WHEN DISTANCE DISAPPEARS: INVENTORS, EDUCATION, AND THE LOCUS OF KNOWLEDGE SPILLOVERS

Paola Giuri and Myriam Mariani*

Abstract—This paper discusses the role of education in shaping the geographical breadth of knowledge spillovers. Data pertaining to 6,051 European inventions reveal that inventors with a high level of education, such as a university or doctoral degree, rely more on external spillovers regardless of the geographical location of their sources. Controlling for this effect, they also access geographically wider knowledge spillovers. This result holds after controlling for alternative explanations, such as the inventors' network and the site where the research is performed. By contributing to individual openness, education thus provides a means to break through geographical barriers to attain knowledge diffusion.

To ask a question, one must know enough to know what is not known.

—Miyake and Norman (1979)

I. Introduction

THE economic importance of knowledge spillovers (Grossman & Helpman, 1991; Romer, 1990), defined as nonmarket-mediated knowledge flows that allow one party's investments in knowledge generation to benefit external parties by facilitating innovation (Jaffe, Trajtenberg, & Fogarty, 2000), has encouraged researchers to document their existence and study their boundaries. A lively debate involves whether the transmission and absorption of knowledge spillovers are sensitive to geographical distance (Thompson, 2006). Although empirical evidence suggests that knowledge spillovers are geographically concentrated between and within countries (Jaffe, Trajtenberg, and Henderson, 1993; Verspagen 1997), more recent work, by Thompson and Fox-Kean (2005), shows modest or no location effects.

This paper considers factors that might make knowledge spillovers sensitive to geographical distance. After controlling for alternative explanations, we study the extent to which the geographical breadth of knowledge spillovers depends on individual characteristics, particularly education. We posit and test whether education contributes to the capacity to scout for and absorb external spillovers, regardless of geographic distance. We then investigate whether, when we control for this capacity, education also facilitates the use of geographically distant knowledge spillovers.

Various researchers have studied the mechanisms by which knowledge spillovers are transmitted, which may affect their geographical breadth. The notion that spillovers appear "in the air" contrasts with suggestions that other mechanisms, such as personal networks, shape the geography of knowledge spillovers or overcome geographic constraints to their diffusion (Agrawal Kapur, & McHale, 2008; Fleming, King, & Juda, 2007; Gittelman 2007; Singh 2005; Summel, Stephan, & Adams 2009). In these studies, though, the role of individual features remains limited to network characteristics. Little work addresses the impact of individuals' intrinsic features, which may reflect differential abilities to use near versus far knowledge spillovers or exogenous factors that drive the formation of networks with varied geographical coverage (Audretsch & Stephan, 1996).

We contribute to this line of research by employing information provided by the inventors of 6,051 European patents (from the PatVal-EU survey), including data on their age, gender, and level of education. As an indicator of knowledge spillovers, we include the inventor's use of interactions, such as meetings, discussions, and circulation of ideas, to pursue research that leads to a patented invention. Although this measure relies on self-assessments, it mimics Marshallian (1890) non–market-mediated knowledge spillovers and thus avoids the problem of using indirect measures such as patent citations.[1] Empirical results from bivariate probit models confirm our prediction: inventors with a high level of education rely more on external knowledge spillovers, regardless of geography, as well as on geographically distant spillovers, even after controlling for alternative explanations such as the use of scientific and professional communities, the nature of the research performed, and the local milieu of the inventive process. This finding suggests that beyond the personal networks it creates, education establishes better capabilities to scout for and absorb knowledge. Education also makes these processes less context specific and less dependent on face-to-face interactions. That is, our work does not deny the role of location with regard to the diffusion of knowledge spillovers, but it adds evidence about the role of individual heterogeneity in determining access to and exploitation of spillovers with varying geographical breadth.

In the next section, we discuss the role of education in shaping the geographical breadth of knowledge flows and derive our expectations. In section III, we present our measure of knowledge spillovers. Section IV outlines the covariates for the empirical analysis and the identification strategy. After discussing the results in section V, we conclude in section VI.

Received for publication April 28, 2009. Revision accepted for publication August 4, 2011.

* Giuri: Bologna University; Mariani: Bocconi University.

We gratefully acknowledge financial support from the European Commission (contract HPV2-CT-2001-00013). We also thank Maryann Feldman, Alfonso Gambardella, Peter Maskell, and Paula Stephan for their helpful and encouraging comments, and we acknowledge the contributions of all participants in conferences, seminars, and workshops where the paper has been presented. The usual disclaimers apply.

[1] Many patent citations are inserted by the patent examiners rather than by the inventors themselves (41% for U.S. and 93% for European patents; Harhoff, Hoisl, & Webb, 2008), which makes citations noisy indicators of knowledge spillovers (Jaffe, Trajtenberg, & Fogarty, 2000; Alcacer & Gittelman, 2006). Using patent citations and survey data from the European Community Survey, Duguet and MacGarvie (2005) find that the legitimacy of patent citations as a measure of knowledge flows varies across geographical regions and mechanisms of technology diffusion. Furthermore, citations capture only knowledge flows that result in patentable technologies.

© 2013 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/449/1917213/rest_a_00259.pdf by guest on 28 September 2021

II. Knowledge Spillovers and Heterogeneity across Inventors

A traditional argument regarding the geographic localization of knowledge spillovers suggests that physical proximity facilitates knowledge interactions and makes it easier for individuals to access knowledge produced by others (Doring & Schnellenbach, 2006; Feldman, 1999; Saxenian, 1994). Yet several scholars explain that geographical reach is shaped by specific mechanisms, such as formal collaborations and mobility of human capital (Zucker, Darby, & Armstrong, 1998; Almeida & Kogut, 1999).

Being a member of a scientific or professional community also provides opportunities for interactions. Allen (1977) shows that researchers use their networks of friends and colleagues as sources of knowledge. Members of scientific communities also tend to adopt a logic of knowledge sharing and circulation (Bercovitz & Feldman, 2008; Dasgupta & David, 1994; Gittelman, 2007; Knorr Cetina, 1999; Merton, 1973), with common norms, codes, and channels that cut across regions and countries such as exchange of papers. Professional networks and prior working relationships facilitate knowledge exchanges (Breschi & Lissoni, 2009; Sorenson & Singh, 2007) because they can produce ties that last even if people change locations (Agrawal, Cockburn, & McHale, 2006; Fleming et al., 2007; Kerr, 2008; Singh, 2005).

Although individual researchers sit at the core of these mechanisms, their role in existing studies remains limited to their networking features. An exception is Audretsch and Stephan's (1996) finding that the geographical breadth of links between U.S. biotechnology firms and scientists depends on the characteristics of the scientists and their role in the companies.[2] To extend this line of research, we posit that by contributing to the individual capacity to scout for and absorb knowledge, education facilitates the use of knowledge spillovers and geographically unbinds them, which enlarges the potential pool of knowledge sources that a scientist can use to develop new ideas, even after controlling for networking and other individual features.

