In & Out:
Inventing with firm internal and external knowledge spillovers

Marco S. Giarratana
Business Department
Universidad Carlos III de Madrid
marco.giarratana@uc3m.es

Myriam Mariani
Department of Institutional Analysis and Public Management, Bocconi University
marianim@unibocconi.it

Abstract

This paper focuses on knowledge spillovers that craft an invention. It uncovers the factors that determine the relative importance of knowledge spillovers coming from outside the inventor organization compared to spillovers internal to the inventor organization. Our empirical test employs a unique and extensive dataset that provides direct measures of such external and internal spillovers from a large scale survey of European patents. Consistently with our theory, we find that the importance of spillovers external to the firm is positively correlated with the costs of the invention and with the level of the knowledge endowment of the geographical location where the research is performed. Conversely, the magnitude of the firm cumulative stock of internal knowledge favors the use of spillover internal to the firm. Our data also corroborate the existence of a competitive effect: Firms rely to a lower extent on external spillovers if high costly inventions are developed in regions characterized by larger knowledge endowments.

VERY PRELIMINARY DRAFT – June 2008

1. Introduction

Inventive capability is an important source of competitive advantage, particularly in high-tech and turbulent industries characterized by fierce product competition. (Schmalensee 2000; Ilinitch et al. 1996; Sorenson, 2000). Yet, in these industries, firm inventive performance does not reside only in the capability of developing completely new ideas, but also in the ability to get access to and to combine knowledge assets from extra-mural sources (Teece 1986, Von Hippel, 1986; Chesbrough, 2003).

In this respect, the production and management of innovation has become an increasingly difficult task, as knowledge created inside the organization is often combined with knowledge generated outside it. The economic and management literature addresses this issue from different perspectives. There are contributions on the role of absorptive capacity in innovation (Cohen and Levinthal 1990), on the complementarity between external and internal sources of knowledge (Arora and Gambardella, 1990; Cassiman and Veugelers, 2006), on the impact of knowledge externalities on research productivity (Laursen and Salter, 2007), on the rise of open innovation business models (Chesbrough 2003; Helfat and Quinn, 2006) and on the importance of alliances and networks for innovation (Sakakibara, 2001; Ahuja, 2000).

More recently, following the Marshallian tradition, a stream of the literature argues that the leverage between the use of firms' internal and external knowledge sources depends also on the characteristics of the location of the firm, as it affects the firm exposure to potential knowledge spillovers (e.g. Saxenian, 1994; Audretsch and Feldman, 1996; Tallman et al., 2004; Alcacer and Chung, 2007).

In order to study the importance of spillovers external to the inventors' organization (hereafter *External*) compared to the internal (hereafter, *Internal*) in the development of patented inventions, we adopt a theoretical perspective that takes into account three layers of analysis: the R&D project level, the firm level and the geographical location level. Accordingly, we the elaborate four hypotheses that state that inventors rely on *External* more than *Internal* knowledge spillovers when: i) the cost of the R&D project is higher; ii) the technological endowment of the region where the firm is located is richer. By contrast, *Internal* spillovers are expected to be more important than those provided externally when iii) the firm cumulative stock of knowledge is larger and; iv) a larger R&D investment is made to develop an invention in a region that own a greater knowledge technological endowment.

We use the patent as the unit of observation and this research is made possible by unique data that we collected by means of a large scale survey of European patents (the PatVal-EU survey, see Giuri et al., 2007). The empirical analysis is based on information provided directly by the inventors of these patents, which we complement with other data on firms and locations derived from standard datasets. The empirical results support our hypotheses.

This study makes four major contributions to the existing literature on this issue. First, in terms of absorptive capacity and R&D complementarity literature, we separate the effects due to R&D project costs from those due to the firm stock of knowledge. Indeed, R&D project costs lead to use *External* spillovers more than *Internal* spillovers; conversely, a larger stock of knowledge accumulated by the firm produces the opposite effect, favoring an inward looking behavior. These findings provide a novel insight on the determinants of the "trade-off between inward looking vs. outward looking absorptive capacity" (Cohen and Levinthal, 1990: 133).

Second, our results show that firms located in regions with a richer knowledge environment do also rely on *External* more than *Internal* spillovers. Theoretically, we support this empirical evidence with a original perspective that treats the geography literature (Saxenian, 1994; Feldman, 1999) and the sociological approach (Fleming, 2007) as complements.

Third, in a paper on firm co-location decisions, Alcacer (2006) describes the trade-off that co-location creates: on the one hand, there are spillover advantages from R&D co-located activities; on the other hand, there are costs due to the increasing competition in the final markets. Following more closely Alcacer and Chung (2007), our study confirms that a similar trade-off exists not only at the level of the firm but also at the level of the invention process, in the choice between the internal and external sources of knowledge. Indeed, we find that as the costs of an invention project increases, the potential threats of imitation that a firm perceives from its close geographical environment hamper the openness of its R&D posture, reducing the importance of external spillovers

Finally, this article uses a direct indicator of the importance of knowledge spillovers, which was provided directly by the inventors of the patents. This avoids the problems of measuring knowledge spillovers with indirect measures like patent citations (Alcacer and Gittelman, 2006; Harhoff et al. 2006).

The paper is organized as follows: the next section develops the hypotheses. Section 3 describes the datasets, the empirical model and the results. Section 4 concludes.

2. Theory

Our theoretical background develops through three levels of analysis: the R&D project or activity, the firm that performs the R&D activity and the location where this R&D activity is conducted.

As far as the first level is concerned – i.e. the research activity or project – we focus on the cost of the R&D project that led to the invention. The cost of producing an invention usually influences the level of innovation complexity and the extent of economies of scale and scope (Krishnan and Ulrich 2001; Henderson and Cockburn, 1994; Danneels, 2002). The second level is the firm. As the neo-Schumpeterian have stressed (Nelson and Winter, 1982; Dosi, 1988), research activity is path-dependent and guided by the past R&D history of the firm. Therefore, we investigate the effect of the knowledge accumulated inside the organization on the use of *External* vs. *Internal* knowledge spillovers. Finally, we move to the regional level and, as suggested by the contributions on the geography of innovation (e.g., Romanelli and Khessina, 2005; Tallman et al., 2004; Saxenian, 1994), we explore the effect of the knowledge endowment of the location where the R&D activity is conducted. Our object of analysis will be the importance of *External* vs. *Internal* knowledge spillovers in crafting a R&D project. Figure 1 summarizes our theoretical background.

Figure 1 About Here

The cost of developing an invention is related to the complexity of the underlying research project (DiBiaggio, 2007). Empirical evidence shows a positive relationship between firms' R&D expenditures and the diversification span of the research projects (Granstrand et al., 1997; Rosenberg, 1976). Part of the cost concerns the human resources involved in developing the inventions: the higher the technological complexity of the project, the larger the array of different and complementary competencies that are needed to develop it, and therefore the cost of recruiting the different competencies. Also, the wider the range of different capabilities needed for complex projects, the lower the probability that they are all available in the same organization, which asks for searching them also outside the firm boundaries (Fleming and Sorenson, 2004). These centripetal forces are stronger if the organization needs to match the benchmark technology in different fields, included "peripheral" technologies in which it has weak or no competencies (Hauser, 1998).

In addition, there is the issue of complementarities. There is a consolidated literature on the super-modularity proprieties of internal and external sources of knowledge (Arora and Gambardella, 1990; Cassiman and Veugelers, 2006). A large investment in a research project results in an intense use of firm internal resources. Also, because of knowledge complementarity, this generates a leverage of external knowledge. However, while internal knowledge sources are limited by factors like the history of the firm, path dependence processes in technology development, financial constraints, differently, external sources are, in principle, unbounded. Therefore, a high complementary between internal and external knowledge would lead to use external sources more than those available internally. This line of reasoning would be consistent with the theories on R&D alliances suggested by Sakakibara (2001) and White and Lui (2005) who demonstrate that firms rely more on R&D alliances if the cost of a research project is high. They explain that this is motivated by costs sharing reasons, as well as by the search for complementary capabilities. This arguments lead to our first hypothesis:

Hypothesis 1: Other things being equal, the higher the cost of a research project, the higher the importance of External knowledge spillovers compared to those accessed Internally.

