Is the aggregate effect of R&D outsourcing positive?

An empirical assessment of the impact of external R&D on innovation in French regions.

Olivier Brossard and Inès Moussa,
LEREPS, Political Studies Institute, Toulouse 1 University

Abstract:
Empirical studies at the individual firm-level most always find a positive impact of R&D outsourcing on innovation when firms' absorptive capacity is high enough or when R&D is outsourced abroad. However, since R&D generates local positive spillovers, aggregate R&D outsourcing may produce a detrimental loss of local knowledge that hinders local innovation. Consequently, aggregate results of R&D outsourcing may differ from individual firm-level results. We estimate knowledge production functions in the 94 metropolitan French NUTS3 regions observed between 1997 and 2008 to test this possible fallacy of composition effect. We use in house and outsourced R&D figures from the official French R&D survey. This allows us to differentiate three categories of outsourced R&D: affiliate R&D outsourcing, non-affiliate R&D outsourced in France and non-affiliate R&D outsourced abroad. We find no positive impact of outsourced R&D, a negative impact of domestic R&D outsourcing and no evidence of a complementarity between outsourced and in house R&D.

Keywords:
Innovation
Knowledge production
R&D
R&D outsourcing
Urbanization economies

JEL codes: R12; R15

1. Introduction

Is it better for innovation performance to promote in house knowledge generation processes based on internal R&D expenses, or is it profitable to replace them, at least partially, with outsourced R&D expenditures? This question has been at the centre of quite a number of research papers in recent years (see, e.g., Hsuan and Mahnke, 2011, for a review). Theoretical contributions clearly show that there can be gains and pains from R&D outsourcing, and firm-level empirical evidence suggests an inverted U-shaped relationship between R&D outsourcing and innovation performance (Grimpe and Kaiser, 2010). However, to the best of our knowledge, there is no empirical study of the *aggregate* effects of R&D outsourcing on innovation performance, and this leaves an important innovation policy issue unexplored: given that a reasonably proportioned externalization of R&D seem to be beneficial at the individual firm level, should one consider that national or regional innovation policies need to encourage R&D outsourcing? In the realm of (knowledge) externalities, nothing is less evident than such a micro-macro transition. In general, externalities produce fallacy of composition effects (Kirman, 1992, Colander et al., 2008), and this is a good reason why localized knowledge spillovers may imply that outsourcing R&D is not necessarily good at the regional or national level.

The pros and cons of R&D outsourcing are well identified. On the one hand, it provides access to new pieces of knowledge that are not available inside the firm (Nelson and Winter, 1982). It also generates new, knowledge-based, inter-firm collaborations in a context of rising open innovation (Chesbrough and Appleyard 2007), and it offers opportunities to reduce R&D costs and make reversible R&D investments with smaller capital stakes and less risk (Narula, 2001). On the other hand, outsourcing R&D may generate a detrimental loss of strategic research competencies (Howells et al., 2008), a “not invented here” syndrome due to excessive cognitive distance (Katz and Allen, 1982, Nooteboom, 2009), an impoverishment of research outputs because only their codified part is transferable across organizations, and an impoverishment of national or regional systems of innovation because of knowledge leakages (Howells, 1996, Narula, 2001). Be that as it may, the balance of empirical results seems to be in favour of a reasonably calibrated R&D outsourcing. Indeed, as shown in the empirical literature review below, the least favourable studies of the impact of external R&D on innovation obtain that it has no influence on process innovation but a positive impact on product innovation, or an impact that exists only when internal R&D is already important enough, or a positive impact that is limited to external R&D outsourced abroad rather than in home country. Does this mean that regional innovation policies should support R&D outsourcing wherever it is efficient for the individual businesses, even if it might prove detrimental for the local flows of knowledge spillovers?

In this paper, we address this question by an empirical investigation of innovation in French regions. We estimate aggregate knowledge production functions (KPF) on a panel dataset made of the 94 French metropolitan regions observed between 1997 and 2008. The R&D inputs are extracted from the French Ministry of Research's R&D survey, which provides detailed information on firms' in house and outsourced R&D. Contrary to the CIS survey, it is possible to aggregate the R&D survey data from the individual firm-level to the NUTS3 regional level without losing representativity. At this level of aggregation, the KPF framework basically correlates R&D activities of a region with its own innovative output, and the intra-regional R&D spillovers are incorporated in the coefficients of the R&D input variables. Moreover, it is possible to use spatial econometric techniques to correlate the innovative output of a region to the R&D inputs of different and distant ones, and thus assess whether there are some R&D spillovers from neighboring regions on top of the intra-region ones. Suppose that outsourced R&D produces a positive effect for the firms that implement R&D externalization but a negative one for the regions hosting them because a generalized outsourcing movement diverts key R&D activities towards other regions. A firm-level estimation of the KPF would result in a positive coefficient of external R&D whereas the regional-level estimation could provide a negative one. In this case, the great advantage of studying the impact of R&D outsourcing on knowledge production at this regional level is that one can account for the potential knowledge leakages generated by the depletion of the aggregate regional R&D effort. However, such an approach requires to control seriously for the regions' characteristics that may affect regional innovation performance. We thus use information on regions' industrial structure to address specialization/diversification economies. We also investigate the impact of urbanization economies using demographic data, examine the effect of the regional stock of human capital using information on workforce education, and we control for public regional R&D and regions' degree of internationalization.

To the best of our knowledge, this is the first study to address the aggregate effects of R&D outsourcing. Our econometric framework uses panel econometrics that seriously account for heterogeneity and endogeneity problems, using the estimator proposed by Hausman and Taylor (1981) along with more classical within regressions. We test the joint impact of in house and outsourced R&D on NUT3 regions' patent applications, controlling carefully for the other acknowledged determinants of regional innovation, namely diversification/specialization externalities, regions internationalization and regions' human capital.

The paper is organised as follows. In section 2, we review the theoretical and empirical literature on the impact of external R&D on innovation. In section 3, we present the data, the research design and the econometric results. Section 4 concludes the paper.

2. The gains, the pains and the fallacy of composition of outsourcing R&D

If the first generation of knowledge production studies have mainly underlined the importance of the *quantity* of R&D capital as an input (Griliches 1979), subsequent researches have put more emphasis on the *nature* of the different types of knowledge inputs involved in learning and discovery processes. This focus on knowledge nature was originally inspired by the seminal work of M. Polanyi (1966) on tacit knowledge, revived by Nelson and Winter (1982) and applied by Gertler (2003) to show the importance of contextualization as a determinant of production, appropriation and exchange of tacit knowledge. Nevertheless, the distinction between tacit, person-embodied, knowledge and codified, explicit, knowledge is difficult to operationalize in econometric frameworks. As a consequence, many knowledge production studies have differentiated public and private R&D, or academic and entrepreneurial R&D, or again basic and applied R&D, but these categories became less differentiated and, thus, less relevant in the age of Mode 2 science (Gibbons et al., 1994). Nevertheless, there is another categorization of knowledge inputs that could prove fertile, both because it better accounts for the importance of tacit and codified knowledge in innovation production processes, and because it better characterizes the actual trade-off faced by innovators between cognitive proximity and knowledge newness: the division between in house and outsourced R&D. To define the latter, one could use a large definition that would encompass licensing and other kinds of technology acquisitions, as well as R&D subcontracting, alliances and formal collaborations, but this would blur the frontiers between inputs and outputs and hinder the identification of the impact of outsourced R&D *inputs* on the production of innovation *outputs*. We will thus prefer a definition of external R&D that does not include the purchase of ready-made technologies and only considers the outsourcing of knowledge *inputs*. Be that as it may, outsourcing R&D is a growing tendency in R&D

intensive industries (Mol, 2005, KPMG, 2008, Howells et al., 2008), generating new challenges for firms' innovation strategies as well as for national or regional innovation policies.