Suppose an inventor faces a problem that he or she, the team, or other people in the organization cannot solve. The inventor can look outside the organization for hints, which may come from people located geographically nearby or from a larger pool of people located geographically farther away. Individual inventors are heterogeneous in their capacity to understand who has potentially useful knowledge and where they are located. They differ in their capacity "to recognize the value of new, external knowledge, assimilate it, and apply it to commercial ends" (Cohen & Levinthal 1990, p.128). We argue that education, especially higher education, shapes this heterogeneous absorptive capacity and thereby lowers the extent to which the absorption of knowledge spillovers is sensitive to geographical distance, for two reasons.

First, education enhances the capacity to scout for knowledge spillovers. In addition to providing students with substantive and technical skills, high-level education trains them to recognize the state of the art in their disciplines (for example, they read scientific and technological literature regularly), which produces awareness about the most recent scientific advancements and those working on them. Therefore, the search for useful knowledge possessed by specific others in specific places should be more efficient for inventors with high-level education.[3]

Second, education confers the ability to absorb knowledge spillovers and use them across contexts and applications. High-level education provides students with general methods and mental categories, as well as a deep understanding of theories about phenomena. This general and abstract knowledge base constitutes a framework to formulate research questions. It also helps researchers translate knowledge across different coding schemes (Allen & Cohen, 1969) and exploit it in different contexts. Haskell (2001) argues that a large knowledge base provides frameworks and mental connections and that failures in the transfer of learning are mostly due to the lack of an adequate knowledge of the underlying theories.[4]

By these means, a high level of education contributes to the capacity to scout for, absorb, and exploit knowledge spillovers. Thus, we posit that inventors with a high level of education use external knowledge spillovers to a greater extent than their less educated peers, regardless of whether the sources of these spillovers are geographically near or far.

Existing contributions maintain that geographical proximity to knowledge sources fosters unplanned interactions and face-to-face contacts and therefore reduces the effort needed to recognize and absorb useful knowledge. This effect should be particularly beneficial to inventors with less education, who have lower general absorptive capacities. It would also make the beneficial effect of education more evident in terms of enabling people to access and absorb knowledge spillovers from distant sources and in different contexts rather than in interactions with geographically close others. In other words, geographical proximity and other localized mechanisms for knowledge transfer could compensate, at least in part, for the lack of general and abstract cognitive skills, but they do not reduce the difficulty associated with gaining access to distant spillovers. Education instead makes access to knowledge spillovers

[2] At the individual researcher level, other studies have considered research productivity over time (Hall, Mairesse, & Turner, 2007; Levin & Stephan, 1991), gender differences (Ding, Murray, & Stuart, 2006), and technology transfers from universities to industry (Bercovitz & Feldman, 2008).

[3] Jones, Wuchty, and Uzzi (2008) show that scientists in top U.S. universities collaborate with geographically distant scientists and that the choice of the partnership is based on who the partner is rather than his or her geographical location, with a positive impact on the quality of research.

[4] In educational psychology, the general principle model indicates that knowledge about abstract general principles underlying a phenomenon supports the application of theory to new and different situations. Haskell (2001) shows the importance of theoretical backgrounds for people's ability to use knowledge in different contexts.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/449/1917213/rest_a_00259.pdf by guest on 28 September 2021

geographically unbound. We therefore expect the probability of using distant spillovers to be higher for inventors with higher education, even after we control for their better ability to absorb knowledge spillovers in general, regardless of geographic origins.[5]

Better access to distant sources of knowledge in turn gives inventors a larger pool of potentially useful spillovers compared with that available to their less educated peers. The larger the pool, the higher the probability is that inventors find hints that fit their needs, so they may substitute for hints received locally with better, more distant spillovers. If such a substitution takes place, we expect the probability of using geographically near spillovers to be lower for inventors with higher education.

III. Measure of Knowledge Spillovers

Because knowledge flows are invisible and leave no paper trail (Krugman, 1991), we collect direct information from the inventors themselves and document the use of knowledge spillovers to produce inventions without resorting to indirect indicators such as patent citations. The PatVal-EU survey provides interviews with inventors of 9,550 patents granted by the European Patent Office (EPO) between 1993 and 1998, located in Denmark, France, Germany, Hungary, Italy, the Netherlands, Spain, and the United Kingdom. We employ information from a subsample of 6,051 patents, which excludes patents with missing data.[6] The PatVal-EU survey is designed to represent the entire universe of patents in EU countries. It covers all technological fields; both for-profit and nonprofit applicants; and small, medium, and large businesses. It collects information about the individual inventors, their inventions, and the inventive processes. Giuri et al. (2007) report the details of the survey and the descriptive statistics. Appendix A describes the sample selection process and some technical issues.

In terms of the specific purpose of studying the geographical breadth of knowledge spillovers, the survey asks inventors to rate the following item from 0 (not used) to 5 (very important):

> "The importance of interactions such as discussions, meetings and sources of ideas with the following types of people (apart from co-inventors) during the research that led to the patented invention: (1) People belonging to unaffiliated organizations that typically takes less than an hour of travel time to reach their office or location (hereafter, *Near* people); (2) People belonging to unaffiliated organizations that typically takes more than an hour of travel time to reach their office or location (hereafter, *Far* people)."

Table 1.—Unconditional Probabilities: Share of patents Invented with Near or Far Interactions

	Far = 0	*Far* = 1	Total
Near = 0	54.1% (63.4%; ***53.7%***; **48.6%**)	15.3% (11.5%; ***15.9%***; **16.6%**)	69.4%
Near = 1	4.7% (5.4%; ***4.6%***; **4.5%**)	25.9% (19.7%; ***25.8%***; **30.3%**)	30.6%
Total	58.8%	41.2%	100%

N = 6,051. High school in plain roman; university in bold italics; PhD in bold. 0 = external interactions not used; 1: external interactions used, regardless of importance.
Source: PatVal-EU data set.

We deliberately define geographical proximity with a dichotomous distinction between less and more than 1 hour of travel time. This measure differs from the frequently used kilometric distance between inventors' addresses and stems from several rationales. First, kilometric distance does not reflect the effort and time needed to cover the distance, which depends on the topography that separates the interacting parties. Our measure overcomes this problem and increases comparability across responses. Second, our dichotomous measure is consistent with the bimodal distribution of distance described by existing studies (Gittelman, 2007). Third, the inventors found the question easy to answer because they did not have to recall or estimate specific kilometer distances. However, our variable does not provide detailed information about the breadth of interactions, particularly those that are more than 1 hour away.