Learning is an important aspect of R&D activities. However, according to Levinthal and March (1993) learning has its own traps. Knowledge development tends to be path dependent and to follow pre-determined and long-run trajectories (Dosi, 1988). The larger the stock of knowledge that a firm owns, the higher the probability that the organization has developed consolidated routines of internal exploration and search (Nelson and Winter, 1982). The scale of a firm knowledge base is strictly related to its R&D experience and identity, which are major sources of the inward-looking behavior in R&D search processes (Tripsas and Gavetti, 2000; Burgelman, 1994). Thus, while on the one hand the cumulated knowledge portfolio of a firm contributes to its past profitability, on the other hand, it favors the inward-looking attitude in research or even the "not-invented-here" syndrome (Allen and Kats, 1982).

In addition, large knowledge repositories internal to the firm increase the likelihood that inventors search inside the organizations when they need to solve specific technical problems. This is not only because of the internal availability of different competences that can be easily employed, but also because it is costly to find motivations of not using internal technological resources. The human resource literature shows that career promotion is more likely to occur when there is deference to internal knowledge (Lee 1997). This is because the results based on internal knowledge exploration and recombination are easier to evaluate (Griffin and Hauser, 1992). The larger is the knowledge library of a firm, the less likely is that inventors will use external knowledge sources, especially in the case of marked vertical hierarchies and strong internal job competition.

Transaction costs also play a role in the decision of whether to use internal and external knowledge sources. The use of internal knowledge requires lower transaction costs, as the resources are accessible at low cost inside the firm boundaries. This suggests that "organizational proximity" could facilitate the exploitation of knowledge spillovers among inventors affiliated to the same organization, without them being geographically close. Also, the reduction of transaction costs could also follow from the fact that firms with a large stock of knowledge provide a structure for technology resources to be organized and easily searchable, both in terms of data warehouse and human resources. As Ahuja (2000) and Uzzi (1997) point out, the creation and flow of knowledge depend on the structure of the network that owns the many pieces of the overall knowledge. Firms with large knowledge bases and technical competencies do also produce a clear understanding about who owns what, i.e. how knowledge is divided and managed among different employees and teams. This reduces transaction costs, as information on the structure of the network is supplied directly by the organization.

Finally, from a strategic point of view, the use of internal knowledge sources and interactions makes the protection of an invention easier and more effective. If inventions are the outcome of knowledge sources owned by the firm, this reduces potential leakages of information to competitors (Arora and Merges, 2004; Menon and Pfeffer, 2003). These arguments lead to our second hypothesis.

Hypothesis 2: Other things being equal, the larger the knowledge base of a firm, the higher the importance of Internal spillovers compared to those available Externally.

Two recent streams of literature suggest that firm geographic location and inventors personal networks might affect the use of external vs. internal knowledge spillovers. Both the geography of innovation literature (Saxenian, 1994; Feldman, 1999; Romanelli and Khessina, 2005) and the sociological literature assume that there are transaction costs in establishing knowledge interactions among individuals and organizations. However, while the geography of innovation approach claims that geographical proximity reduces these costs (i.e. proximity matters), the sociological view (Fleming, 2007) attributes this merit to the pre-existing network ties, which can be irrespective of geographical proximity (i.e proximity does not matter).

The traditional argument about knowledge spillovers being geographically localized stems from the idea that physical proximity makes it easier to access information produced by others (for a survey, see Doring and Schnellenbach 2006, Feldman 1999), that is, knowledge is like a radio frequency with a geographical bounded range. The evidence suggests that inventive activities benefit more than manufacturing from co-location, particularly in skilled and R&D-intensive industries and in sectors that rely to a greater extent on tacit knowledge and learning-by-doing (Pavitt 1987, Audretsch and Feldman 1996). Some authors also argue that knowledge flows are stimulated in some regions more than in others according to their local technological endowment (Almeida and Kogut, 1999; Thompson, 2006, Jaffe et al., 1993). This is because a large pool of co-localized inventors facilitates knowledge interactions due to proximity, cluster identity and recognition (Romanelli and Khessina, 2005; Sexanian, 1994), localized labor mobility (Almeida and Kogut, 1999), the possibility to avoid arm-length contracts (Storper, 1993), and the creation of shared regional knowledge to which outsiders cannot get access (Tallman et al. 2004, Lampel and Shamsie 2003).

Somehow differently, the sociological literature argues that the individuals' pre-existing social links reduce transaction costs in knowledge interactions (see, among others, Breschi and Lissoni 2001, Sorenson and Singh 2007). By studying patenting co-authorship in the US, Fleming et al. (2007) argue that previous working relationships among inventors produce robust ties that are then used for future interactions, also after the inventors move geographically (see also Agrawal et al. 2006). Thus, irrespective of location, personal connections to a rich social network give access to external knowledge flows.

We argue that the sociological view could complement the geographical proximity paradigm. Location might matter for establishing external interactions also when the inventors are connected to a rich social network. This might be because both geographical proximity and social networks are important to reduce transaction costs, or because proximity and networks are not independent one of another: geographical proximity might foster social networking (Bell and Zaheer, 2007) and, vice-versa; social networking can generate proximity (i.e. Klepper, 2007).

Also, proximity to knowledge sources and personal network ties could be complements even when the personal network ties are completely exogenous to location. Indeed, better-networked inventors might have better skills to absorb knowledge produced by others. In other words, inventors in a well-structured network are also endogenously more productive in using external knowledge sources of their location. Thus, if they located in a knowledge intensive region, they have an advantage in exploiting it. Moreover, the reverse effect might also occur: serendipitous interactions fostered by geographical proximity lead to a frequent updating on the type of knowledge that is pivotal for invention. This facilitates the search for this knowledge in the personal social network. In other words, being in a knowledge intensive location increases the frequency of declarative or procedural knowledge renewals (knowing what and how) that, in turn, creates a more intense use of the external network ties (knowing who) (Borgatti and Cross, 2003). This leads to our third hypothesis:

Hypothesis 3: Other things being equal, the richer the knowledge endowment of the geographical location where the research is performed, the higher the importance of External spillovers compared to those Internal to the organization.

The three hypotheses developed so far predict one negative and two positive effects (i.e the firm stock of knowledge vs. the costs of the project and the regional knowledge endowment) on the relative importance of *External* vs. *Internal* spillovers. A natural question is whether the interaction between the two factors that positively affect the reliance on *External* spillovers reinforces this effect. Based on the above theory, we would expect a positive joint effect. However, if one takes into account the strategic behavior of the firm, then the joint increase in the costs of the research project and the knowledge endowment of the firm location could lead to the opposite conclusion. Alcacer and Chung (2007) show that firm location decision is endogenous and that firms locate strategically in order to minimize knowledge spin-out and maximize spin-in. As Alacer and Chung (2007) point out, there is a tradeoff in the location decision: on the one hand, a knowledge intensive location facilitates the use of external spillovers; on the other hand, the richer the knowledge endowment of the location, the higher the probability that valuable knowledge leaks out to competitors. This means that when the cost of the innovation rises, so does the potential loss of the firm in case of direct imitation. Indeed, since expensive projects do also imply larger sunk costs, higher risks and, most likely, higher expected future value, then, also the fear of imitation is higher for the firm. This suggests that it is more likely that the firm will adopt an inward-looking behavior when the R&D costs are higher in order to avoid inadvertent outward knowledge flows if the firm location is extremely conducive to knowledge spillovers. This is consistent with all the motivations provided for our *Hypothesis 2*. As a matter of fact, McEvily and Chakravarthy (2002) find that the complexity and tacitness of technological knowledge can prevent firm's major innovation improvements from imitation.

Hypothesis 4: Other things being equal, the joint effect of the cost of the research project and the knowledge endowment of the region where the R&D project is conducted has a negative impact on the importance of External compared to Internal knowledge spillovers.