There are strong forces behind R&D outsourcing. Firstly, costs and risks optimization is certainly the main driver of the externalization of less strategic R&D activities (Narula, 2001, Mudambi and Tallman, 2010). For instance, after the discovery of a new drug, pharmaceutical firms frequently outsource the trial work to a firm specialized in clinical testing. Howells and al. (2008) reveal the emergence of such contract R&D companies belonging to the fast growing sector of "R&D services" (SIC code 73.10). Innovative pharmaceutical companies are also using the R&D services of software consultancy, data processing and knowledge management firms that provide, for example, useful genomic information services. This example taken in the pharmaceutical industry is topical of the phases of the R&D processes that can be most easily outsourced: not the upstream exploration stages that are complex and strategic, but rather the downstream examination or exploitation ones wherein the tradeoff between appropriation and accessibility is less compelling (March, 1991, Cooke, 2006, Gilsing and Nooteboom, 2006, Brossard and Vicente, 2010).

In exploratory phases, another motive for outsourcing R&D may come into play: the awareness that true knowledge breakthroughs require the combination of heterogenous pieces of knowledge provided by cognitively and geographically distant actors (Nelson and Winter, 1982, March, 1991). Outsourcing part of the exploratory R&D provides access to new talents and new knowledge inputs. A firm contracting a R&D project with a university lab specialized in nanotechnologies or robotics, for instance, might get access to high-potential ideas. However, it could also appear difficult for the firm's engineers to understand what to do with basic research results because the key knowledge that is needed to exploit these results is tacit rather than codified, and because the tacit is much harder to transfer than the codified. Several authors argue that outsourced R&D mostly supplies codified results and does not provide a lot of person-embodied knowledge (e.g., Howells, 1996, Cowan and Foray, 1997, Cantwell and Santangelo, 1999 or Narula, 2001). Serious arguments support this view: tacit knowledge transfers require frequent interactions in a trust climate that is more easily attained when people belong to the same company and share the same routines and norms. However, the importance of internal learning processes is not justified only by the tacit dimension of knowledge: in a case study of the Brescia mechanical cluster, Lissoni (2001) has clearly shown that codified knowledge is also better exploited inside firms' boundaries because the understanding of the codes requires firm-specific skills. In fact, what is really important for knowledge production and diffusion is the existence of common cognitive routines. This highlights the other reason why transferring a truly new knowledge is not an easy task: cognitive distance. As Nooteboom (2009) clearly demonstrated, too much cognitive distance between the members of an organization can generate misunderstanding in many processes wherein knowledge exchange is required. Codification does not necessarily abolish cognitive distance and, therefore, does not always make the transfer of new knowledge easy. Similarly, R&D contracts, rent-sharing and co-patenting agreements provide legal solutions for the protection of intellectual property, but they do not offer any solution to misunderstanding problems. To deal with this challenge, the key capability is firms' absorptive capacity in the sense of Cohen and Levinthal (1990), that is to say a capability to bridge the cognitive distance between themselves and their knowledge suppliers. Internal R&D is considered the main source of absorptive capacity because developing one's knowledge base is a way to increase cognitive capacity and therefore gain ability to identify, interpret and exploit new knowledge (Cohen and Levinthal, 1990, Griffith et al., 2004, Cassiman and Veugelers, 2006, Bertrand and Mol, 2013). This suggests a complementarity rather than a substitutability relationship between in house and outsourced R&D.

If we regard tacitness and cognitive distance as the two main reasons why R&D outsourcing may prove inefficient in some cases, there are also other downsides of R&D externalization. Howells et al. (2008) underline the risk of detrimental loss of strategic research competencies; Katz and Allen (1982), the “not invented here” syndrome; and Howells (1996), Narula (2001), and Hsuan and Mahnke (2011) point the impoverishment of national or regional systems of innovation, the brain-drain of researchers and the loss of innovation-based first-mover advantage. It is not surprising, hence, that R&D back-sourcing has been decided by a significant number of firms (Mahnke, 2007, Zirpoli and Becker, 2014).

In summary, the pains of R&D outsourcing may beat the gains when absorptive capacity is not strong enough. As a consequence, there should be at the firm-level a complementarity relationship between outsourced R&D and the main determinants of absorptive capacity (Internal R&D, experience of previous research collaborations, quality and frequency of the interactions with external knowledge providers, efficiency of internal knowledge management processes). Whether it is necessarily true at the aggregate level is another issue that we now want to discuss because it seems to be neglected in empirical works.

More and more empirical studies at the firm level provide evidence of the complementarity between internal and external knowledge inputs (see, e.g., Cassiman and Veugelers 2006, Tsai and Wang 2008, Hagedoorn and Wang 2012). For instance, Grimpe and Kaiser (2010) study a sample of 3966 innovative firms in Germany and find evidence of an inverse U-shaped relation between R&D outsourcing and innovation performance. They show that the tipping point of this relation depends on firms’ absorptive capacity since it is at higher levels of outsourced R&D when internal R&D is larger and when firms have more formal innovation-related collaborations. Nevertheless, a few studies find no significant complementarity between internal and external R&D (Vega-Jurado *et al*. 2009, Hess and Rothaermel 2011). Most interestingly, in a study based on the official R&D Survey of the French Ministry of Research, Bertrand and Mol (2013) obtain that the impact of external R&D is strongly positive on product innovation when R&D is outsourced abroad, but they also find negative impacts of affiliate R&D outsourcing and domestic R&D outsourcing. They interpret these results as evidence that outsourcing R&D is positive only when the cognitive distance with the source is high enough. However, results seem to vary across countries since Arvanitis and Loukis (2012) find positive effects of external R&D both for product and process innovation in Switzerland, but no significant effect on product and process innovation in Greece.