In addition, the spillover indicator is based on inventors' self-evaluations, which could undermine the cross-country and cross-individual comparability of the variable. In our survey, cross-country differences in the ratings of questions that require a subjective assessment exist, but they are not systematic. To reduce potential bias, we transformed the 0–5 spillover scores into two dichotomous variables: *Near* equals 0 if the inventors do not establish any interactions with geographically near others during the research leading up to the patent (score = 0), and it equals 1 if they interact with nearby people, regardless of the importance of the interactions (score = 1–5). The same applies to *Far* (0/1) for interactions with distant people. Although the 1–5 score might be subject to the use of different metrics by respondents, this problem is unlikely to affect the distinction between no interactions at all and the use of interactions regardless of their importance. The genuine cross-country and cross-regional differences are addressed by the country (regional) dummy controls. Table 1 reports the share of patents invented with either near or far interactions.

More than half the patents (54.1%) do not include interactions with people external to the inventor's organization. The inventors of 25.9% of the patents establish external interactions regardless of distance (*Near* and *Far*). The share of patents developed in interactions with near others is only 4.7%, lower than the share of patents with only far interactions (15.3%). These numbers are consistent with

[5] The breadth of background and experience affects the capability to link different ideas or knowledge (Lazear, 2004; Singh & Fleming, 2010). We consider this possibility and test if education has a net effect on the probability of using external, distant spillovers, after controlling for the diversity of research experience.

[6] This subsample of 6,051 patents does not contain any French, Danish, and Hungarian inventors, because we do not have data for these countries about past patents and past coinventors, key variables for our empirical analysis.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/449/1917213/rest_a_00259.pdf by guest on 28 September 2021

previous results drawn from patents and publication data (Audretsch & Stephan, 1996; Gittelman, 2007; Jaffe et al., 1993). In line with our hypotheses about the role of education, we find that the share of patents developed through both *Near* and *Far* or only *Far* interactions is higher among people with doctoral or university degrees compared with inventors with high school diplomas. In contrast, the share of patents invented with only *Near* interactions is greater for people with high school diplomas.

To improve our understanding of the nature of these interactions, we conducted a follow-up survey of a sample of 100 inventors who also responded to the PatVal-EU survey. The survey was conducted from January to April 2010 by e-mail and telephone and was pretested in five pilot interviews. Consistent with our framework, we find that the inventors did not turn to external interactions if they, their team, or their employer already possessed the necessary competencies to develop the invention. In addition, the survey supported our sense that geographical proximity can foster "accidental," or unplanned, knowledge spillovers, but education provides the cognitive skills that enable inventors to scout for and plan interactions. Inventors reported that chance determined 25% of their geographically near versus 15% of their far interactions. For inventors with high school diplomas, chance mattered in 50% of their near and 27% of their far interactions, whereas for inventors with university or PhD degrees, chance drove only 20% of their near interactions and 12% of the far ones. Thus, geographically far interactions are planned in advance, and inventors with more education plan both near and far interactions more than do their less-educated peers.[7]

IV. Empirical Analysis: Method and Measures

The dichotomous *Near* and *Far* measures serve as dependent variables in two equations that we estimate with a bivariate probit model. This estimation enables us to compute the marginal effects of the regressors on the predicted probabilities of the four combinations of outcomes and then obtain the effect of the covariates on the geographical breadth of knowledge spillovers after we control for the probability of resorting to external spillovers irrespective of geography.[8]

A. Level of Education and other Inventors' Controls

The level of inventors' education is our key covariate. The survey provides information about the highest degree earned by the inventors at the time of their invention. We employ three dummy variables—secondary or high school (High School Degree), university BSc or master's (University Degree), and PhD (PhD Degree) degrees—to test our hypothesis. To estimate the effect of education net of other factors, we employ a large set of controls, including the location of prior coinventors, invention, organization, and location characteristics.

From the EPO-PatStat database, we downloaded all the patents to which our inventors had contributed in the ten years prior to the focal patent and collected the names of all coinventors listed in those patents. We describe the search and matching procedure in appendix A. By using the address of each past coinventor, we define him or her as "geographically near" the PatVal-EU inventor if the address was within the same NUTS3 region where the research that led to the PatVal-EU patent had been performed. The coinventor is "geographically far" if the address was in a different NUTS3. We then construct the variable Far Past Coinventors, or the ratio of the number of geographically far to all past coinventors.[9] Because the network of past coinventors might depend on the length of the inventor's career, we control for the year of the inventor's first patent (Experience). A dummy variable (Coinventors Dummy) controls for whether the inventor collaborated with at least one coinventor in the ten years before the focal patent.

For personal connections established during educational training that did not result in past joint patents, we include a dummy variable (Country of Degree Dummy) that indicates whether the inventor's last degree was earned in the country where the research that led to the patent was conducted (0) or in a different country (1). This variable controls for situations in which spillovers come from former classmates or mentors located in a different country. Because inventors who perform more science-based research also might interact more with members of epistemic communities, we include the variable Science to indicate the importance of scientific literature as a source of knowledge for the research leading to the patent. This survey variable ranges between 0 (not used) and 5 (very important). Finally, we include the inventor's age (Age) and gender (Male: 1 if male; 0 if female), as well as two dichotomous variables that indicate the inventor's job mobility in the ten years before the surveyed patent: Mobility Out Region equals 1 if the inventor has changed employer and NUTS3 location, and 0 otherwise; Mobility in Region equals 1 if the inventor has changed employer within the same NUTS3 region, and 0 otherwise. To build these variables, we asked the inventors to reveal the name of their previous employers, if any,

[7] We thank two anonymous referees whose questions prompted us to undertake the small-scale survey. Space considerations prevent us from reporting more data, such as who the interacting parties are or the answers to a set of anchoring vignettes for cross-country comparability. These data are available on request.

[8] As a robustness check, we used the 0–5 importance score of near and far interactions as dependent variables of two ordered probit regressions. There were no relevant changes in the signs or statistical significance of the estimated effects compared with those from the bivariate probit model.

[9] The PatVal-EU survey asks for the postal code of the location where the research that led to the patent was conducted. We assign each reported postal code to its corresponding NUTS2 (for example, *Länder* in Germany, *regioni* in Italy) and NUTS3 region (for example, districts in Germany, provinces in Italy). The list of regions is available at http://epp.eurostat.ec.europa.eu. If a coinventor appears in multiple previous patents, he or she is counted once. As robustness checks, we employ the share between distant and total ties (instead of inventors) and the share between distant and near coinventors, with no relevant changes.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/449/1917213/rest_a_00259.pdf by guest on 28 September 2021

before working for the organization where they conducted the research leading to the patent. With information from Amadeus and company websites, we traced their mobility across employers and regions.[10] A dummy variable (Missing Mobility Dummy) controls for cases in which we could not find information on the inventors' previous employer or its location.

Because variety in inventors' skills also might affect their capacity to absorb and link knowledge from different sources (Lazear, 2004; Singh & Fleming, 2010), we constructed a variable based on inventive experience, which indicates the change in research fields between inventions (see also Jones, 2009): Experience Herfindahl is a Herfindahl index computed for the inventor's patents across technological classes.[11] A dummy variable controls for whether inventors have at least one patent before the surveyed invention (Previous Patents Dummy = 0 if the inventor has no other patents; 1 otherwise).