3. Empirical evidence

3.1 The PatVal-EU Survey

This paper employs a unique and extensive survey-based dataset that provides direct information on the use of *External* and *Internal* spillovers, as well as additional information, at the level of the specific research project leading to a patent. The database is the outcome of the PatVal-EU survey that interviewed the inventors of 9,550 patents granted by the European Patent Office (EPO) in Denmark, France, Germany, Hungary, Italy, the Netherlands, Spain and the United Kingdom. The patents in the sample have with priority dates between 1993 and 1997.

The PatVal-EU survey was designed to be representative of the universe of patents in the 8 EU countries. Compared to previous surveys on patents, which had limited European coverage and are mostly biased towards large companies, the PatVal-EU survey has a much broader coverage in terms of European countries, technologies, types and size of the applicant organizations. The PatVal-EU's main objective was to collect information on the individual inventors, the underlying invention process, and the resulting patents, including their economic value (see, on this matter, Gambardella et al, 2008). Giuri et al. (2007) report the details of the survey and the key descriptive statistics. This paper uses information on a sub-sample of 6,016 patents invented in Germany, Spain, Italy, the Netherlands and the United Kingdom.

3.2 Econometric Estimation

3.2.1 Dependent Variable and Method of Estimation

One of the merits of this paper is that it documents the use of knowledge spillovers in the inventive process without resorting to indirect indicators like patent citations. To this specific purpose, and more specifically to study the relative importance of spillovers with people external to the inventor's organization compared to spillovers with people affiliated to the same organization, we asked the inventors the following question: "*Were interactions such as discussions, meetings and sources of ideas with the following types of people (apart from co-inventors) important during the research that led to the patented invention? (0 = not used, 1 = not important, 5 = very important):*

- *People belonging to your organization including affiliates;*
- *People belonging to other unaffiliated organizations*." [1]

We use this question to construct the dependent variable of our empirical analysis. Since the unit of the analysis is the patent, for each of them in our sample we use the 0-5 scores of *Internal* and *External* interactions to compute the following standardized difference:

$$StdExtInt = \frac{(External + 1) - (Internal + 1)}{(External + 1) + (Internal + 1)}$$

[1] We explicitly asked the inventors to exclude interactions with co-inventors. We did not ask to exclude informal interactions set up within other forms of collaborative agreements.

This ratio returns the relative importance of *External* vs. *Internal* knowledge spillovers. In so doing it also takes into account the fact that the same difference between *External* and *Internal* might derive from different absolute scores (e.g. 2-1 has a different meaning compared to 5-4). The correlation coefficient between this variable and the -5 to +5 variable computed as *External-Internal* is 0.98 and it is statistically significant at 1% level.

The unconditional probabilities of *Internal* spillovers being more, less or equally important as *External* spillovers are reported in Figure 2. The Figure shows the number and the share of patents with the score of *Internal* being larger than *External* (*StdExtInt*<0), the score of *Internal* being equal to *External* (*StdExtInt*=0) and the score of *External* being larger than *Internal* (*StdExtInt*>0).

Figure 2 and Table 1 About Here

In over half of the patents (60.5%) *Internal* spillovers are more important than *External* spillovers; in 29.0% of the cases *Internal* and *External* spillovers are equally important to develop the inventions; only in the remaining 10.5% of the patents *External* spillovers are more important than *Internal* spillovers. The average importance of *Internal* spillovers (3.15 on a 0-5 scale) is higher than the average importance of *External* spillovers (1.42).

Table 1 reports the share of patents invented with either *Internal* or *External* spillovers, with both *Internal* and *External* spillovers, and with none of them. It shows that only 18.0% of the patents are invented with no *Internal* spillovers. Moreover, in 53.0% of the patents *Internal* spillovers are very important to develop the patent (i.e. score of 4 and 5, not shown in the Table). The share of patents that do not use *External* spillovers is 53.6%, a much higher share than “no *Internal* spillovers”. Also, in most of the remaining cases, *External* spillovers contribute to the invention together with those accesses internally.

Not surprisingly, these results indicate that “organizational proximity” (i.e. the fact that people are affiliated to the same organization) encourages knowledge interactions among people. Differently, the extent to which inventors rely on knowledge spillovers external to the organization is much more limited compared to the use of spillovers inside the organization. This is true, however, unconditionally.

The variable *StdExtInt* is the dependent variable of our equation. Since it ranges between two limits, i.e. -0.714 and 0.714, we employ a two-limit tobit model, which is used when the dependent variable is truncated at both high and low values (Maddala, 1983). To test our hypotheses, we include covariates for the characteristics of the research project, the applicant organization, and the technological endowment of the location, as well as controls for the inventors, the technology and the country of the invention. Table 2 provides the descriptive statistics of the variables. Table 3 shows the correlation matrix. Appendix 1 lists the technological classes used in the regressions.

Tables 2 and 3 About Here

3.2.2 Core covariates

The cost of the research project. The inventors were asked to report the number of men-months required by the research leading to the patent (MMONTHS) We use a variable that ranges between 1 and 8 that indicates whether the research project required: less than 1 man-month; from 1 to 3 man-months; 4-6 man-months; 7-12 man-months; 13-24 man-months; 25-48 man-months; 49-72 man-months; more than 72 man-months.

The knowledge base of the firm. By using data from the European Patent Office we compute the number of patents granted to the consolidated applicant firm in 1993-1997 (FIRM_PATS). In the robustness checks, in place of FIRM_PATS we include two other variables: one is a measure of diversification of the firm knowledge portfolio (i.e. NR_TECHCLASSES, computed as number of ISI-INPI-OST technological classes of the firm's patents); the other one proxies for the experience in producing inventions, and it is given by number of years since the first patent invented by the inventor in our sample (YEAR_FIRSTPAT). Both variables are calculated by using the PatVal-EU data.

The regional knowledge endowment. We complement the PatVal-EU database with information on the technological endowment of the regions where the research project was performed, and we incorporate this information in different specifications of the econometric model. We start with a measure of the "general" technological endowment of the regions. This is REG_PATS and it is given by the total number of patents applied in all sectors (average in 1994-1996) and invented in the NUTS3 region in which the R&D project that led to a patent was carried out (source: Regio Eurostat). REG_PATS excludes the number of patents granted to the applicant firm of the patent.

Additionally, we develop two measures of the knowledge endowment in the *specific* technology of the patent-unit of observation (see, for example, Jaffe 1989, Furman et al. 2007). From the Regio-Eurostat database we collect the 1994-1996 number of regional patents applied at the EPO in each specific ISI-INPI-OST class of the patent (REG_PATS_TECH). Then, we compute the ratio between the patents invented in the region *i* in the specific technology *t* and the total number of patents invented in that technology in all regions (REG_PATS_SHARE).[2] We use REG_PATS_TECH and REG_PATS_SHARE as different specifications of the regional knowledge endowment.

Interaction term: cost of the invention and regional knowledge endowment. For each of the regional technological variables included in the regression, we also add its interaction with the cost of the invention (respectively, REG_PATS*MMONTHS, TECH_PATS*MMONTHS, REG_PATS_SHARE*MMONTHS).

3.2.3 Control Covariates

Applicant controls. About 92% of the patents in our database are granted to business companies. In the remaining 8% of the cases they are granted to individual inventors and public research organizations including universities. We use three dummy variables for the type of applicant organization: PRI_APPLIC takes the value 1 if the applicant organization is a university or a public research institution, INDIVIDUAL_APPLIC takes the value 1 if the applicant is an individual inventor. The baseline is FIRM_APPLIC.

For patents granted to private companies we add information on the size and R&D intensity of the firms as the average in the period of the patent priority dates. We collect these data from Compustat (1998) and Amadeus (2005). Both variables are at the level of the parent company. The number of employees (EMPLOYEES) is a proxy for the size of the firms, while the ratio between R&D expenditure and sales

[2] These dummy variables are calculated at the NUTS2 regional level because NUTS3 level data by micro technological classes are not available from Regio-Eurostat.