To the best of our knowledge, all the empirical tests on this issue use firm-level datasets. An interesting study by Rondé and Hussler (2005) provides evidence of a positive influence of "external competencies" on regional innovation using French data aggregated at the NUTS3 level. However, since they do not exploit R&D data, they do not include information on in house or outsourced R&D but only aggregated measures of some firms' competences that they consider favorable to external knowledge exploitation[1]. As a consequence, there is a complete dearth of empirical evidence that individual firm-level results on the impact of R&D outsourcing could be generalized at an aggregate, regional or national, level. Thus, an important innovation policy question remains unanswered: given that outsourced R&D seem to improve individual innovation performance under some conditions that are rather easy to realize, should regional innovation policies encourage R&D outsourcing? Knowledge externalities may perturb this micro-macro generalisation because of

[1] They use a survey implemented in 1997 by the SESSI (a research department of the French Ministry of Industry) providing information on the set of competences that French firms possess. They classify these competences to differentiate internal and external ones in relation to the innovation process. Their results show the importance of external competences in regional innovation processes, which we consider as valuable empirical evidence of the importance of absorptive capacity. Nevertheless, the complementarity between these factors contributing to build a good absorptive capacity and external R&D is not tested.

a fallacy of composition effect (Kirman, 1992, Colander et al., 2008). Indeed, it is widely acknowledged that R&D produces localized knowledge spillovers, that is to say, positive externalities on innovation in the neighbourhood wherein R&D is implemented (Jaffe, 1986, Griliches, 1991, Audretsch, 1998, Bottazzi and Peri, 2003, Feldman and Kelley, 2006). Massive R&D externalization could lead some places to deprive themselves from these positive spillovers. Firms may have self-interest in outsourcing R&D, but they may have no interest in seeing their neighbours doing the same thing. If this fallacy of composition exists, individual firm-level studies will not detect the negative externality of massive R&D outsourcing. That is why we implement an empirical investigation of the impact of external R&D on innovation in French regions rather than in French firms.

3. Empirical assessment

The knowledge production function approach introduced by Griliches (1979) and Jaffe (1986) is highly appreciated as a means of detecting and quantifying knowledge flows and knowledge externalities. We select this approach because it can provide direct measures of the impact of various forms of R&D on regional innovation.

3.1. Sample and variables construction

We estimate our model on the so-called French “départements” between 1997 and 2008. These administrative units created in 1789 correspond to NUTS 3 regions in the Eurostat classification. We exclude overseas «départements», as well as Southern and Northern Corsica, to circumvent discontinuity problems. Consequently, we work with 94 metropolitan «départements», observed during twelve years regarding patents and fourteen years concerning R&D. Working on geographical units rather than on individual firms is useful if one desires to detect the effects of the possible decline in regional knowledge spillovers due to R&D outsourcing. The NUT3 aggregation level is relevant therein because many innovation, labour market or educational policies are implemented at this level and contribute to create measurable differences between NUTS3 regions[2]. The descriptive statistics of the variables are presented in **Table 1**.

3.1.1. The dependent variable: innovation output

Despite its imperfections, the patent count indicator is a widely accepted proxy for the innovative output. The caveats are well known: some valuable innovations are not patented, and many patents will prove to have low economic value. In addition, the design of the patent system, the type of R&D implemented (e.g., business versus basic R&D), and the variety of science and technology policies may all influence the patenting strategies through a propensity to patent effect (de Rassenfosse and van Pottelsberghe de la Potterie, 2009). Nevertheless, there are means to control for this effect and the novelty content of patented innovations is warranted by the patenting procedure, whereas it is much more problematic to assess the newness of the product or process innovations added up in innovation surveys (Griliches, 1990).

The French National Institute of Industrial Property (INPI) provided us with a count of published patent applications of French origin between 1997 and 2008. Patents have been distributed across the French NUTS 3 regions ("départements") according to the address of the inventor. In case of multiple co-inventors residing in different NUTS3 regions, an even fraction of the patent is granted to each region. If one of the co-inventors does not live in France, the corresponding fraction of the patent is not counted. These counts include all patent applications of French origin published by any possible patent office, that is to say, the national one (INPI), the European one (EPO), the American one (USPTO) and so on. They also include all applications filled under the Patent Cooperation Treatise (PCT). To avoid multiple counting, only first fillings are considered. All industries are covered, including, for

[2] For a convincing argumentation in favor of studying innovation processes at the level of a geographical unit, one can read Rondé and Hussler, 2005, for instance.

instance, the patenting of financial innovations. Counting evenly all the possible sorts of patents has caveats and advantages. One main problem is that it amounts to considering that all patents have the same innovative content. However, the inclusion of all patent categories provides a more comprehensive account of the innovativeness of each region[3]. This inclusive approach is also interesting because it softens the propensity to patent problem: some unobserved regional characteristics may influence the propensity to patent in general and also the propensity to file patents at one particular office rather than the others. The latter problem is eliminated by the inclusion of all types of patents in the count. We can also mention that, contrary to many studies, time-smoothing of the patent count proved unnecessary because this inclusive approach of counting patents results in the absence of zeros in regional patent counts[4]. This allowed us to maintain a panel data structure.

3.1.2. R&D independent variables

The main independent variables are internal and external R&D expenses of region *i* over the year *t*. We extracted these figures for the period 1995-2008 from the national R&D survey implemented yearly by the French Ministry of Research. The French R&D survey is implemented since 1967, but internal R&D expenses and R&D workforce are localized at the NUTS3 level only since 1993. For the private sector dataset that we exploit here, 11000 firms are surveyed[5]. The sampling method warrants size and sector representativity both at the

---

[3] Because the INPI only provided us with the patent counts and not the detailed data, we could not implement the solutions that are sometimes used in the literature. Nevertheless, we can argue that these solutions are not necessarily satisfying: weighting the patents by their number of citations has disadvantages since many citations are imposed by patent examiners according to criteria that do not really reflect economic value. Moreover, very new ideas may not be cited for quite a while. In addition, the other solution that would consists in running the regressions with only the patents of one single office would imply a large measurement error in the dependent variable; a solution we consider worse than the disease.

[4] This is also due to the NUTS3 aggregation level.

[5] More precisely all the 243 French firms classified as “large” are surveyed each year, representing 87% of the French internal R&D expenses, and the rest of the firms are partially surveyed each year. The parent population is made of all the firms that implement R&D, which represents approximately 23000 companies on a total of 3,14 million businesses in France (of which 3 million are classified as “micro-firms”, 138000 are “SMEs”, 5000 are “intermediary size firms” and 243 are “large enterprises”). The drawing is exhaustive for two categories of

national and at the NUT3 level. It is a compulsory survey and there is an adjustment of the results to correct non-responses. The content of the survey is detailed. Firms give their principal activity code, their research activity codes if they have several, various information on their size and structure and, then, figures on their R&D effort. Respondents are asked to report R&D according to the definition of the Frascati Manuel. As a consequence, for example, they are explicitly requested not to report patent or license purchase as external R&D expense. Eligible R&D expenses are wages and taxes, other current expenses, lands, buildings, machinery and equipment, software, capital expenses. Regarding internal R&D, respondent firms are asked to give its distribution across six technological domains (software development, biotechnologies, etc.), its distribution across the firms establishments in NUTS3 regions, its allocation by nature of expense (wages, general expenses, building and real estate, equipment and so on), and its division across fundamental research, applied research and experimental research. The firms are also asked to distribute their total R&D workforce in the NUTS3 regions and precise whether these are researchers and engineers, technicians and so on. Finally, there are also questions on external R&D expenses and external resources obtained for R&D (public or private subsidies). External R&D expenses are not localized at the NUTS3 level unfortunately, but there is very interesting information regarding whether this research has been outsourced to public sector organizations (Universities, national research labs, etc.), whether it has been outsourced to foreign firms, and whether it has been externalized to affiliate or non-affiliate firms. This will allow us to differentiate forms of R&D outsourcing in a way that provides proxies of the geographical and cognitive distances characterizing the external knowledge purchased by French firms.