B. *Patented Invention*

The variable Number of Inventors refers to the number of inventors listed in the patent and helps control for the scale and complexity of the research project (Gittelman & Kogut, 2003; Jones, 2009). We also control for the use that the inventors foresee for the patent, which influences the extent to which they interact with external parties. The PatVal-EU survey provides information about the importance of commercial exploitation (Commercial Exploitation), licensing (Licensing), and prevention from imitation (Prevent Imitation) as reasons to patent. We measure this importance on a scale from 0 (not important) to 5 (very important).

C. *Applicant Organization*

The attributes of the applicant organization may affect the need, costs, and benefits of using near versus far interactions. Approximately 93% of the patents in our database are owned by businesses; the remaining 7% are granted to individual inventors and public research organizations, such as universities, which we code as PRI Applicant (1 = applicant is a public research institution) or Individual Applicant (1 = individual inventor). For patents granted to private companies, we collect information about the size and R&D intensity of the firms at the parent company level (average 1990–1996) from the Compustat (1998) and Amadeus (2005) databases. We use the number of employees (Employees) as a proxy for the size of the firms and the ratio between R&D expenditure and sales (R&D Intensity) as a measure of R&D intensity.[12]

D. *Regional and Other Controls*

To control for any possible effect of the research milieu, we use regional dummies for the NUTS2 regions or NUTS3 provinces in which the research that led to the patent was performed. In alternative specifications, instead of regional dummies, we employ country dummies and measures of the technological, economic, and demographic characteristics of the locations. We also include dummy variables for the application year (1993–1998) and for the thirty ISI-INPI-OST technological classes of the patents, which we list in Appendix B. In table 2, we provide the descriptive statistics of the variables. Table 3 contains the correlation matrix among the variables, including those used in the robustness checks in section VC.

V. Results

A. *Univariate Probabilities*

Table 4 reveals the marginal effects of a one-unit change in the covariates on the probability of Near and Far interactions. The results support the positive effect of a high level of education on the probability of gathering knowledge spillovers from distant others.

All specifications in table 4 include dummies for regions, application year, and technological field of the patent. They cluster robust estimators on firms, and because our sample differs from the complete data set in terms of nonresponses and oversampling of "important" patents, all specifications include sampling weights for the hypothetical unbiased sample distribution (see Giuri et al., 2007).

The signs of University Degree and PhD Degree are positive and statistically significant on Far (1% level), but other than in model 1, where we do not control for the networking capabilities of the inventors, they do not correlate with the probability of Near interactions. These results are robust to the inclusion of all networking controls (models 2–3) and to the use of the NUTS3 regional dummies (model 4) in place of the NUTS2 dummies. Also, in the baseline model 1, the magnitude of the effect of education is higher than in models 2 to 4.[13]

We also note the signs of the control variables that are statistically significant. Inventors rely less on external spillovers as they age: Age relates negatively and significantly (at 1% level) to both Near and Far. More science-based research instead relies more on external spillovers: Science is positive for both Near and Far at the 1% level. Mobility

[10] We purposefully limit our investigation to one employer change to keep the questionnaire simple and brief.

[11] Unfortunately, we do not have information about the subject matter studied during formal schooling. Technological fields are defined according to the 30 ISI-INPI-OST classes. Jones (2009) notes that technical classes are assigned to patents rather than inventors. This limitation is less relevant in our study because we need a variable that measures the extent to which inventors are "contaminated" by knowledge from different fields during their research experience.

[12] Because the Employees variable data are available for 80% of the patents, whereas data on R&D Intensity are available for 45% of the patents, we include two dummy variables for missing observations (Missing Employees Dummy and Missing R&D Dummy).

[13] The null hypothesis that the estimated effects of University Degree and PhD Degree are statistically not different is rejected for both near ($p = 0.03$) and far ($p = 0.00$) interactions, according to model 1 in table 4. In the remaining specifications, the difference is significant only in the Far equation.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/449/1917213/rest_a_00259.pdf by guest on 28 September 2021

TABLE 2.—DESCRIPTIVE STATISTICS

	Mean	SD	Minimum	Maximum
Dependent variables				
Near	0.31	0.46	0	1
Far	0.41	0.49	0	1
Inventor characteristics				
Age	44.76	9.61	20	78
Male	0.98	0.15	0	1
High School Degree	0.18	0.39	0	1
University Degree	0.55	0.50	0	1
PhD Degree	0.27	0.44	0	1
Far Past Coinventors[a]	0.60	0.38	0	1
Coinventors Dummy	0.52	0.50	0	1
Experience	1989	5.5	1,977	1,997
Science	2.63	1.86	0	5
Country of Degree Dummy	0.04	0.19	0	1
Mobility in Region	0.07	0.25	0	1
Mobility Out Region	0.20	0.40	0	1
Missing Mobility Dummy	0.07	0.25	0	1
Experience Herfindahl[b]	0.78	0.26	0.12	1
Previous_Patents Dummy	0.64	0.48	0	1
Inventor Past Patents	5.04	12.10	0	327
Conferences	1.72	1.73	0	5
Applicant controls				
Employees[c]	89,052	116,337	0	723,329
Missing Employees Dummy	0.20	0.40	0	1
R&D Intensity[d]	0.05	0.03	0	0.41
Missing R&D Dummy	0.55	0.50	0	1
PRI Applicant	0.03	0.17	0	1
Individual Applicant	0.04	0.20	0	1
Patent controls				
Number of Inventors	2.36	1.49	1	15
Commercial Exploitation	3.79	1.55	0	5
Licensing	2.07	1.52	0	5
Prevent Imitation	3.81	1.55	0	5
Regional and other controls				
GDPPC	23,202.06	9,165.44	8,677.90	76,910.80
Pop	737.35	897.27	37.60	4,634.40
Area	1,519.40	1,934.58	35.60	17,252.00
Regional Patents	123.72	137.40	0.83	543.21
TOP1% in Technology	0.15	0.36	0	1
NR Research Universities	0.56	0.78	0	5
Research University Score	1.20	6.00	0	34.10
Country				
Germany	0.44	0.50	0	1
Italy	0.17	0.37	0	1
Spain	0.03	0.17	0	1
Netherlands	0.16	0.37	0	1
United Kingdom	0.19	0.40	0	1
Year				
1993	0.03	0.17	0	1
1994	0.28	0.45	0	1
1995	0.27	0.44	0	1
1996	0.23	0.42	0	1
1997	0.15	0.36	0	1
1998	0.04	0.20	0	1

$N = 6{,}051$.[a] $N = 3{,}134$.[b] $N = 3{,}877$.[c] $N = 4{,}850$.[d] $N = 2{,}706$.

and diverse inventive experience positively correlate with the probability of resorting to distant interactions. Both mobility variables are positive for Far (at 10% level), though the effect decreases in model 4. Experience Herfindahl is negative for Far, at less than the 5% level.[14] Inventors also engage more in external interactions if the invention is designed to be licensed; Licensing has a similar marginal effect for both Near and Far. As expected, the higher the R&D intensity of the employer firm, the lower is the probability of establishing external interactions (statistically significant at the 5% level).[15] Thus, even after control-

[14] In an unreported specification, in place of the Herfindahl variable, we employ a dummy variable that takes a value of 0 if all the inventor's past patents are in the same technological class, and 1 otherwise. The estimated marginal effect of the dummy variable is positive for Far.