(R&DINT) measures their R&D intensity. For missing observations we include two dummy variables: D_MISS_EMPLOYEES and D_MISS_R&D.[3] By controlling for both firm size and R&D intensity, we separate the effect of the scale of the organization from the resources devoted to innovation.

Inventor controls. The PatVal-EU survey provides information on the individual characteristics of the inventor. We control for the level of education of the inventors by employing a dummy variable that is 0 if the inventor has a degree up to the High School; 1 if he has a University BSc, Master or PhD degree (UNIPHD_DEGREE). We also included a variable for the gender the inventors(MALE, which is 1 if male; 0 if female) and for their age (AGE), which is calculated as the years between the date of birth and the date of the patent application**.**

Other regional controls. In order to estimate the net effect of the technological characteristics of the regions we include two exogenous regional controls at the NUTS3 level: population density (POPDENSITY, ratio between population in thousands of people living in the region - average 1994-1996 - and the area of measured in square kilometers) and the general economic conditions (GDPPC, i.e., regional per capita Gross Domestic Product in thousands of purchasing power parity corrected for inflation, average 1994-1996).

All regressions include dummies for the application year (1994 to 1998), country of the invention (DE, ES, IT, NL, UK) and the 30 ISI-INPI-OST technological classes.

3.3 Results

Table 4 shows the estimated results of our empirical model. The three specifications differ for the inclusion of different measure of the regional technological endowment: REG_PATS, TECH_PATS and REG_PATS_SHARE, and the corresponding interaction term. The covariates with a large range of variation (FIRM_PATS, REG_PATS, TECH_PATS, REG_PATS_SHARE, EMPLOYEES, AGE, GDPPC) are in logs as indicated in the Table. Though they are not shown in the Table, all specifications include dummies for missing values in EMPLOYEES and R&DINT, as well as dummies for the country of the inventor, year of application, and technological field of the patent.

[3] Data on EMPLOYEES are available for 77.78% of the patents; data on R&DINT for 41.92% of the patents.

Table 4 About Here

Our data support *Hypothesis 1*. The larger the investment to develop the invention, the higher the likelihood that the inventors rely on *External* rather than *Internal* spillovers in the inventive process. As expected, the coefficient of MMONTHS is positive and statistically significant at 1% level. This holds across all three specifications, though the coefficient is smaller when we control for the external technological endowment specific to the technology of the patent.

Results are also in line with *Hypothesis 2*. The larger the stock of knowledge accumulated inside the firm, the lower the importance of *External* spillovers compared to those *Internal* to the organization. The coefficient of FIRM_PATS is negative and statistically significant at 1% level. This inward looking behavior of companies with larger patent portfolios holds across the three specifications.

What about the effect of the external technological environment? *Hypothesis 3* says that the richer the knowledge endowment of the region where the R&D project is carried out, the higher the importance of *External* compared to *Internal* knowledge spillovers. Model 1 estimates the effect of REG_PATS, the general technological endowment. As expected, the effect of REG_PATS is positive and statistically significant at 1% level. The effect persists also when we employ the technology specific measure of the local knowledge endowment: REG_PATS_TECH in Model 2 and REG_PATS_SHARE in Model 3 are positive and statistically significant, confirming that, as expected, the richer the technological environment, also in the specific technological class of the patented invention, the higher the importance of *External* compared to *Internal* spillovers.

Finally, *Hypothesis 4* states the joint effect of the cost of the innovation project and the level of knowledge endowment of the region is negatively correlated with the importance of *External* vs. *Internal* spillovers is also confirmed. In all three models, the interaction term between the regional technological environment and the cost of the R&D project (respectively, REG_PATS*MMONTHS, TECH_PATS*MMONTHS and REG_PATS_SHARE*MMONTHS) is negative and statistically significant at 1% level. Thus, inventors rely less on *External* than *Internal* spillovers when they develop potentially valuable inventions in a regional endowments that is conducive to knowledge spillovers.

As far as the applicant control covariates are concerned, some of them are statistically significant. This is the case of the R&D intensity of the firms. The higher the R&D intensity of the firm, the lower the importance of *External* spillovers compared to those *Internal* to the organization. Also, public research institutes and individual inventors are more open to *External* spillovers. At the level of the individual inventors, only AGE is positively correlated with the importance of *External* spillovers: older inventors are more open (or have better opportunities) to interact with people external to the organization. The other inventor level variables do not show any correlation with our dependent variable. Finally, as far as the regional controls are concerned, the level of economic development of the region is negatively correlated with the importance of *External* interactions, which, again, might control for the level of local competition.

3.4 Robustness checks: The firm stock of knowledge and bivariate probabilities

In order to control for the robustness of our results we employ alternative measures of the knowledge base of the firm in place of FIRM_PATS. We first use a measure of diversification of the firm knowledge portfolio (NR_TECHCLASSES), and then a proxy for the experience in producing inventions (YEAR_FIRSTPAT). Table 5 shows the estimated results.

Table 5 About Here

Both measures are negatively correlated with our dependent variable, suggesting that, again, the larger the stock of knowledge of the firm, either measured in terms of the technological diversification of its patent portfolio or the experience in producing inventions, the more inventors rely on knowledge spillovers *Internal* to the organization compared to those accessed though *External* interactions. The same applies when the two variables are included together in the same equation (not shown here). They are both negatively correlated with our dependent variable. The estimated results of the other covariates do not change compared to those shown in Table 4.

We perform an additional robustness check by estimating the marginal effects of the covariates on the predicted probabilities of either *External* or *Internal* spillovers. In so doing, we change the estimation technique and so the specification of our dependent variable. We first construct two dichotomous variables: *Internal* that takes the value 1 if the inventors use *Internal* spillovers, irrespective of their importance, and 0 otherwise; and *External* that takes the value 1 if the inventors use *External* spillovers, irrespective of their importance, and 0 otherwise. These two variables are the dependent variables of a Bivariate Probit regression. We then estimate the marginal effects of the covariates on the predicted probabilities of the four combinations of outcomes (i.e., *External*=1&*Internal*=1, *External*=0&*Internal*=0 *External*=1&*Internal*=0 *External*=0&*Internal*=1) computed after the Bivariate Probit estimates. Table 6 reports the marginal effects on the predicted probabilities of *Internal*=1&*External*=0 (i.e. P10) and *Internal*=0&*External*=1 (i.e. P01).

Table 6 About Here

The results in Table 6 show the net effect of the covariates on the use of either *External* (right-end column) or *Internal* spillovers (left column). The results are perfectly consistent with those reported in Table 4 on the standardized difference. First of all, MMONTHS (i.e. the cost of the research project) is negatively correlated with the probability of the inventor to use *Internal* interactions, and it is positively correlated with the use of *External* interactions. Both marginal effects are statistically significant at 1% level. Also, the effect on *Internal* is larger than the effect on *External*. Since MMONTHS in Table 4 is also positively correlated with the standardized difference between *External* and *Internal*, this suggests that not only the cost of the project affects the reliance on either *Internal* or *External* spillovers when they are used exclusively in the inventive process, but also their relative importance when they are both used to develop an invention.

Consistently with the results in Table 4, the stock of knowledge accumulated by the firm (FIRM_PATS) positively affects the probability to rely on *Internal* spillovers, while being negatively correlated with *External* spillovers. The former effect is also larger than the latter, therefore contributing to the general effect on the standardized difference. The other two firm level covariates that we included in place of FIRM_PATS also behave as expected: the marginal effects of NR_TECHCLASSES and YEAR_FIRSTPAT are both positive and statistically significant on *Internal* and negative on *External*. Again, the marginal effect on *Internal* is larger than that on *External*.

As far as the regional knowledge endowment is concerned, all the three covariates (i.e. REG_PATS, REG_PATS_TECH and REG_PATS_SHARE) positively affect the likelihood of the inventors to rely on *External* spillovers. The marginal effects on *External* are also statistically significant, which might explain the positive effect on the standardized variable in Table 4, as the negative marginal effect on *Internal* is not statistically significant (it is only at 10% level for REG_PATS_SHARE).