---

firms (6000 entities): those that make more than 750K€ of internal R&D expense and those that recently entered the parent population because they started doing R&D. 5000 entities are drawn among the 17000 remaining ones.

Previous generations of this R&D survey have already been exploited by Autant-Bernard (2001), Autant-Bernard et al. (2011) and Mairesse & Mulkey (2008), but they only used internal R&D and R&D staff figures. Bertrand and Mol (2013) have recently exploited the external R&D figures from the 1995-2004 surveys, but at the individual firm-level. We do not know of any study that would have aggregated these figures at a regional level. Given the regional policy perspective of this paper, the French R&D Survey has two interesting advantages over the community innovation survey (CIS). Firstly, R&D figures are collected both at the firm level and at the establishment level. The latter statistics are necessary if one seeks to trace precisely the locus of R&D activities. Secondly, the R&D survey data are representative of firms' sizes and sectors both at the national and at the regional level, that is to say, in the territorial units we study (the French «départements»).

Because we want to account for all the R&D implemented in each of the 94 French regions, we need to recount all the R&D expenses of all the business units present in each region. This is straightforward for in-house R&D, since the figures are available at the establishment level, but external R&D is only available at the firm level and has to be re-allocated to firms' business units. An important point needs to be mentioned here: contrary to internal R&D, external R&D is *implemented* somewhere and *exploited* somewhere else. If we were mostly seeking to detect the knowledge spillovers generated by external R&D in the neighborhood around which it is implemented, it would be necessary to know the place where it is realized. This information is not available in the R&D survey, but we know whether R&D has been outsourced domestically or abroad, and whether it has been outsourced to affiliate or non-affiliate organizations, which will prove useful to differentiate external R&D that is being implemented at larger cognitive and geographical distances. Since we are seeking to measure the impact of outsourced R&D on the regions wherein it is ordered and exploited, not wherein it is implemented, it is only necessary to find a methodology to redistribute the total external R&D expense of multi-establishment firms across their regional business units. A simple way to allocate the outsourced R&D of multi-establishment firms across their business units would be to distribute the company external R&D expenses in proportion of the share of the company's total internal R&D expense or total R&D workforce that each establishment receives. In favor of this method, one could argue that the R&D outsourced by a company is certainly exploited in priority in the places wherein an internal research capacity still exists and allows interpreting and exploiting the external research results. However, this could artificially reinforce the correlation between internal and external R&D expenses of NUTS3 regions because we would allocate more external R&D to the regions that already implement more internal R&D. This could therefore artificially induce the complementarity between internal and external R&D that we want to test. We addressed this potential problem by testing several distribution methods for external R&D: a) even distribution across firms' establishments; b) distribution according to establishments' internal R&D expense; c) distribution according to establishments' entire R&D workforce; d) allocation according to the number of researchers in the establishments. We also tested an allocation method e) that computes the external/internal R&D ratio at the company-level and applies it to the average internal R&D expense of all the establishments located in the same NUTS3 region belonging to other firms and possessing the same research code, that is to say specialized in similar technological domains[6]. To save space, we only display results obtained with method d) and e). We can provide the other results upon request; they do not differ noticeably from those presented here.

We finally end up with several time-varying regional internal and external R&D variables (Table 1). As usual, we consider that there is a time lag between R&D expenses and innovation. In his seminal paper, Jaffe (1986) argues that "we expect knowledge production

[6] We thank referee two for this suggestion.

to depend on a distributed lag of R&D, but this lag structure is difficult to identify, and much of the weight appears to fall on the contemporaneous R&D". Accordingly, Acs, Anselin and Varga (2002) and Gumbau and Albert (2009), regress patents on contemporaneous R&D variables. However, in a paper that explicitly deals with the issue of identifying the lag between R&D and patents, Hall et al. (1984) concluded that "there is a significant effect of R&D on patenting (with most of it occuring in the first year)". Accordingly, Fritsch and Slavtchef (2011) regress patents on a one-year-lagged R&D workforce variable, and Gurmu et al. (2010) use 18 months-lagged R&D expense variables. We consider that using contemporaneous R&D reinforces endogeneity issues and, therefore, we choose the one year lag as our preferred specification. We also tested specifications, not displayed here, with R&D variables averaged over years t-1 and t-2, which implies an average lag of 18 months, and with two years-lagged R&D variables.

3.1.3. Control variables

The literature on regional innovation production has abundantly demonstrated that the R&D effort is not the sole determinant of patent production at this level of aggregation. One must also account for the influence of specialization externalities (Marshall, 1890), diversity and urbanization externalities (Jacobs, 1969), regional human capital (Lucas, 1993), and trade linkages created by the internationalization of regional firms (Amin and Cohendet, 2004, Nooteboom, 2009, Boschma and Iammarino, 2009).

We construct an index of regions' relative industrial specialization based on the industry classification of R&D employment. This seems more relevant than the use of total employment because diversification economies refer not only to the variety of industrial sectors but also to the diversity of cognitive and technological competencies. Note also that the use of R&D *employment* rather than R&D expenses produces a less volatile index. Since we have R&D workforce data at the business-unit level, we can compute the R&D

employment share of each industry in each region. Our data allowed us to differentiate industries at the 'NAF60' level. 'NAF' is the French "Nomenclature d'Activités Française" similar to the SIC classification. It classifies each business unit according to its principal activity. 'NAF60' means that we are able to differentiate sixty different activities in our index. It roughly corresponds to a two-digit SIC classification. This remains a fairly aggregated level that probably tends to overestimate positive specialization externalities (Beaudry and Shiffauerova (2009)). However, this will not prove to be a problem since what we eventually detect in our econometric estimates is positive *diversification* externalities. Regarding the mathematical formula of the specialization/diversification indicator, we did not opt for a simple Herfindahl index (location quotient) because it does not account for the heterogeneity of business units dispersion across regions. We therefore prefer an Ellison-Glaeser index (Ellison and Glaeser, 1997). It follows the formula:

$$EGindex_{it} = -\frac{G_{it} - H_{it}}{1 - H_{it}}$$

with:

$$H_{it} = \sum_{e} \left(\frac{RD_{et}}{RD_{it}} \right)^{2} \text{ and } G_{it} = \frac{\sum_{k} (S_{ikt} - S_{kt})^{2}}{1 - \sum_{k} S_{kt}^{2}}$$

where S_{ikt} is the share of sector *k* R&D in region *i* R&D employment at year *t*, S_{kt} is the share of sector *k* R&D in national R&D employment at year *t*, RD_{et} is establishment *e* R&D employment at year *t* and RD_{it} is region *i* R&D employment at year *t*. Regions with a high *EGindex* display a high diversity of their R&D activities. In contrast, regions with a low *EGindex* are characterized by R&D activities that are more concentrated on some specific sectors.