[15] Only a few technological dummies are statistically significant and different from 0. To check the effect of the technological field, we compared the results in table 4 with specifications that exclude the technological dummies and found no relevant changes. We also estimated the models for subsamples of technologies selected according to the importance of Science. The few observations available in many of the thirty classes prevented us from running separate regressions. With some minor differences, education matters in both more and less science-intensive technologies.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/449/1917213/rest_a_00259.pdf by guest on 28 September 2021

Table 3.—Correlation Matrix.

	Variable	1	2	3	4	5	6	7	8	9	10
1	Age										
2	Male	.126*									
3	University Degree	−.074*	−.021								
4	PhD Degree	−.011	−.015	−.666*							
5	Far Past Coinventors	.099*	.020	−.051*	.196*						
6	Experience	.353*	.054*	.069*	.172*	.468*					
7	Science	−.078*	−.073*	−.064*	.286*	.118*	−.087				
8	Country of Degree Dummy	.032*	−.024	−.030*	.059	.010	−.001	.031*			
9	Mobility In Region	−.052*	.003	−.007	−.034*	−.094*	.084*	−.010	−.027*		
10	Mobility Out of Region	−.049*	.023	−.030*	.062*	−.053*	.058*	.004	.002	−.136*	
11	Missing Mobility Dummy	.002	−.011	.003	−.010	−.009	.037	−.006	.012	−.071	−.134
12	Experience Herfindahl	.089*	.031*	−.014	.075*	.417*	−.429*	.021	.006	−.017	−.037*
13	Employees	−.019	.023	.008	.125*	.086*	−.125*	.068*	.031*	−.060*	−.044*
14	Missing Employees Dummy	.050*	−.003	.018	−.100*	−.111*	.140*	−.069*	.016	.039*	.021
15	R&D Intensity	−.056*	−.013	−.051*	.200*	.123*	−.166*	.124*	.011	−.055*	−.006
16	Missing R&D Dummy	.055*	.013	.061*	−.209*	−.135*	.174*	−.108*	−.008	.064*	−.034*
17	PRI Applicant	−.034*	−.045*	−.035*	.103*	−.021	.063*	.096*	−.003	.018	.018
18	Individual Applicant	.070*	.027*	−.020	−.060*	−.089*	.076*	−.048*	.011	−.004	−.026*
19	Number of Inventors	−.073*	−.086*	−.098*	.239*	.159*	−.114*	.215*	−.001	−.032*	.002
20	Commercial Exploitation	.035*	.009	.008	−.016	−.002	−.008	.037*	−.003	.023	.042*
21	Licensing	−.064*	.006	−.011	.088*	.002	.048*	.152*	−.005	.001	.003
22	Prevent Imitation	−.063*	−.009	.011	−.039*	.059*	−.016	.005	−.015	.017	−.027*
23	GDPPC	.014	.027*	−.085*	.192*	.263*	−.206*	.108*	.014	−.037*	−.114*
24	Population	−.076*	−.111*	.073*	−.114*	−.105*	.076*	.024	.006	.110*	−.046*
25	Area	−.050*	−.070*	.031*	−.150*	−.236*	.174*	−.044*	−.007	.113*	−.012
26	Reg Patents	−.045*	−.027*	.052*	−.001	−.005	−.081*	.047*	.031*	.073*	−.074*
27	ToP1% in Technology	.024	.023	−.086*	.158*	.152*	−.183*	.041*	.010	−.041*	−.083*
28	NR Research Universities	−.071*	−.069*	.038*	.007	−.049*	.046*	.052*	.012	.089*	−.023
29	Research University Score	.005	.014	.008	.046*	.072*	−.054*	.010	.008	.013	−.042*
30	Inventor Past Patents	.171*	.037*	−.113*	.207*	.318*	−.514*	.095*	−.015	−.037*	−.061*
31	Conferences	.050*	−.027*	−.034*	.188*	.108*	−.076*	.428*	.024	−.023	−.005

	Variable	11	12	13	14	15	16	17	18	19	20
1	Age										
2	Male										
3	University Degree										
4	PhD Degree										
5	Far Past Coinventors										
6	Experience										
7	Science										
8	Country of Degree Dummy										
9	Mobility In Region										
10	Mobility Out of Region										
11	Missing Mobility Dummy										
12	Experience Herfindahl	−.005									
13	Employees	−.042	.072*								
14	Missing Employees Dummy	.036	−.126*	−.323*							
15	R&D Intensity	−.051	.102*	.596*	−.343*						
16	Missing R&D Dummy	.059	−.096*	−.585*	.438*	−.779*					
17	PRI Applicant	.007	−.044*	−.110*	.343*	−.119*	.153*				
18	Individual Applicant	.017	−.109*	−.133*	.411*	−.143*	.184*	−.035*			
19	Number of Inventors	−.034	.048*	.086*	−.137*	.147*	−.158*	.045*	−.138*		
20	Commercial Exploitation	−.012	.009	−.085*	.029*	−.041*	.019	−.089*	.031*	.015	
21	Licensing	.015	−.046*	.042*	.114*	−.003	.024	.141*	.155*	.031*	.095*
22	Prevent Imitation	−.010	.058*	−.012	−.026*	.010	.005	−.087*	−.042*	.010	.149*
23	GDPPC	−.032	.099*	.231*	−.144*	.193*	−.155*	−.057*	−.047*	.163*	−.106*
24	Population	.014	−.037*	−.075*	−.019	−.062*	.069*	.012	−.009	−.002	−.005
25	Area	−.023	−.088*	−.149*	.046*	−.167*	.165*	.034*	.050*	−.083*	.027*
26	Reg Patents	−.020	.042*	.211*	−.100*	.154*	−.098*	−.050*	−.048*	.033*	−.086*
27	TOP1% in Technology	−.032	.113*	.184*	−.121*	.163*	−.140*	−.054*	−.045*	.106*	−.082*
28	NR Research Universities	.022	−.023	.058*	−.018	.049*	−.033*	.059*	−.019	.006	−.014
29	Research University Score	−.008	.050*	.195*	−.042*	.103*	−.055*	.005	−.009	−.003	−.029*
30	Inventor Past Patents	−.032	.179*	.084*	−.091*	.140*	−.145*	−.032*	−.052*	.138*	.000
31	Conferences	.001	.039*	.109*	−.076	.101*	−.084*	.087*	−.091*	.142*	.003