Finally, in all three models, the interaction term between the regional technological environment and the cost of the project is negative on *External* and statistically significant. The positive effect on the use of *Internal* spillovers applies only when we consider the quality of the knowledge endowment specific to the technology of the patent.

4. Conclusions

This article employs unique data on a sample of 6,016 patents that we collected by means of a large scale survey (the PatVal_EU survey). We uncover the factors that determine the relative importance of *External* knowledge spillovers (i.e. spillovers from sources outside the inventor organization) compared to *Internal* spillovers (i.e. spillovers internal to the inventor organization) in the inventive process. Information are provided directly by the inventors of the patents, and then complemented with other data on firms and locations derived from standard datasets.

Our theory develops through three layers of analysis: the R&D project that led to the invention, the applicant organization, and the geographical region in which the inventive process takes place. This is the base to develop the four hypotheses on the factors that drive the relative importance of external vs. internal spillovers. Given the nature of our dependent variable, the empirical model estimates a two-limit tobit regression, followed by robustness checks. Empirical evidence supports our hypotheses.

Specifically, absorptive capacity and R&D complementarity lead to use External spillovers more than Internal spillovers. However, a larger stock of knowledge accumulated by the firm produces the opposite effect, favoring the inward looking behavior of the organizations. These results provide an additional insight on the determinants of the trade-off between the inward looking and outward looking absorptive capacity of the firms.

Also, our results show that firms located in regions with a richer knowledge environment do also rely on External more than on Internal knowledge spillovers. We provide arguments for these results based on a novel perspective that treats the geography literature (Saxenian, 1994; Feldman, 1999) and the sociological tradition (Fleming, 2007) as complements.

Finally, our findings corroborate the idea that a trade-off exists between the benefits to R&D produced by local knowledge spillovers and the costs of involuntary leakages of valuable knowledge to local competitors. We find that, as the costs of an invention project increases, the potential threats of imitation that a firm perceives from its close geographical environment hamper its openness and reduce the use of external spillovers.

5. References

Agrawal, A., I. Cockburn, J. McHale. 2006. Gone But Not Forgotten: Labor Flows, Knowledge Spillovers and Enduring Social Capital. *Journal of Economic Geography* 6 (5), 571-591.

Alcacer, J., W. Chung. 2007. Location Strategies and Knowledge Spillovers. *Management Science* 53(5), 760-776.

Alcacer, J., M. Gittelman. 2006. Patent Citations as a Measure of Knowledge Flows: The Influence of Examiner Citations. *Review of Economics and Statistics* 88(4), 774-779.

Alcacer, J.2006, Location Choices across the Value Chain: How Activity and Capability Influence Collocation, Management Science 52(10) pp. 1457-1471.

Allen TJ, Katz R. 1982. Investigating the not invented here (NIH) Syndrome. A Look at the Performance, Tenure and Communication Patterns of 50 R&D Project groups. R&D Management 12(1): 7-19

Almeida, P., B. Kogut. 1999. Localization of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science* 45(7), 905-17.

Arora A. e Gambardella A., 1990, Complementarity and External Linkages: The Strategies of the Large Firms in Biotechnology, *Journal of Industrial Economics*, 38 (4), pp.361-379.

Arora A, Merges RP. 2004. Specialized supply firms, property rights and firm boundaries. Industrial and Corporate Change. 13(3): 451-475

Audretsch, D.B., M.P. Feldman. 1996. R&D Spillovers and the Geography of Innovation and Production. *American Economic Review* 86(3), 630-640.

Ahuja G., 2000 . Collaboration Networks, Structural Holes and Innovation: A Longitudinal Study. *Administrative Science Quarterly*, 45:425-455

Bell G.G., Zaber, A. 2007. Geography, Networks and Knowledge Flows. *Organization Science*, 18(6) pp. 955-972

Borgatti SP, Cross R. 2003. A relational view of information seeking and learning in social networks. Management Science, 49(4): 432-445

Breschi, S., F. Lissoni. 2001. Localized knowledge spillovers vs. innovative milieux: Knowledge 'tacitness' reconsidered. *Regional Science* 80(3), 255-73.

Burgelman RA. 1994. Fading Memories. A Process Theory of Strategic Business Exit in Dynamic Enviroment, Administrative Science Quarterly, 39(1): 24-56

Cassiman B. and Veugelers R., 2006, In Search of Complementarity in Innovation Strategy: Internal R&D and External Knowledge Acquisition**,** Management Science, 52, pp 68 – 82

Chesbrough, H., 2003, Open Innovation: the new imperative for creating and profiting from technology, Harvard Business School Press

Cohen W., and D. Levinthal. 1990, Absorptive capacity: a new perspective on learning and innovation, Administrative Science Quarterly, 35(1) pp 128-152.

Danneels E. 2002. The dynamics of product innovation and firm competences. Strategic Management Journal, 23(12): 1095-1121

Dibiaggio, L. 2007. Design complexity, vertical disintegration and knowledge organization in the semiconductor industry. Industrial and Corporate Change, 16(2): 239-267

Dosi, G. 1988. Sources, Procedures, and Microeconomic Effects of Innovation, Journal of Economic Literature, 26(3): 1120-1171

Doring, T., J. Schnellenbach. 2006. What do we know about geographical knowledge spillovers and regional growth?: A survey of the literature. *Regional Studies* 40(3), 375-395.

Feldman, M. 1999. The new economics of innovation, spillovers and agglomeration: A review of empirical studies. *Economics of Innovation and New Technology* 8(1-2), 5-25.

Fleming, L., C. King, A. Juda. 2007. Small Worlds and Regional Innovation. Organization Science 18(6), 938-954

Fleming L., Sorenson O., 2004. Science and the diffusion of knowledge. Research Policy, 33 (10): 1615-1634

Gambardella, A., Harhoff, D., Verspagen, B., 2008, The Value of European Patents, European *Management Review*, forthcoming

Giuri P., Mariani M., Brusoni S., Crespi G., Francoz D., Gambardella A., Garcia-Fontes W., Geuna A., Gonzales R., Harhoff D., Hoisl K., Lebas C., Luzzi A., Magazzini L., Nesta L., Nomaler O., Palomeras N., Patel P., Romanelli M., Verspagen B., Inventors and Invention Processes in Europe. Results from the PatVal-EU survey, 2007, *Research Policy*, 36(8), pp 1107-1127

Granstrand O, Patel P, Pavitt K. 1997. Multi-technology corporations: Why they have "distributed" rather than "distinctive core" competencies. California Management Review. 39(4): 8-18

Griffin and J.R. Hauser, 1992, Patterns of communications among marketing, engineering and manufacturing. A comparison between two new product teams, *Management Science* **38** (3), pp. 360–373.

Harhoff, D., K. Hoisl, C. Webb. 2006. European patent citations. How to count and how to interpret them, Unpublished Manuscript. *Ludwig-Maximilian University of Munich*.

Hauser JR. 1998. Research, development, and engineering metrics. Management Science. 44(12) 1670-1689

Helfat C.E and Quinn J. B., 2006, Open innovation: The new imperative for creating and profiting from technology. *Academy of Management Perspectives*, 20, 86-88

Henderson R., Cockburn I 1994. Measuring Competence. Exploring firm effects in Pharmaceutical Research. Strategic Management Journal, 15: 63-84

Ilinitch, A., R. D'Aveni, and A. Lewin. 1996 New organizational forms and strategies for managing in hypercompetitive environments. Organization Science 7/3: 211–220

Jaffe, A., M. Trajtenberg, R. Henderson. 1993. Geographic knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics* 108(3), 577-98.

Klepper S. 2007. Disagreements, spinoffs, and the evolution of Detroit as the capital of the US automobile industry . Management Science, 53(4): 616-631

Krishnan V, Ulrich KT. 2001. Product development decisions: A review of the literature. Management Science, 47(1): 1-21

Lampel, J., J. Shamsie 2003. Capabilities in motion: New organizational forms and the reshaping of the Hollywood movie industry. Journal of Management Studies 40 2189-2210.