To account for urbanization economies, we introduce regional population densities computed with the annual population estimations provided by the French National Institute of

Statistics (INSEE) divided by NUTS3 regions' surfaces. This variable also controls for size effects since it includes the time-varying regional population at the numerator.

Since we do not have data on NUTS3 trade balances, we construct a proxy of regions internationalization based on information extracted from the R&D survey. For each establishment located in a NUTS3 region, the survey indicates whether it belongs to a domestic company or to a foreign one. We thus compute the percentage of each region's establishments belonging to a foreign company and use it as a regional internationalization index.

Finally, we construct a proxy of regional human capital. We use the regional population censuses provided by the French National Institute of Statistics to compute the share of people aged between 25 and 54 holding a graduate or post-graduate diploma. Since this information is only available every ten years, we could only compute this proxy for year 1999. This is therefore a time-invariant variable in the panel regressions, which prevented us from using fixed-effects estimators once this control was introduced.

PLEASE INSERT TABLE 1

3.2. Econometric Methodology

The general form of the knowledge production equation that we estimate is:

(Equation 1)

$$\log(pat_{it}) = \alpha + \beta_1 \log(R\,\&\,D\ variables_{it-1}) + \gamma_1 Time - varying\ controls_{it} + \lambda_1 Time\text{ - }invariant\ controls_i + \sum_t \mu_t\, time_t + u_i + \varepsilon_{it}$$

Where pat_{it} is the total number of patents filed by region i at year t, $time_t$ is a time dummy equal to 1 at years t=1997…2008, u_i is an unobserved individual effect and ε_{it} is the usual idiosyncratic error term. The other variables are defined in Table 1. We introduce the R&D covariates one by one and then altogether. The main time-varying controls are introduced in each regression. We then introduce the human capital index as a time-invariant control.

We first implemented random and fixed-effects regressions. The Hausman test always suggested to reject the hypothesis that u_i is uncorrelated with the covariates. Normally, this leads to implement individual fixed-effects regressions that are robust to this correlation. However, the rejection of the null hypothesis in the Hausman test can always come from the fact that the model is mispecified because an important time-invariant independent variable could not be introduced in the FE regression. Since we want to use such a covariate (the regional human capital index), we have to implement another kind of estimator that is robust to the suspected correlation between some covariates and the unobserved individual effect u_i. That is the reason why we use the estimator proposed by Hausman-Taylor (1981). This method is designed for panel-data random effects models wherein some covariates are correlated with the unobserved individual-level random effect. It assumes however that the covariates are independent of the idiosyncratic error term $\varepsilon_{it,}$, which means that the endogeneity problem is located in the individual rather than in the time dimension. This is a logical hypothesis here because our R&D covariates are lagged and vary a lot in the individual dimension (see Table 1). The Sargan-Hansen overidentifying restrictions tests validate this assumption in each regression. The Hausman-Taylor estimator is based on GLS instrumental variables regressions producing consistent and efficient estimators of the coefficients, provided that the instruments respect some conditions. No external instruments are needed. Internal instruments are constructed using averaged and demeaned covariates.

3.3. Results of panel estimates

Table 2 presents the results of within (FE) and Hausman-Taylor (HT) regressions wherein we introduce in house and external R&D variables independently and then jointly. In the FE regressions, the standard-errors are computed with the Huber-White-Sandwich method to account for between-heteroskedasticity and within-autocorrelation.

The only control variable that appears significant across all estimators is the internationalization index. It has the expected positive sign. In Hausman-Taylor estimations, the industrial diversity index is significant with a positive sign suggesting the presence of positive diversification externalities, and the population density is positive significant which suggests positive urbanization economies. The human capital index is never significant except if we remove the population density variable (results not displayed here).

The elasticity of the internal R&D variable has a level, sign and significance that is consistent across all regressions: a 1% increase in regional in house R&D produces a 0,1% rise in patenting. It is coherent with comparable empirical studies on French regions (Massard and Riou, 2002, Autant-Bernard and LeSage, 2011), and with similar studies on other European regions (e.g., Bottazzi and Peri, 2003, Ponds et al., 2010). The introduction of external R&D, and human capital as a supplementary control, never changes this elasticity.

External R&D is not significant when introduced alone in the FE regression controlled with the three time-variant controls (population density, diversity index, internationalization index). Introducing internal R&D and human capital does not change this result but the coefficient of outsourced R&D becomes negative and its Student statistic becomes much larger. This happens in the fixed-effect regression of column (3) and in the Hausman-Taylor regression of column (4). We thus suspect that some subcategories of outsourced R&D may have in fact a negative impact. We check that, introducing a differentiation between domestic outsourced R&D, foreign outsourced R&D and affiliate outsourced R&D in the Hausman-Taylor estimation displayed in column (5). We obtain that the former has a significant negative impact, a result already present in Bertrand and Mol (2013). The two other external R&D variables have no significant impact. Note that we obtain the same result in a fixed-effect regression, not displayed here, wherein the time-invariant human capital variable must be removed. Note also that in the results displayed in Table 2, outsourced R&D is allocated in regions with the methodology e) mentioned in section 3.1.2. above (columns (2), (3) and (4)), and with methodology d) in columns (5). We tested all the other methods and this produced no change except that foreign outsourced R&D also becomes significant negative when we use method c). We can provide the results upon request.

In summary, our aggregate regional-level estimations provide results that differ markedly from those obtained by the previous studies implemented at the individual firm-level. Total outsourced R&D is far from being significant at the regional level, and domestic R&D outsourcing has a significant negative impact. This latter result is also found at the firm-level by Bertrand and Mol (2013) who also work on France. They interpret this as evidence that domestic R&D outsourcing does not bring enough new knowledge because of too limited cognitive distance between the R&D buyer and the R&D provider. We share this interpretation but, contrary to them, we do not find a positive impact of foreign R&D outsourcing that could be interpreted as evidence that R&D outsourcing at larger cognitive and geographical distance is positive for innovation. This does not work at the regional level.

Outsourced R&D alone could very well be non-significant but still positive significant when combined with a sufficient level of internal R&D. Therefore, we now have to test at the aggregate level the potential complementarity between internal and outsourced R&D. For that purpose, we introduce crossed variables interacting in house and outsourced R&D. The results are displayed in Table 3. To warranty robustness, we again implemented these tests with outsourced R&D expenses computed with all the methods mentioned in section 3.1.2. above. To save space, we only display here the results obtained when external R&D is computed according to methods e) (columns (1) and (2)) and with method d) (columns (3) and (4)). Again, we found no different results with the other methods.

