Technology Transfer* 30(4):355–370.

Appendix

Figure A1: Precision, recall et F_β (when $\beta = 0.5$, $\beta = 1$ or $\beta = 2$) for different threshold probability values.

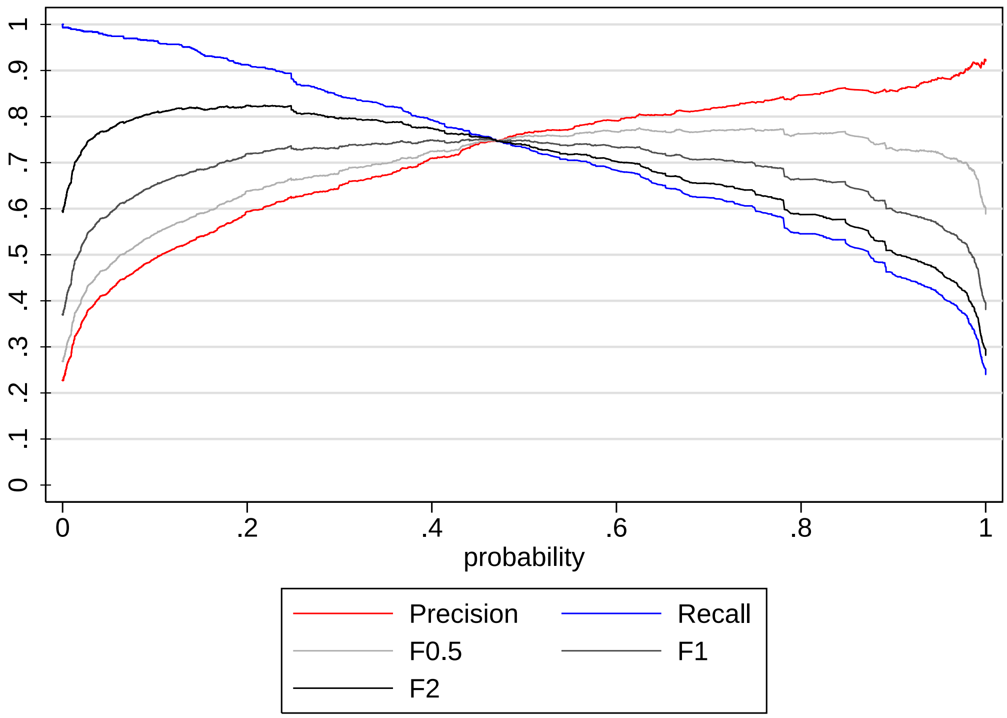
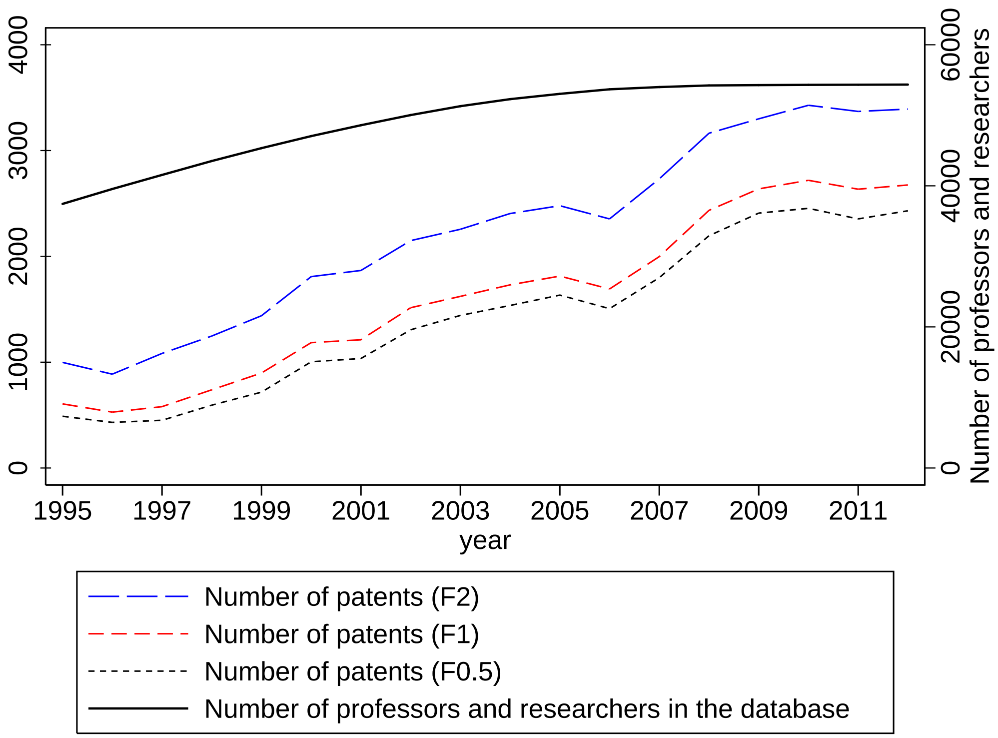


Figure A2: The evolution of academic patenting in France, with respect to the reference population.