Laursen, K, and Salter, A. 2006. Open for innovation: The role of openness in explaining innovation performance among UK manufacturing firms, Strategic Management Journal, 27, pp.131 – 150

Lee, F. 1997. When the going gets tough, do the tough ask for help? Help seeking and power motivation in organizations. Organizational Behavior and Human Decision Processes. 72 (3): 336-363

Levinthal D.A., March J.G. 1993. The Myopia of Learning, Strategic Management Journal, 14: 95-112

Maddala G.S., 1983, Limited-dependent and qualitative variables in econometrics, Cambridge University Press

McEvily, S., Chakravarthy, B., 2002, The persistence of knowledge-based advantage: an empirical test for product performance and technological knowledge, Strategic Management Journal, 23 pp.285-305

Menon T, Pfeffer J. 2003. Valuing internal vs. external knowledge: Explaining the preference for outsiders. Management Science. 49(4): 497-513

Nelson, R., Winter, S., 1982. An Evolutionary Theory of Economic Change. Harvard University Press, Cambridge, MA.

Pavitt, K. 1987. On the nature of technology. *University of Sussex*, SPRU, Brighton

Romanelli E, Khessina OM. 2005. Regional industrial identity: Cluster configurations and economic development . Organization Science. 16(4): 344-358

Rosenberg, N. 1976. Technological Expectations. Economic Journal, 86(343): 523-535

Sakakibara M., 2001, "Cooperative research and development: who participates and in which industries do projects take place?", Research Policy 30(7), 993-1018

Saxenian, A.L. 1994, *Regional Advantage, Culture and Competition in Silicon Valley and Route 128*, Cambridge (MA) and London.

Schmalensee, R. 2000. Antitrust issues in Schumpeterian industries. American Economic Review 90 (2): 192–197.

Sorenson, O. 2000. Letting the market work for you: An evolutionary perspective on product strategy. Strategic Management Journal 21(2): 277–292.

Sorenson, O. and Singh, J., 2007, Science, social networks and spillovers, *Industry & Innovation*, 14, pp. 219-238

Storper M., 1993, Regional Worlds of Production: Learning and Innovation in the Technology Districts of France, Italy and the USA, Regional Studies, 27 (5), pp. 433-455

Tallman S., M. Jenkins, N. Henry, S. Pinch 2004. Knowledge, cluster and competitive advantage. Academy of Management Review 29 258-271.

Teece, D. 1986. Profiting from technological innovation: Implications for integration, licensing and public policy. Research Policy 15(1) 285–305

Tripsas M, Gavetti G. 2000. Capabilities, cognition, and inertia: Evidence from digital imaging. Strategic Management Journal, 21(10-11): 1147-1161

Thompson, P. 2006. Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-added Citations. *Review of Economics and Statistics* 88(2), 383-388.

Uzzi B., 1997. Social Structure and Competition in Interfirm Network: The Paradox of Embedness. *Administrative Science Quarterly*, 42:35-67

Von Hippel, E. 1986. Lead users: A source of novel product concepts. Management Science. 32(7): 791–805

White S, Lui SSY. 2005. Distinguishing costs of cooperation and control in alliances. Strategic Management Journal. 26(10): 913-932

Appendix 1. List of ISI-INPI-OST technological classes used in the paper and descriptive statistics.

	Mean	Std. Dev.
Electrical devices, engineering, energy	0.078	0.268
Audio-visual technology	0.020	0.139
Telecommunications	0.033	0.180
Information technology	0.022	0.146
Semiconductors	0.010	0.100
Optics	0.020	0.142
Analysis, measurement, control technology	0.056	0.231
Medical technology	0.022	0.147
Organic fine chemistry	0.065	0.247
Macromolecular chemistry, polymers	0.063	0.243
Pharmaceuticals, cosmetics	0.016	0.125
Biotechnology	0.007	0.086
Materials, metallurgy	0.032	0.176
Agriculture, food chemistry	0.013	0.113
Chemical&petrol, basic materials chem.	0.039	0.193
Chemical engineering	0.028	0.166
Surface technology, coating	0.016	0.124
Materials processing, textiles, paper	0.056	0.230
Thermal processes and apparatus	0.021	0.143
Environmental technology	0.017	0.131
Machine tools	0.036	0.186
Engines, pumps, turbines	0.031	0.173
Mechanical Elements	0.043	0.203
Handling, printing	0.076	0.266
Agricultural&food proc-machin-apparatus	0.019	0.136
Transport	0.070	0.255
Nuclear engineering	0.004	0.060
Space technology weapons	0.004	0.066
Consumer goods and equipment	0.046	0.209
Civil engineering, building, mining	0.036	0.186

Figures

Figure 1: Theoretical background.
Factors affecting the use of Internal vs. External knowledge use in the invention processes

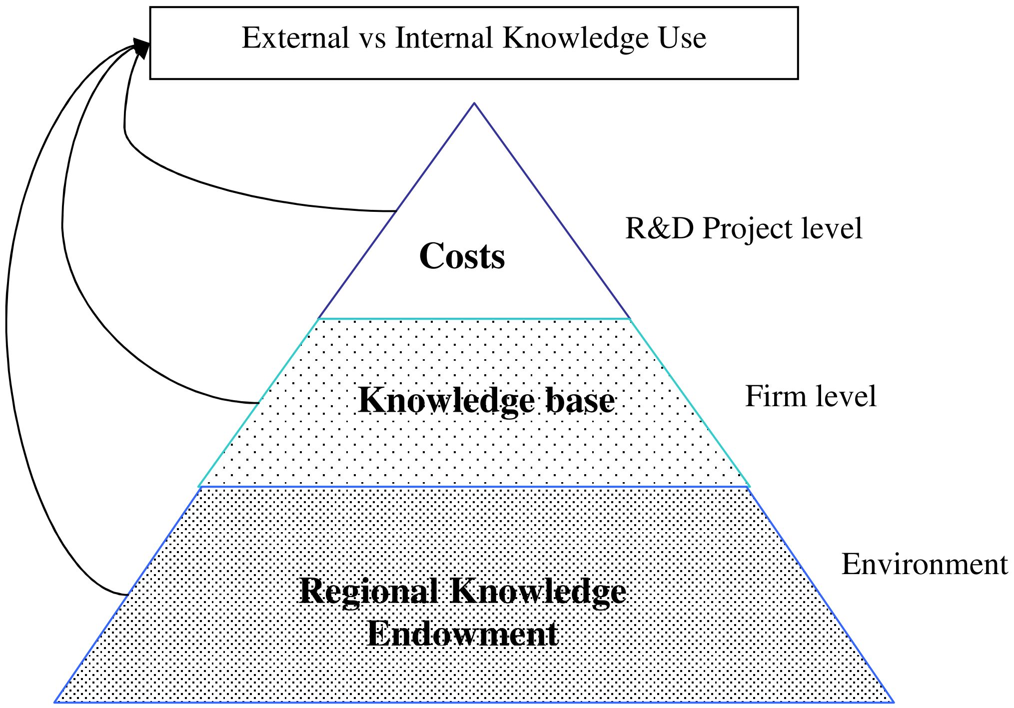
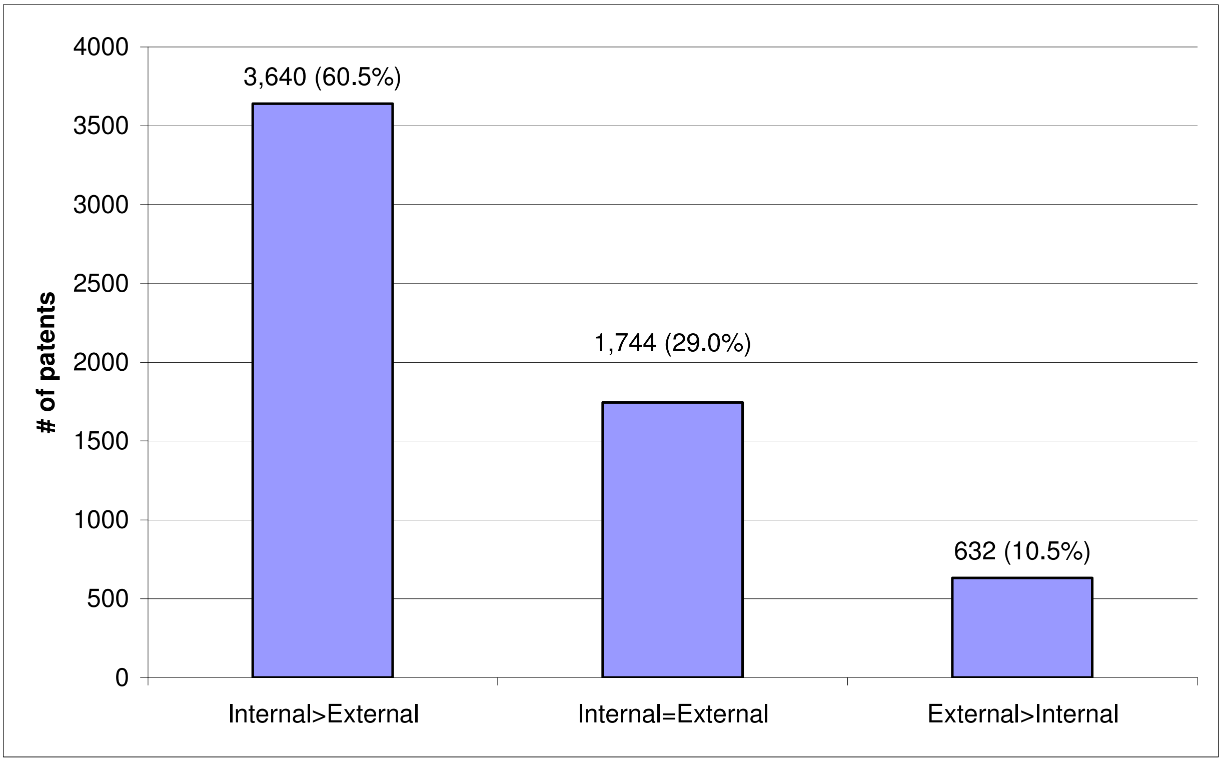