We first cross the total outsourced R&D variable with in house R&D (columns (1) and (3) in Table 3) in specifications that also include the direct effects of internal and external R&D as well as the controls previously employed. The crossed R&D variables are never significant, and the coefficients of the other covariates do not change in comparison to previous estimations. There could be a threshold effect in the complementarity between the two R&Ds implying that it appears only at high levels of internal R&D. We test this with an interaction term wherein we no longer use the level of in house R&D but, instead, a dummy variable equal to one whenever the internal R&D of region *i* at year $t{-}1$ is in the top 33% of regions' internal R&D expenses in this year. Again the crossed R&D variables are never significant, but we obtain now a negative impact of the direct external R&D variable when it is estimated with method d) (column (4)). All in all, we find no evidence of a complementarity between in house and outsourced R&D at the aggregate regional level.

We also have to mention that we tested the robustness of all the results displayed in Tables 2 and 3 in various ways. Firstly, we implemented regressions where the R&D variables are averages of their $t{-}1$ and $t{-}2$ values, which amounts to creating R&D lags of 18 months. The results are not changed. However, at larger lags (t-2 and more), the R&D variables become non-significant. We also implemented the estimations removing the French "départements" composing respectively the "Régions" "Ile de France" (The Paris NUTS2 region, composed of 8 NUTS3 regions) and "Rhône-Alpes" (The Lyon NUTS2 region, composed of 8 NUTS3 regions), because patenting and R&D is highly concentrated within these regions, which imply that they could over-determine the results. Again, we obtained results that do not differ significantly. Lastly, we have tested whether our regional innovation and R&D data could be subject to a spatial autocorrelation that may have biased the estimates. To check this point, we tested various spatial specifications of equation 1 but never obtained that controlling for spatial autocorrelation changes the previous results. This is in line with other studies on French regions that get weak spatial spillovers at this level of aggregation. For example, Autant-Bernard (2001) finds that the spillovers of public R&D in France do not diffuse beyond the frontiers of NUTS3 regions. In a comprehensive study of the localization of French innovative activities across NUTS3 regions, Moussa (2012) applies various spatial autocorrelation tests and shows that R&D and patent counts are strongly concentrated but only weakly auto-correlated across French "départements". At the European level, Bottazi and Peri (2003) find, that spatial R&D spillovers are weak and do not diffuse beyond 300km circles.

4. Conclusions

We addressed an issue that has been largely ignored by the empirical literature dealing with innovation production: even if it is generally positive at the firm level, should we consider that the influence of R&D externalization is also positive for innovation at aggregate regional or national levels? If the answer is yes, regional innovation policies should support R&D outsourcing wherever it is efficient for the individual businesses. We have argued that knowledge externalities could contradict this view because they may create a fallacy of composition of R&D outsourcing: even if individual firms that outsource R&D become more innovative, it could very well be that too much R&D externalization produces an impoverishment of territories' R&D, and therefore a negative knowledge externality leading to less innovation.

Our empirical results provide evidence in favour of this fallacy of composition of R&D outsourcing: in French NUTS3 regions, we find that total outsourced R&D has no significant impact on regions' patenting. Moreover, we obtain that the impact of domestic R&D outsourcing is negative, and we find no evidence of a complementarity between in house and outsourced R&D at this level of aggregation.

Since there is a dearth of empirical studies of R&D outsourcing at the aggregate level, these results are still to be confirmed by other similar studies on different geographical units. If the results were to be confirmed, regional policy makers would have to consider cautiously the R&D externalization strategies that are sometimes supported in the name of "smart specialization". Outsourcing strategies may allow regions to specialize in the R&D activities they best perform, and provide new knowledge from abroad, but local knowledge production by in house R&D may remain the only form of R&D having a significant and positive impact on innovation production at the aggregate level.

References

Amin, A. and Cohendet, P. (2004) *Architectures of Knowledge: Firms, Capabilities and Communities* (Oxford: Oxford University Press).

Acs, Z., Anselin, L. and Varga, A. (2002) Patents and innovation counts as measures of regional production of new knowledge, Research Policy, 31.

Arvanitis, S. and Loukis, E.N. (2012) Outsourcing and firm performance: a comparative study of Swiss and Greek firms, *Industrial and Corporate Change*, 22(3), pp. 771-806.

Audrescht, D. (1998) Agglomeration and the location of innovative activities, *Oxford Review of Economic Policy*, 14(2).

Autant-Bernard, C. and LeSage, J.P. (2011) Quantifying Knowledge Spillovers Using Spatial Econometric Models, *Journal of Regional Science* 51(3) 471-496.

Beaudry, C. and Schiffauerova, A. (2009) Who's right, Marshall or Jacobs? The localization versus urbanization debate, *Research Policy* 38(2) 318-337.

Boschma, R.A. and Iammarino, S. (2009) Related Variety, Trade Linkages and Regional Growth, *Economic Geography* 85 (3): 289-311.

Bottazzi, L. and Peri, G. (2003) Innovation and spillovers in regions: evidence from European patent data, *European Economic Review* 47 (7) 687-710.

Brossard, O. and Vicente, J. (2010) Knowledge phases, cognitive and relational distance in ICT alliance networks, in Cooke Ph., De Laurentis C., Macneill S. & Collinge Ch. (eds.), *Platforms of Innovation : dynamics of new industrial knowledge flows*, Edwar Elgar.

Cantwell, J. and Santangelo, G.D. (1999) The frontier of international technology networks: sourcing abroad the most highly tacit capabilities, *Information Economics and Policy*, 11(1): 101-123.

Cassiman, B. and Veugelers, R. (2006) In search of complementarity in innovation strategy: internal R&D and external technology acquisition, *Management Science* 52, 68–82.

Chesbrough, H.W. and Appleyard, M.M. (2007) Open innovation and strategy, *California Management Review*, 50, 1, 57-76.

Cohen, W. and Levinthal, D. (1990) Absorptive capacity: a new perspective on learning and innovation, *Administrative science quarterly* 35(1).

Colander, D., Howitt, P., Kirman, A., Leijonhufvud, A. and Mehrling, P. (2008) Beyond DSGE Models: Toward an Empirically Based Macroeconomics, *The American Economic Review*, 98, 2.

Cooke, P. (2006) Rational durg design, the knowledge value chain and biosciences megacentres, *Cambridge Journal of Economics*, 29, 325-341.

Cowan, R. and Foray, D. (1997) The economics of codification and the diffusion of knowledge, *Industrial and Corporate Change* 6 595-622.

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

de Rassenfosse, G. and van Pottelsberghe de la Potterie, B. (2009) A policy insight into the R&D–patent relationship, *Research Policy* 38 779–792.

Ellison, G. and Glaeser, E. L. (1997) Geographic concentration in U.S. manufacturing industries: A dartboard approach, *Journal of Political Economy* 105: 889-927.

Feldman, M. P. and Kelley, M.R. 2006, The *ex ante* assessment of knowledge spillovers: Government R&D policy, economic incentives and private firm behavior, *Research Policy* 35, 1509–1521.