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/449/1917213/rest_a_00259.pdf by guest on 28 September 2021

Table 3.—(Continued)

	Variable	21	22	23	24	25	26	27	28	29	30
1	Age										
2	Male										
3	University Degree										
4	PhD Degree										
5	Far Past Coinventors										
6	Experience										
7	Science										
8	Country of Degree Dummy										
9	Mobility In Region										
10	Mobility Out of Region										
11	Missing Mobility Dummy										
12	Experience Herfindahl										
13	Employees										
14	Missing Employees Dummy										
15	R&D Intensity										
16	Missing R&D Dummy										
17	PRI Applicant										
18	Individual Applicant										
19	Number of Inventors										
20	Commercial Exploitation										
21	Licensing										
22	Prevent Imitation	.004									
23	GDPPC	.002	.044*								
24	Population	−.009	−.042*	.028*							
25	Area	−.016	−.034*	−.295*	.438*						
26	Reg Patents	.023	−.072*	.271*	.586*	.030*					
27	TOP1% in Technology	.025	.046*	.370*	−.105*	−.190*	.224*				
28	NR Research Universities	.045*	−.070*	.196*	.618*	.123*	.545*	−.016			
29	Research University Score	.019	.028*	.282*	.078*	−.113*	.404*	.183*	.338*		
30	Inventor Past Patents	−.023	.040*	.231*	−.064*	−.118*	.034*	.179*	−.056*	.009	
31	Conferences	.149	.031*	.090*	−.030*	−.095*	.060*	.082*	.020	.056*	.082*

*$p < 0.10$.

ling for all these other factors, higher education predicts better access to spillovers produced far away.

B. *Bivariate Probabilities*

Our theoretical framework posits that education increases the capacity to absorb external knowledge spillovers regardless of the geographical location of their sources. When we control for this factor, education should predict better access to geographically distant spillovers as well. The results of the bivariate probit regressions confound these two effects, in that the dependent variables capture both dimensions of the spillovers. To disentangle them, we estimate the marginal effects of the regressors on the predicted probabilities of the four combinations of outcomes computed from the bivariate probit model: Near&Far Interactions, corresponding to the predicted probability of Near = 1 and Far = 1; OnlyNear Interactions, which is the predicted probability of Near = 1 and Far = 0; OnlyFar Interactions, or the predicted probability of Near = 0 and Far = 1; and No Interactions, the predicted probability of Near = 0 and Far = 0.

The estimated results corroborate our hypothesis that apart from increasing the capacity to absorb external knowledge spillovers in general, education predicts better access to spillovers from geographically distant individuals. Table 5 reports the results estimated for model 3 in table 4.[16]

[16] We obtain similar results for the specification that employs NUTS3 regional dummies.

The first two columns in table 5 indicate the effects of the covariates on the use of external spillovers. The first column shows the effect of the covariates on the probability of interacting with people external to the inventor's organization, regardless of their geographical distance (that is, the predicted probability of Near&Far Interactions). The second reports the effects on the probability of not having any external interactions (the predicted probability of No Interactions). The last two columns in table 5 show the effect of the regressors on the geographical breadth of the spillovers (the predicted probability of OnlyNear versus OnlyFar).

The marginal effects in the first and second columns confirm that higher education confers boundary-spanning roles on people (Allen & Cohen, 1969), in that it affects the probability of using external spillovers, both near and far. PhD Degree is positive and statistically significant for the predicted probability of Near&Far Interactions but negative and statistically significant for the predicted probability of No Interactions. The same role pertains to more science-based research (Science) and inventions that are patented to be licensed (Licensing). In contrast, older inventors and higher firm R&D intensity both lower the likelihood of resorting to external interactions. This effect of Age is consistent with results that indicate individual research productivity decreases with age (Haeussler, 2011; Levin & Stephan, 1991), and the effect of R&D intensity confirms that inventors do not resort to external interactions when their team or organization has the competencies needed for

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/449/1917213/rest_a_00259.pdf by guest on 28 September 2021

TABLE 4.—BIVARIATE PROBIT ESTIMATION: MARGINAL EFFECTS ON THE UNIVARIATE PROBABILITY OF *NEAR* AND *FAR*

	1. University Degree and PHD Degree		2. Far Past Coinventors and Science		3. Experience Herfindahl		4. NUTS3 regions	
	Near	*Far*	*Near*	*Far*	*Near*	*Far*	*Near*	*Far*
Inventor characteristics								
Age	−0.004***	−0.003***	−0.004***	−0.003***	−0.004***	−0.003***	−0.004***	−0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Male	0.016	−0.032	0.023	−0.019	0.023	−0.023	−0.008	−0.055
	(0.045)	(0.044)	(0.044)	(0.043)	(0.044)	(0.044)	(0.046)	(0.045)
University Degree	0.035*	0.090***	0.015	0.071***	0.015	0.069***	0.019	0.070***
	(0.021)	(0.022)	(0.021)	(0.022)	(0.021)	(0.022)	(0.020)	(0.023)
PhD Degree	0.072***	0.141***	0.036	0.104***	0.035	0.097***	0.039	0.101***
	(0.025)	(0.026)	(0.025)	(0.027)	(0.025)	(0.027)	(0.025)	(0.029)
Far Past Coinventors	–	–	0.018	0.033	0.020	0.033	0.026	0.041
			(0.027)	(0.030)	(0.027)	(0.030)	(0.026)	(0.031)
Science	–	–	0.029***	0.038***	0.029***	0.038***	0.029***	0.039***
			(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
Experience	–	–	−0.002	−0.001	−0.002	−0.000	−0.002	0.000
			(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Country of Degree Dummy	–	–	–	–	−0.043	−0.018	−0.031	−0.030
					(0.029)	(0.035)	(0.030)	(0.036)
Mobility In Region	–	–	–	–	0.040	0.053*	0.027	0.043
					(0.030)	(0.030)	(0.028)	(0.031)
Mobility Out Region	–	–	–	–	0.021	0.031*	0.016	0.022
					(0.019)	(0.019)	(0.018)	(0.019)
Experience Herfindahl	–	–	–	–	−0.015	−0.094**	−0.019	−0.111***
					(0.037)	(0.044)	(0.034)	(0.042)
Applicant controls								
log(Employees)	0.001	0.005	−0.001	0.004	−0.001	0.005	−0.006	0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
R&D Intensity	−0.745**	−0.825**	−0.798***	−0.881**	−0.803***	−0.859**	−0.728**	−0.703**
	(0.304)	(0.347)	(0.331)	(0.365)	(0.300)	(0.382)	(0.293)	(0.352)
PRI Applicant	−0.011	0.051	−0.025	0.031	−0.024	0.033	−0.025	0.015
	(0.042)	(0.044)	(0.039)	(0.043)	(0.040)	(0.044)	(0.038)	(0.044)
Individual Applicant	−0.045**	−0.057**	−0.044**	−0.061***	−0.043**	−0.058**	−0.041**	−0.059**
	(0.020)	(0.024)	(0.020)	(0.024)	(0.020)	(0.023)	(0.019)	(0.024)
Patent controls								
Number of Inventors	0.001	0.008	−0.002	0.004	−0.002	0.004	0.001	0.006
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
Commercial Exploitation	0.004	0.008	0.003	0.006	0.003	0.006	0.003	0.007
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
Licensing	0.028***	0.033***	0.024***	0.027***	0.023***	0.027***	0.022***	0.029***
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
Prevent Imitation	0.009**	0.007	0.008*	0.006	0.008*	0.006	0.008*	0.005
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
Region controls								
NUTS2 regional dummies		Yes		Yes		Yes		
NUTS3 regional dummies								Yes
N		6,051		6,051		6,051		6,051
Ll		−29,231.27		−28,976.35		−28,923.83		−2,7452.49
Rho	0.818***	(0.012)	0.814***	(0.012)	0.814***	(0.012)	0.836***	(0.011)