Figure 2. Relative importance of *Internal* and *External* interactions. Number and share of patents where Internal>External (StdExtInt<0), Internal=External (StdExtInt=0) and External>Internal (StdExtInt>0).



Source: PatVal-EU dataset. Note: the average score of Internal is 3.15 (s.d. 1.85); the average score of External is 1.42 (s.d. 1.83)

Tables

Table 1. Unconditional probabilities: share of patents invented with *Close* and/or *Distant* interactions (0 = external interactions not used; 1: external interactions used, regardless of their importance).

	External = 0	*External* = 1	Total
Internal = 0	14.4%	3.6%	18.0%
Internal = 1	39.2%	42.8%	82.0%
Total	53.6%	46.4%	100%

Source: PatVal-EU dataset
N = 6,016

Table 2. Descriptive statistics

	Mean	St. Dev.	Min	Max
	Dependent variables			
StdExtInt	-0.273	0.361	-0.714	0.714
External	0.464	0.499	0	1
Internal	0.819	0.385	0	1
	Explanatory variables			
MMONTHS	3.407	1.776	1	8
FIRM_PATS	232	410	1	1729
REG_PATS	118.340	131.138	0.497	542.880
REG_PATS_TECH	59.262	70.596	0	424.918
REG_PATS_SHARE	0.017	0.014	0	0.063
NR_TECHCLASSES	7.124	6.825	1	23
YEAR_FIRSTPAT	8.263	5.330	1	21
	Applicant controls			
EMPLOYEES[a]	90200.39	117170.9	1	723328.60
D_MISS_EMPLOYEES	0.199	0.399	0	1
R&DINT[b]	0.054	0.033	0	0.412
D_MISS_R&D	0.552	0.497	0	1
PRI_APPLIC	0.027	0.161	0	1
INDIVIDUAL_APPLIC	0.042	0.200	0	1
	Inventor controls			
AGE	45.045	9.670	20	81
MALE	0.978	0.150	0	1
UNIPHD_DEGREE	0.811	0.392	0	1
	Regional controls			
GDPPC	23358.56	9255.649	8677.9	76910.8
POPDENSITY	1.066	1.070	0.126	8.787
Other Controls				
DE	0.472	0.499	0	1
ES	0.028	0.168	0	1
IT	0.161	0.367	0	1
NL	0.150	0.357	0	1
UK	0.188	0.391	0	1
AppYear1994	0.285	0.451	0	1
AppYear1995	0.269	0.443	0	1
AppYear1996	0.231	0.421	0	1
AppYear1997	0.147	0.354	0	1
AppYear1998	0.041	0.197	0	1

Note: N = 6,016 [a] N=4820 [b]N= 2697

Table 3. Correlation matrix.

	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	MMONTHS	1.000															
2	FIRM_PATS	-0.024	1.000														
3	NR_TECHCLASSES	0.022	0.920*	1.000													
4	YEAR_FIRSTPAT	-0.058*	0.237*	0.209*	1.000												
5	REG_PATS	-0.027*	0.240*	0.212*	0.082*	1.000											
6	REG_PATS_TECH	-0.147*	0.368*	0.302*	0.226*	0.497*	1.000										
7	REG_PATS_SHARE	-0.138*	0.234*	0.228*	0.195*	0.483*	0.746*	1.000									
8	EMPLOYEES	-0.020	0.769*	0.734*	0.186*	0.143*	0.232*	0.160*	1.000								
9	R&DINT	0.031*	0.603*	0.587*	0.151*	0.149*	0.205*	0.130*	0.565*	1.000							
10	PRI_APPLIC	0.129*	-0.052*	-0.017	-0.073*	-0.055*	-0.100*	-0.084*	-0.268*	-0.116*	1.000						
11	INDIVIDUAL_APPLIC	0.030*	-0.297*	-0.255*	-0.080*	-0.051*	-0.068*	-0.027*	-0.358*	-0.146*	-0.034*	1.000					
12	AGE	-0.039*	-0.112*	-0.079*	0.336*	-0.028*	-0.005	0.047*	-0.085*	-0.065*	-0.030*	0.081*	1.000				
13	MALE	-0.100*	-0.013	-0.019	0.060*	-0.004	0.027*	0.037*	0.006	-0.015	-0.050*	0.010	0.122*	1.000			
14	UNIPHD_DEGREE	0.085*	0.235*	0.227*	0.090*	0.057*	0.063*	0.073*	0.190*	0.165*	0.072*	-0.097*	-0.119*	-0.043*	1.000		
15	GDPPC	-0.059*	0.300*	0.284*	0.149*	0.361*	0.480*	0.490*	0.217*	0.190*	-0.062*	-0.049*	0.008	0.033*	0.090*	1.000	
16	POPDENSITY	-0.006	0.235*	0.241*	0.098*	0.259*	0.297*	0.352*	0.189*	0.159*	0.026*	-0.053*	0.006	0.014	-0.134*	0.575*	1.000