Frenken, K., Van Oort, F. and Verburg, T. (2007) Related Variety, Unrelated Variety and Regional Economic Growth, *Regional Studies* 41 685–697.

Gertler, M.S. (2003) Tacit knowledge and the economic geography of context, or the indefinable tacitness of being (there), *Journal of Economic Geography*, 3 75-99.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P. and Trow, M. (1994) *The new production of knowledge: the dynamics of science and research in contemporary societies* (London: Sage).
Gilsing, V., Nooteboom, B. (2006) Exploration and exploitation in innovation systems: the case of pharmaceutical biotechnology, *Research Policy* 35, 1-23.
Griffith, R., Redding, S., Van Reenen, J. (2004) Mapping the two faces of R&D: productivity growth in a panel of OECD industries, *The Review of Economics and Slatisiics,* November, 86(4): 883-895
Griliches, Z. (1979) Issues in Assessing the Contribution of R&D to Productivity Growth, *Bell Journal of Economics,* Spring 10 92-116.
Griliches, Z. (1990) Patent statistics as economic indicators: A survey, *Journal of Economic Literature* 28 1161-1707.
Griliches, Z. (1991) The search for R&D spillovers, *NBER Working Paper,* n° 3768, July.
Grimpe C, Kaiser U. (2010) Balancing internal and external knowledge acquisition: the gains and pains from R&D outsourcing, *Journal of Management Studies* 47: 1483–1509.
Gurmu, S., Black, G. and Stephan, P. (2010), The knowledge production function for university patenting, *Economic Inquiry*, 48(1).
Hagedoorn, J. and Wang, J. (2012) Is there complementarity or substitutability between internal and external R&D strategies?, *Research Policy,* 41.
Hall B.H., Hausman J.A. and Griliches, Z. (1984) Is there a lag? *NBER Working Paper,* n°1454, September.
Hausman, J. A. and Taylor, W.E. (1981) Panel data and unobservable individual effects, *Econometrica,* 49: 1377-1398.
Hess, A.M. and Rothaermel, F.T. (2011) When are assets complementary? Star scientists, strategic alliances, and innovation in the pharmaceutical industry, *Strategic Management Journal* 32.
Howells, J. (1996) Tacit knowledge, innovation and technology transfer. *Technology Analysis and Strategic Management*, 8, 91–106.
Howells J, Gagliardi D, Malik K. (2008) The growth and management of R&D outsourcing: evidence from UK pharmaceuticals. *R&D Management* 38: 205–219.
Hsuan J. and Mahnke V. (2011) Outsourcing R&D: a review,model, and research agenda, *R&D Management*, 41, 1.
Jacobs, J. (1969) *The Economy of Cities* (New York: Random House, Vintage).
Jaffe, A. (1986) Technological Opportunity and Spillovers of R&D: Evidence from Firms Patents, Profits, and Market Value, *American Economic Review* 76(5) 984-1001.
Katz, R. and Allen, T. J. (1982) Investigating the not invented here (NIH) syndrome. A look at the erformance, tenure, and communication patterns of 50 R&D project groups, *R&D Management*, 12, 7–19.
Kirman, A. (1992) Whom or What Does the Representative Individual Represent? *The Journal of Economic Perspectives* 6(2): 117-36.
KPMG (2008). 'Knowledge process outsourcing'. *Global Issues and Insights*. 14 February.
http://www.kpmg.com/Global/en/IssuesAndInsights/ArticlesPublications/Pages/Knowledgeprocess-outsourcing-article.aspx
Lissoni, F. (2001) Knowledge codification and the geography of innovation: the case of Brescia mechanical cluster, *Research Policy,* 30(9), December.
Lucas, R.E., 1993. Making a miracle. *Econometrica* 61, 251–272.
Mahnke, V. (2007) Communication metaphors-in-use: technical communication and offshore systems development. IEEE Transactions on Professional Communication, 50, 2, 93–108.
March JG. (1991) Exploration and exploitation in organizational learning. *Organization Science* **2**: 71–87.

Marshall, A. (1890) Principles of Economics: An Introductory Volume, *Rod Hay's Archive for the History of Economic Thought* (McMaster University, Canada, Online Version).
Massard, N. and Riou, S. (2002) L'impact des Structures Locales sur l'Innovation en France: Spécialisation ou Diversité? *Région et Développement* 16 111-136.
Mol MJ. (2005) Does being R&D intensive still discourage outsourcing? Evidence from Dutch manufacturing. *Research Policy* **34**: 571–582.
Moussa, I. (2012) Proximité géographique, diffusion des connaissances et innovation: une étude sur les départements français métropolitains, PhD thesis dissertation, Toulouse 1 University.
Mudambi, S. M. and Tallman, S. (2010) Make, buy or ally? Theoretical perspectives on knowledge process outsourcing through alliances. *Journal of Management Studies*, 47:8.
Narula, R. (2001) Choosing between Internal and Non-internal R&D Activities: Some Technological and Economic Factors, *Technology Analysis and Strategic Management* 13(3) 365-87.
Nelson, R. and Winter, S. (1982) *An Evolutionary theory of economic change,* Production Sets and Organizational Capabilities (Belknap/Harvard).
Nooteboom B. 2009. *A Cognitive Theory of the Firm: Learning, Governance and Dynamic Capabilities*. Edward Elgar: Cheltenham, UK.
Polanyi, M. (1966) *The tacit dimension* (New York: Doubleday).
Ponds, R., Oort, F. and Frenken, K. (2010) Spillovers and university–industry collaboration: an extended knowledge production function approach, *Journal of Economic Geography* 10 231–255.
Rondé, P. and Hussler, C. (2005) Innovation in regions: what does really matter? *Research Policy* 34 1150-1172.
Tsai, K.H. and Wang, J.C. (2008) External technology acquisition and firm performance: a longitudinal study, *Journal of Business Venturing* 23 91-112.
Vega-Jurado, J., Gutiérrez-Gracia, A. and Fernández-de-Lucio, I. (2009) Does external knowledge sourcing matter for innovation? Evidence from the Spanish manufacturing industry" *Industrial and Corporate Change* 18, 637–670.