Notes:
- This graph displays fractional counts for academic patents.
- The left scale is the number of patents and the right scale the number of professors and researchers in our database.
- The number of patents are computed for each threshold : F1 (respectively F05 and F2) relates to $\beta = 1$ (respectively $\beta = 0.5$ and $\beta = 2$).

Table A1: Logistic regressions on the benchmark, per office

	(1) EPO	(2) EPO	(3) INPI
Name similarity	22.427***	21.834***	12.740***
	[3.635]	[3.117]	[2.002]
Inventor's name frequency	1.234***	1.239***	0.645***
	[0.165]	[0.169]	[0.055]
Assignee/employer consistency	6.244***	6.231***	1.273***
	[0.675]	[0.662]	[0.455]
Tech/discipline consistency	0.566**	0.591***	0.425***
	[0.193]	[0.202]	[0.087]
Age/year consistency	0.666		1.245***
	[0.554]		[0.244]
Constant	-32.188***	-29.189***	-21.841***
	[5.046]	[3.618]	[2.392]
Observations	682	682	2829

Notes:
- Bootstrap standard errors into brackets.
- Significance levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A2: Optimal thresholds for each office and β values $(0.5, 1, 2)$

Patent office	β	Optimal threshold	Precision	Recall	F-measure	Number of validated patents
EPO	2	0.14	0.53	0.95	0.82	20,648
	1	0.44	0.84	0.81	0.82	12,166
	0.5	0.46	0.88	0.80	0.86	11,995
INPI	2	0.20	0.52	0.93	0.80	20,202
	1	0.45	0.66	0.82	0.73	12,620
	0.5	0.74	0.77	0.62	0.74	7,898

Notes:
- Interpretation: For patents filed at the EPO and the when precision is valued the most (thus $\beta = 0.5$), the maximum F-measure is 0.86 for a threshold probability set at 0.46. In the dataset, 11,995 patents have a probability higher than or equal to 0.46 and are thus validated as academic patents.

Table A3: Expected number of academic patents for several β values (from 1995 to 2012)

Office	$\hat{x}_2$		$\hat{x}_1$		$\hat{x}_{0.5}$		All French-invented patents
EPO	11072	(6.1%)	12320	(6.9%)	12325	(7%)	177286
INPI	11173	(4.5%)	10804	(4.3%)	10473	(4.1%)	250605
Total	22245	(5.1%)	23123	(5.4%)	22798	(5.3%)	427891

Notes:
– This table displays fractional counts.
– The shares of academic patents - by office and overall - over all patents invented in France are placed in parentheses.

Table A4: RTA and CAGRs for patents in 35 technology classes (1995-2012)

Technology class	RTA	CAGR	Growth ratio	Share
Micro-structural and nano-technology	5.40	-1.5%	5.5	0.2%
IT methods for management	0.41	-1.9%	3.5	0.5%
Digital communication	0.53	-0.0%	2.0	3.7%
Semiconductors	2.76	4.1%	1.8	2%
Engines, pumps, turbines	0.81	-3.2%	1.7	3.7%
Computer technology	0.86	1.5%	1.7	5.3%
Analysis of biological materials	3.40	2.6%	1.6	0.8%
Transport	0.34	-2.4%	1.5	8.7%
Measurement	1.99	-0.6%	1.1	4.5%
Thermal processes and apparatus	0.73	1.5%	1.1	1.5%
Surface technology, coating	1.64	0.8%	1.1	1.4%
Environmental technology	1.43	-1.2%	1.1	1.6%
Electrical machinery, apparatus, energy	0.96	1.4%	1.0	5.5%
Food chemistry	0.76	-3.9%	1.0	1%
Civil engineering	0.47	-4.3%	0.9	4.1%
Other consumer goods	0.32	0.4%	0.9	2.6%
Telecommunications	0.82	0.8%	0.8	2.8%
Control	0.65	0.4%	0.8	1.7%
Medical technology	0.89	2.3%	0.8	4.1%
Audio-visual technology	0.77	-1.7%	0.8	2.6%
Machine tools	0.65	-2.6%	0.8	2%
Other special machines	0.57	-1.4%	0.7	3.5%
Handling	0.38	-1.1%	0.7	3.4%
Biotechnology	3.16	-0.5%	0.6	2.7%
Mechanical elements	0.42	-2.6%	0.6	3.8%
Pharmaceuticals	1.57	1.1%	0.6	4.8%
Furniture, games	0.23	3.1%	0.6	2.6%
Materials, metallurgy	1.41	0.7%	0.6	1.9%
Basic materials chemistry	1.52	-1.0%	0.6	2.1%
Macromolecular chemistry, polymers	1.03	1.2%	0.6	1.9%
Basic communication processes	1.32	4.8%	0.6	1%
Chemical engineering	1.71	-1.0%	0.5	2.7%
Optics	1.61	1.2%	0.5	1.9%
Organic fine chemistry	0.82	1.5%	0.4	6.1%
Textile and paper machines	0.44	0.7%	0.0	1.3%
Total	1	0%	1	100%