Note: * significant at 5%

Table 4. Two-limit tobit estimation. Dependent variable: StdExtInt. Models 1-3

	Model 1: REG_PATS	Model 2: REG_PATS_TECH	Model 1: REG_PATS_SHARE
Explanatory variables			
MMONTHS	0.04*** (0.01)	0.02*** (0.01)	0.02*** (0.00)
Log(FIRM_PATS)	-0.02*** (0.00)	-0.02*** (0.00)	-0.02*** (0.00)
Log(REG_PATS)	0.10*** (0.03)	-	-
Log(REG_PATS)*MMONTHS	-0.10*** (0.03)	-	-
Log(REG_PATS_TECH)	-	0.05*** (0.02)	-
Log(REG_PATS_TECH)*MMONTHS	-	-0.05*** (0.01)	-
REG_PATS_SHARE	-	-	1.64* (0.87)
REG_PATS_SHARE*MMONTHS	-	-	- 0.59*** (0.22)
Applicant controls			
Log(EMPLOYEES)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
D_MISS_EMPLOYEES	0.03 (0.03)	0.03 (0.03)	0.03 (0.03)
R&DINT	-0.56** (0.26)	-0.58** (0.26)	-0.55** (0.26)
D_MISS_R&D	0.01 (0.02)	0.00 (0.02)	0.01 (0.02)
PRI_APPLIC	0.10*** (0.04)	0.10*** (0.04)	0.10*** (0.04)
INDIVIDUAL_APPLIC	0.27*** (0.03)	0.27*** (0.03)	0.27*** (0.03)
Inventor controls			
Log(AGE)	0.08*** (0.03)	0.08*** (0.03)	0.08*** (0.03)
MALE	-0.01 (0.04)	-0.01 (0.04)	-0.01 (0.04)
UNIPHD_DEGREE	-0.02 (0.01)	-0.02 (0.01)	-0.02 (0.01)
Regional controls			
Log(GDPPC)	-0.06*** (0.02)	-0.06*** (0.02)	-0.06*** (0.02)
POPDENSITY	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)
cons	-0.06 (0.24)	-0.04 (0.24)	-0.09 (0.24)
N	6016	6016	6016
Ll	-3663.46	-3661.38	-3665.62
Chi squared	497.91	502.07	493.58

Note: Standard errors are in parentheses. Coefficient significant at * p<0.10, **p<0.05, ***p<0.01. All regressions include dummies for *Inventor country, Year of application* and *Technological field* (30 ISI-INPI-OST classes).

Table 5. Two-limit tobit estimation. Robustness checks. Dependent variable: StdExtInt. Models 4-6 with NR_TECHCLASSES; Models 7-9 with YEAR_FIRSTPAT.

	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Explanatory variables						
MMONTHS	0.04*** (0.01)	0.02*** (0.01)	0.02*** (0.00)	0.04*** (0.01)	0.02*** (0.01)	0.02*** (0.00)
Log(NR_TECHCLASSES)	-0.03*** (0.01)	-0.02*** (0.01)	-0.03*** (0.01)	-	-	-
Log(YEAR_FIRSTPAT)	-	-	-	-0.03*** (0.01)	-0.03*** (0.01)	-0.03*** (0.01)
Log(REG_PATS)	0.10*** (0.03)	-	-	0.10*** (0.03)	-	-
Log(REG_PATS)*MMONTHS	-0.10*** (0.03)	-	-	-0.10*** (0.03)	-	-
Log(REG_PATS_TECH)	-	0.04** (0.02)	-	-	0.04** (0.02)	-
Log(REG_PATS_TECH)*MMONTHS	-	-0.05*** (0.01)	-	-	-0.05*** (0.01)	-
REG_PATS_SHARE	-	-	1.59* (0.86)	-	-	1.60* (0.86)
REG_PATS_SHARE*MMONTHS	-	-	- 0.59*** (0.22)	-	-	- 0.60*** (0.22)
Applicant controls						
Log(EMPLOYEES)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.01 (0.00)	0.00 (0.00)	0.01 (0.00)
D_MISS_EMPLOYEES	0.02 (0.03)	0.03 (0.03)	0.03 (0.03)	-0.02 (0.03)	-0.02 (0.03)	-0.02 (0.03)
R&DINT	-0.62** (0.26)	-0.64*** (0.26)	-0.61** (0.26)	-0.72*** (0.26)	-0.74*** (0.26)	-0.71*** (0.26)
D_MISS_R&D	0.01 (0.02)	0.00 (0.02)	0.01 (0.02)	0.01 (0.02)	0.00 (0.02)	0.01 (0.02)
PRI_APPLIC	0.10*** (0.04)	0.09*** (0.04)	0.10*** (0.04)	0.08** (0.04)	0.07** (0.04)	0.08** (0.04)
INDIVIDUAL_APPLIC	0.28*** (0.03)	0.28*** (0.03)	0.28*** (0.03)	0.28*** (0.03)	0.28*** (0.03)	0.28*** (0.03)
Inventor controls						
Log(AGE)	0.08*** (0.03)	0.08*** (0.03)	0.08*** (0.03)	0.12*** (0.03)	0.12*** (0.03)	0.12*** (0.03)
MALE	-0.01 (0.04)	-0.01 (0.04)	-0.01 (0.04)	-0.01 (0.04)	-0.01 (0.04)	-0.01 (0.04)
UNIPHD_DEGREE	-0.02 (0.01)	-0.02 (0.01)	-0.02 (0.01)	-0.02 (0.01)	-0.02 (0.01)	-0.02 (0.01)
Regional controls						
Log(GDPPC)	-0.06*** (0.02)	-0.06*** (0.02)	-0.06*** (0.02)	-0.06*** (0.02)	-0.07*** (0.02)	-0.07*** (0.02)
POPDENSITY	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)
Cons	-0.03 (0.24)	-0.03 (0.24)	-0.07 (0.24)	-0.07 (0.24)	-0.05 (0.24)	-0.10 (0.24)
N	6016	6016	6016	6016	6016	6016
Ll	-3668.14	-3665.41	-3670.09	-3665.32	-3662.74	-3667.70
Chi squared	488.55	494.01	484.64	494.19	499.36	489.43

Note: Standard errors are in parentheses. Coefficient significant at * p<0.10, **p<0.05, ***p<0.01. All regressions include dummies for *Inventor country, Year of application* and *Technological field* (30 ISI-INPI-OST classes).

Table 6. Bivariate probit estimation. Marginal effects on the bivariate probabilities of *Internal* (P10) and *External* (P01)

	pr (Internal=1, External=0)	**pr (Internal=0, External=1)**
Explanatory variables		
MMONTHS	-0.038*** (0.012)	0.009*** (0.003)
Log(FIRM_PATS)	0.014*** (0.006)	-0.003** (0.001)
Model 1		
Log(REG_PATS)	-0.034 (0.035)	0.028*** (0.008)
Log(REG_PATS)*MMONTHS	0.039 (0.035)	-0.029*** (0.008)
Model 2		
Log(REG_PATS_TECH)	-0.029 (0.020)	0.012*** (0.004)
Log(REG_PATS_TECH)*MMONTHS	0.030* (0.016)	-0.012*** (0.003)
Model 3		
REG_PATS_SHARE	-1.618* (0.980)	0.538* (0.209)
REG_PATS_SHARE*MMONTHS	0.588** (0.259)	-0.169*** (0.058)
Additional explanatory variables		
Log(NR_TECHCLASSES)	0.020** (0.010)	-0.004* (0.002)
Log(YEAR_FIRSTPAT)	0.025** (0.011)	-0.010*** (0.002)
Applicant controls		
Log(EMPLOYEES)	-0.007 (0.005)	0.001 (0.001)
D_MISS_EMPLOYEES	-0.058 (0.036)	0.010 (0.009)
R&DINT	0.517* (0.306)	-0.073 (0.070)
D_MISS_R&D	-0.010 (0.025)	0.005 (0.005)
PRI_APPLIC	-0.105*** (0.039)	0.037** (0.016)
INDIVIDUAL_APPLIC	-0.202*** (0.027)	0.110*** (0.020)
Inventor controls		
AGE	0.015 (0.029)	0.028*** (0.006)
MALE	0.004 (0.043)	-0.005 (0.011)
UNIPHD_DEGREE	-0.011 (0.016)	-0.009** (0.004)
Regional controls		
Log(GDPPC)	0.048* (0.026)	-0.009 (0.006)
POPDENSITY	-0.006 (0.007)	0.000 (0.002)

Note: Robust standard errors are in parentheses adjusted for clusters by firms' identifier. Coefficient significant at * p<0.10, **p<0.05, ***p<0.01. All regressions include dummies for *Missing value for EMPLOYEES*, *Missing values for R&DINT*, *Inventor country, Year of application* and *Technological field* (30 ISI-INPI-OST classes).