Table 1 : Variables definitions and descriptive statistics

Variable	Definition		Mean	Std. Dev.	Min	Max	Observations
Patents	Number of patent applications (all patents) of region i at year t	overall	159.8367	249.4873	.9	1952.756	N = 1128
		between		236.5444	2.99418	1260.317	n = 94
		within		82.68485	-179.2697	1093.836	T = 12
In house R&D	In house R&D expense of region i at year t	overall	692262.4	1885195	1	2.25e+07	N = 1128
		between		1454179	1750.35	1.01e+07	n = 94
		within		1208291	-6421794	1.32e+07	T = 12
Outsourced R&D	Estimated R&D outsourcing of region i at year t, with various estimation methods (see main text): here the outsourced R&D is distributed accross firms' business units according to their share in firms' total researchers workforce	overall	142524.5	556317.7	1	7395033	N = 1128
		between		409110.3	67.83333	2927453	n = 94
		within		379146	-1913943	4610104	T = 12
Population density	Population density of region i at year t	overall	543.1186	2376.758	14.17051	21059.97	N = 1128
		between		2387.962	14.53687	20502.26	n = 94
		within		47.69581	144.3691	1100.826	T = 12
Domestic R&D outsourcing	(Estimated) domestic outsourcing of region i at year t. Represents all spending on R&D transactions with independent R&D suppliers located in France.	overall	82960.9	357898.6	1	4749196	N = 1128
		between		261852.5	65.16667	1815864	n = 94
		within		245344.4	-1418150	3016292	T = 12
Foreign R&D outsourcing	(Estimated) offshore outsourcing of region i at year t. Represents all spending on R&D transactions with independent R&D suppliers located abroad.	overall	18236.17	83780.56	.5999985	1347282	N = 1128
		between		57103.69	1	449623.7	n = 94
		within		61564.42	-330344.8	1013511	T = 12
Affiliate R&D sourcing	(Estimated) affiliate R&D sourcing of region i at year t. Includes all R&D procurement from other units of company.	overall	42335.27	146160.8	1	2495240	N = 1128
		between		102310.7	1	793256.6	n = 94
		within		104869.8	-634944.5	1744318	T = 12
Industrial diversity index	Ellison-Glaeser index of technological and industrial diversity of region i at year t	overall	.0214649	.1219489	-1.424775	.9928296	N = 1128
		between		.069782	-.2313505	.3071793	n = 94
		within		.1002474	-1.17196	1.135218	T = 12
Internationalization index	Share of the regions' business units belonging to a foreign company in region i at year t	overall	.1807886	.1050502	0	1	N = 1128
		between		.0775435	.0138889	.3403622	n = 94
		within		.0712828	-.0692114	.9307886	T = 12
Human capital index	Share of people aged between 25-54 holding a graduate or post-graduate diploma in region i in year 1999	overall	.215445	.0592637	.1510136	.541601	N = 1128
		between		.0595551	.1510136	.541601	n = 94

Table 2 : Panel estimates of patenting in French regions

	(1: FE) log(pat_{it})	(2: FE) log(pat_{it})	(3: FE) log(pat_{it})	(4: HT) log(pat_{it})	(5:HT) log(pat_{it})
log(In house $R\&D_{it-1}$)	0.0827+		0.0974*	0.101***	0.0996***
	(0.0427)		(0.0481)	(0.0255)	(0.0236)
log(Outsourced $R\&D_{it-1}$)		0.00361	-0.0191	-0.0183	
		(0.0245)	(0.0246)	(0.0140)	
log(Domestic outsourced $R\&D_{it-1}$)					-0.0180**
					(0.00672)
log(Foreign outsourced $R\&D_{it-1}$)					0.00115
					(0.00716)
log(Affiliate outsourced $R\&D_{it-1}$)					-0.000226
					(0.00684)
log(Population $density_{it}$)	1.044	1.220	1.071	0.628***	0.630***
	(1.504)	(1.487)	(1.487)	(0.0949)	(0.0990)
Industrial diversity $index_{it}$	0.271	0.291	0.290	0.280*	0.297**
	(0.208)	(0.227)	(0.208)	(0.114)	(0.113)
Internationalization $index_{it}$	0.407*	0.469*	0.417*	0.460**	0.434**
	(0.182)	(0.188)	(0.184)	(0.168)	(0.167)
Human capital $index_{it}$				3.067	3.023
				(1.934)	(2.013)
Constant	-1.706	-1.483	-1.813	-0.493	-0.491
	(6.804)	(6.892)	(6.716)	(0.359)	(0.370)
Observations	1128	1128	1128	1128	1128
Sargan-Hansen statistics				Chi2(3)=4.06	Chi2(3)=4.01
				p-value=0.26	p-value=0.27

Standard errors in parentheses. There are computed according to the Huber-White-Sandwich method in FE regressions.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Columns (1), (2) and (3) are fixed-effect regressions with time and individual (region) fixed-effects. Columns (4) and (5) display Hausman-Taylor instrumental variables estimators. The instrumented variables are log(in house $R\&D_{it-1}$) and log(Outsourced $R\&D_{it-1}$) in column (4) and log(in house $R\&D_{it-1}$), log(Domestic outsourced $R\&D_{it-1}$), log(Foreign outsourced $R\&D_{it-1}$) and log(Affiliate outsourced $R\&D_{it-1}$) in column (5). The instruments are means of covariates and demeaned covariates. See Hausman and Taylor (1981). Sargan-Hansen statistics show that the instruments are valid.

Table 3: test of the complementary between in house and outsourced R&D in Hausman-Taylor IV regressions

	(1) log(patit)	(2) log(patit)	(3) log(patit)	(4) log(patit)
log(In house $R\&D_{it-1}$)	0.0904***	0.103***	0.115***	0.115***
	(0.0273)	(0.0260)	(0.0265)	(0.0249)
log(Outsourced $R\&D(1)_{it-1}$)	-0.0443	-0.0179		
	(0.0285)	(0.0141)		
Log(In house $R\&D_{it-1}$)×Log(Outsourced $R\&D(1)_{it-1}$)	0.00252			
	(0.00241)			
log(Outsourced $R\&D(1)_{it-1}$)×dummy for regions wherein internal R&D is high		-0.00178		
		(0.00476)		
log(Outsourced $R\&D(2)_{it-1}$)			-0.0333	-0.0410**
			(0.0283)	(0.0138)
Log(In house $R\&D_{it-1}$)×Log(Outsourced $R\&D(2)_{it-1}$)			-0.000826	
			(0.00253)	
log(Outsourced $R\&D(2)_{it-1}$)×dummy for regions wherein internal R&D is high				-0.00375
				(0.00488)
log(Population $density_{it}$)	0.619***	0.627***	0.647***	0.643***
	(0.0964)	(0.0948)	(0.0986)	(0.0982)
Industrial diversity $index_{it}$	0.294*	0.279*	0.300**	0.302**
	(0.114)	(0.114)	(0.114)	(0.114)
Internationalization $index_{it}$	0.483**	0.458**	0.478**	0.480**
	(0.169)	(0.168)	(0.169)	(0.168)
Human capital $index_{it}$	2.794	3.111	3.293	3.296
	(1.975)	(1.935)	(2.020)	(2.006)
Constant	-0.341	-0.530	-0.534	-0.546
	(0.390)	(0.367)	(0.394)	(0.375)
Observations	1128	1128	1128	1128
Sargan-Hansen statistics	Chi2(3)=3.76	Chi2(3)=4.11	Chi2(3)=4.17	Chi2(3)=4.14
	p-value=0.29	p-value=0.25	p-value=0.24	p-value=0.25
Wald Chi2 test[a]		1.84		
Wald Chi2 test[b]				9.6***

Standard errors in parentheses; + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$