Notes:

- For a technology class i, the revealed technological advantage is RTA= $\frac{A_i}{B_i} \times \frac{\sum_i B_i}{\sum_i A_i}$. Column one displays its average value over the years 1995-2012.
- For column two, we first calculate the RTA of each technological class every year, then we display its Compound Annual Growth Rate between 1995 and 2012.
- We calculate the CAGR of French patents overall ($nCAGR$) and the CAGR of those patents in a technology class i ($nCAGR_i$). Thus, in column three we have the growth ratio = $\frac{nCAGR_i}{nCAGR}$.
- For the fourth and last column, share = $\frac{\#patents_i}{\#patents}$ where $\#patents_i$ is the number of French patents in technological class i, and $\#patents$ the total number of French patents.
- A patent may belong to more than one technology class so we use fractional counts.

Table A5: Repartition of professors and researchers involved in academic patenting by scientific discipline (1995 – 2012), for a chosen $\beta = 0.5$ at the filtering stage.

Scientific field	professors-inventors		All professors
Chemistry	2,327	(32.8%)	7,093
Applied Bio. Ecology	1,924	(22.7%)	8,469
Fundamental Biology	2,992	(23.7%)	12,639
Medicine	2,892	(25.3%)	11,409
Engineering Sciences	2,626	(24.2%)	10,862
Mathematics	1,545	(21.2%)	7,295
Physics	2,039	(24.5%)	8,309
Universe Science	420	(12.4%)	3,383
Total	7,503	(21.8%)	34,439

Notes:
– The scientific field is part of the employment data collected from the French Ministry of Higher Education and Research and follows the OST classification (Observatoire des Sciences et Techniques).
– 17,347 professors and researchers in Social Sciences and Humanities are not represented in this table. 670 of them have invented at least one patent over the period (3.9%). If these SSH inventors are included in the full sample (51,786 researchers), the global share of academic inventors goes down to 15.8%.

Table A6: Repartition of professors and researchers involved in academic patenting by scientific discipline (1995 – 2012), for a chosen $\beta = 2$ at the filtering stage.

Scientific field	professors-inventors		All professors
Chemistry	2,500	(35.2%)	7,093
Applied Bio. Ecology	2,121	(25%)	8,469
Fundamental Biology	3,299	(26.1%)	12,639
Medicine	3,174	(27.8%)	11,409
Engineering Sciences	2,923	(26.9%)	10,862
Mathematics	1,742	(23.9%)	7,295
Physics	2,282	(27.5%)	8,309
Universe Science	538	(15.9%)	3,383
Total	8,441	(24.5%)	34,439

Notes:
– The scientific field is part of the employment data collected from the French Ministry of Higher Education and Research and follows the OST classification (Observatoire des Sciences et Techniques).
– 17,347 professors and researchers in Social Sciences and Humanities are not represented in this table. 1,086 of them have invented at least one patent over the period (6.3%). If these SSH inventors are included in the full sample (51,786 researchers), the global share of academic inventors goes down to 18.4%.

Table A7: The social and individual factors of academic patenting.

	(1)	(2)	(3)	(4)	(5)	(6)
UnivExp	0.086***		0.063***	0.084***		0.062***
	(0.019)		(0.015)	(0.020)		(0.016)
LabExp		0.151***	0.148***		0.151***	0.148***
		(0.017)	(0.017)		(0.017)	(0.017)
UnivSize				-0.000	-0.000	-0.000
				(0.000)	(0.000)	(0.000)
LabSize				0.000	0.000	0.000
				(0.000)	(0.000)	(0.000)
Age	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age squared	-0.000*	-0.000*	-0.000*	-0.000*	-0.000*	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cohort 2	0.003**	0.003**	0.003**	0.003**	0.003**	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Cohort 3	0.003	0.002	0.002	0.003	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Cohort 4	0.004	0.004	0.004	0.004	0.004	0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Professor	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Associate Researcher	0.003***	0.003***	0.002***	0.003***	0.003***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Researcher	0.011***	0.011***	0.011***	0.011***	0.011***	0.011***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	658415	661645	657662	658415	661645	657662
Adjusted R^2	0.450	0.451	0.452	0.450	0.451	0.452
Mean dep variable	0.023	0.023	0.023	0.023	0.023	0.023
F statistic	16.824	17.555	20.664	14.657	15.430	18.886

Notes: All regressions also control for university, field of science, technological specialization and cohort fixed effects. Standard errors into parentheses are clustered at the individual, laboratory and university levels. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.