Robust standard errors are in parentheses, adjusted for clusters by firms' identifier. Models 1–3 include dummies for NUTS2 regions; model 4 uses dummies for NUTS3 regions. All regressions include dummies for missing values for employees, missing values for R&D intensity, missing values for MOBILITY, year of application, and technological field (30 ISI-INPI-OST classes). Specifications 2–4 include a dummy variable for inventors with no past coinventors; specifications 3 and 4 include a dummy variable for inventors with no prior patents. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

the project.[17] Compared with an average R&D-intensive firm, a firm in the 90th percentile of the distribution exhibits a 2.8% lower probability of using external spillovers, and one in the 99th percentile is 6.6% less likely to need them. Comparatively, the marginal effect of PhD Degree is 4.8% for the predicted probability of Near&Far and –8.5% for the predicted probability of No Interactions. A change in Science, from the sample mean to the 99th percentile, corresponds to a 7.3% increase in external interactions. A ten-year increase in Age decreases the probability of using external spillovers by about 3.3%. Finally, the mobility variables are positive but not statistically significant for the predicted probability of Near&Far Interactions and negative and statistically significant for No Interactions.

The marginal effects in the third and fourth columns of table 5 support our hypothesis that education contributes to people's access to geographically wider knowledge spillovers. Both Phd and University Degree are positive and statistically significant at the 1% level on the predicted

[17] By using information about interactions with people internal to the organization, we also note that inventors with higher education interact more with people internal to the employer organization, compared with their less educated peers. Older inventors interact less with people internal to the organization. We tried a specification with Age^2 and Age to test for possible nonlinear effects (Cole, 1979), but Age^2 was never statistically significant.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/449/1917213/rest_a_00259.pdf by guest on 28 September 2021

TABLE 5.—BIVARIATE PROBIT ESTIMATION (MODEL 3 IN TABLE 4): MARGINAL EFFECTS ON THE BIVARIATE PROBABILITIES OF *NEAR* AND *FAR*

	Near&Far Interactions	*No Interactions*	*OnlyNear Interactions*	*OnlyFar Interactions*
Inventor characteristics				
Age	−0.003***	0.004***	−0.001***	0.001
	(0.001)	(0.001)	(0.000)	(0.001)
Male	0.010	0.010	0.012	−0.033
	(0.037)	(0.043)	(0.011)	(0.032)
University Degree	0.028	−0.056***	−0.013*	0.042***
	(0.017)	(0.021)	(0.007)	(0.015)
PhD Degree	0.048**	−0.085***	−0.013*	0.050***
	(0.022)	(0.026)	(0.007)	(0.018)
Far Past Coinventors	0.022	−0.032	−0.002	0.011
	(0.024)	(0.030)	(0.007)	(0.017)
Science	0.029***	−0.038***	0.000	0.009***
	(0.004)	(0.004)	(0.001)	(0.003)
Experience	−0.002	0.001	−0.001	0.001
	(0.002)	(0.002)	(0.001)	(0.001)
Country of Degree Dummy	−0.033	0.028	−0.010	0.015
	(0.025)	(0.034)	(0.008)	(0.022)
Mobility In Region	0.040	−0.052*	0.000	0.013
	(0.026)	(0.029)	(0.009)	(0.020)
Mobility Out Region	0.022	−0.030*	−0.001	0.010
	(0.016)	(0.018)	(0.006)	(0.013)
Experience Herfindahl	−0.034	0.074*	0.019**	−0.060***
	(0.033)	(0.042)	(0.009)	(0.022)
Applicant controls				
Log(Employees)	0.001	−0.003	−0.001	0.004
	(0.003)	(0.004)	(0.001)	(0.003)
R&D Intensity	−0.733***	0.929***	−0.070	−0.126
	(0.259)	(0.358)	(0.079)	(0.200)
PRI Applicant	−0.009	−0.018	−0.015	0.043
	(0.034)	(0.042)	(0.009)	(0.032)
Individual Applicant	−0.042**	0.056***	0.001	−0.016
	(0.017)	(0.023)	(0.006)	(0.014)
Patent controls				
Number of Inventors	0.000	−0.002	−0.002	0.005
	(0.004)	(0.005)	(0.001)	(0.003)
Commercial Exploitation	0.004	−0.006	−0.001	0.003
	(0.004)	(0.005)	(0.001)	(0.003)
Licensing	0.022***	−0.028***	0.001	0.005
	(0.004)	(0.005)	(0.001)	(0.003)
Prevent Imitation	0.007*	−0.007	0.001	−0.001
	(0.003)	(0.005)	(0.001)	(0.003)

Robust standard errors are in parentheses, adjusted for clusters by firms' identifier. All regressions include dummies for NUTS2 regions, missing values for employees, missing values for R&D intensity, missing values for MOBILITY, year of application and technological field (30 ISI-INPI-OST classes), a dummy variable for inventors with no prior patents, and one for inventors with no past coinventors. *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$.

probability of OnlyFar Interactions but negative and statistically significant at the 10% level on the predicted probability of OnlyNear Interactions.[18] In addition, consistent with the existing literature, we find that more scientific research (Science) correlates with the probability of interacting with distant others. Older inventors rely less on external knowledge spillovers, but when they do, they interact less with geographically near people than do younger inventors. Diverse inventive experience contributes to the absorption of distant knowledge spillovers: Experience Herfindahl is positive and statistically significant for the predicted probability of OnlyNear Interactions and negative and statistically significant for the predicted probability of OnlyFar Interactions. Of the different factors, inventors' education and diversity of research experience have the greatest marginal effects on the probability of distant interactions.[19] In summary, even after controlling for alternative explanations, inventors with a high level of education use external spillovers more, regardless of the geographical location of their sources, and they also access geographically wider knowledge spillovers than do their less educated peers.

C. Robustness Checks

We report some additional checks of the robustness of our finding that higher education enhances access to external and geographically distant knowledge spillovers, after

[18] A Guttman-type variable for education in place of the dummy variables confirms our findings. We thank an anonymous reviewer for this suggestion.

[19] Although not shown here, this result holds for experience diversity measured with the dummy variable. We also checked if the impact of the covariates is influenced by education and reestimated our models for three samples of inventors, distinguished by their level of education. These results are available from the authors. The impact of Science decreases for inventors with PhD degrees. For inventors with high school diplomas, diversity of experience is not statistically significant, and R&D Intensity is positive and statistically significant for the probability of relying on both external and geographically close spillovers.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/449/1917213/rest_a_00259.pdf by guest on 28 September 2021

TABLE 6.—ROBUSTNESS CHECKS: BIVARIATE PROBIT ESTIMATION: MARGINAL EFFECTS ON THE BIVARIATE PROBABILITIES OF *NEAR* AND *FAR* (MODEL 5 IN TABLE 5)

	Near&Far Interactions	*No Interactions*	*OnlyNear Interactions*	*OnlyFar Interactions*
Robustness check 1				
University Degree	0.034*	−0.069***	−0.016**	0.052***
	(0.020)	(0.024)	(0.008)	(0.018)
PhD Degree	0.020	−0.065**	−0.020**	0.065**
	(0.026)	(0.032)	(0.008)	(0.024)
Far Past Coinventors	0.020	−0.031	−0.002	0.013
	(0.024)	(0.030)	(0.007)	(0.017)
University Degree × Coinventors Dummy	0.011	0.030	0.011	−0.030
	(0.035)	(0.042)	(0.013)	(0.026)
Phd Degree × Coinventors Dummy	0.044	0.032	0.022	−0.035
	(0.043)	(0.050)	(0.016)	(0.027)
Robustness check 2				
University Degree	0.020	−0.052*	−0.017*	0.048**
	(0.022)	(0.028)	(0.009)	(0.020)
PhD Degree	0.080**	−0.138***	−0.018*	0.077**
	(0.039)	(0.042)	(0.011)	(0.032)
Science	0.029***	−0.041***	−0.001	0.013
	(0.007)	(0.009)	(0.003)	(0.008)
University Degree × Science	0.002	−0.001	0.002	−0.003
	(0.007)	(0.010)	(0.003)	(0.008)
Phd Degree × Science	−0.009	0.016	0.002	−0.009
	(0.010)	(0.012)	(0.005)	(0.011)
Robustness check 3				
University Degree	0.022	−0.049**	−0.013*	0.040***
	(0.017)	(0.022)	(0.007)	(0.015)
PhD Degree	0.033	−0.067**	−0.013*	0.046**
	(0.022)	(0.027)	(0.007)	(0.019)
Science	0.013***	−0.017***	0.000	0.004
	(0.004)	(0.005)	(0.001)	(0.003)
Conferences	0.045***	−0.061***	0.000	0.016***
	(0.004)	(0.005)	(0.001)	(0.003)
Robustness check 4				
University Degree	0.027	−0.057***	−0.013**	0.042***
	(0.017)	(0.022)	(0.007)	(0.015)
PhD Degree	0.047**	−0.086***	−0.014**	0.052***
	(0.022)	(0.026)	(0.007)	(0.018)
Science	0.029***	−0.038***	0.000	0.009***
	(0.004)	(0.004)	(0.001)	(0.003)
Inventor Past Patents	0.003	−0.003	0.001	0.001
	(0.010)	(0.014)	(0.004)	(0.009)

Robust standard errors are in parentheses, adjusted for clusters by firms' identifier. All regressions include dummies for NUTS2 regions, missing values for employees, missing values for R&D intensity, missing values for Mobility, year of application and technological field (30 ISI-INPI-OST classes), a dummy variable for inventors with no prior patents, and one for inventors with no past coinventors. *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$.

controlling for other sources of heterogeneity. We first consider the possibility that the effect of PhD and university degrees relates to the networking ability of the inventors. We construct two interaction variables, PhD Degree × Coinventors Dummy and University Degree × Coinventors Dummy, where Coinventors Dummy indicates whether the inventor has collaborated with at least one other coinventor in the ten years before the focal patent. We then estimate model 3 with these two terms. The upper part of table 6 shows the marginal effects of the five covariates of interest on the predicted probabilities of the four combinations of outcomes, computed after the bivariate probit model. We do not provide the estimates of the other covariates, which remain largely unchanged from those in table 5.

The effect of education holds for both inventors with and without past collaborations. The marginal effects and statistical significance of Phd and University Degree are similar to those in table 5, and the interacted terms are not statistically significant for any output variables.

As a second check, we consider the role of scientific communities that might not appear in the Science variable or translate into past coinventions. If inventors with PhD or university degrees take advantage of their scientific community, Science as a source of knowledge should be more important for them than for inventors with less education, who are less likely to be part of these scientific communities. We therefore incorporate the interaction terms Phd Degree × Science and University Degree × Science in model 3. The estimated results confirm that education matters beyond personal scientific communities. The marginal effects of the interacted terms are not statistically significant in terms of the probability of accessing either Near or Far spillovers. The terms University Degree and Phd Degree retain their signs and statistical significance.

Because people from the same scientific communities likely meet at conferences and workshops, we incorporate a variable provided by the PatVal-EU survey that measures, on a scale of 0 to 5, the importance of conferences and workshops as sources of knowledge. The results indicate

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/449/1917213/rest_a_00259.pdf by guest on 28 September 2021

TABLE 7.—BIVARIATE PROBIT ESTIMATION WITH REGIONAL VARIABLES: MARGINAL EFFECTS ON THE BIVARIATE PROBABILITIES OF *NEAR* AND *FAR*

	Near&Far Interactions	*No_ Interactions*	*OnlyNear Interactions*	*OnlyFar Interactions*
Inventor characteristics				
Age	−0.003***	0.003***	−0.001**	0.000
	(0.001)	(0.001)	(0.000)	(0.001)
Male	0.010	0.015	0.015	−0.039
	(0.038)	(0.043)	(0.011)	(0.032)
University Degree	0.029*	−0.056***	−0.013*	0.040***
	(0.017)	(0.021)	(0.007)	(0.014)
PhD Degree	0.049**	−0.084***	−0.012*	0.047***
	(0.022)	(0.026)	(0.007)	(0.019)
Far Past Coinventors	0.015	−0.019	−0.001	0.002
	(0.023)	(0.029)	(0.007)	(0.017)
Science	0.030***	−0.039***	0.001	0.008***
	(0.004)	(0.004)	(0.001)	(0.003)
Experience	−0.001	0.001	−0.001	0.001
	(0.002)	(0.002)	(0.001)	(0.001)
Country of Degree dummy	−0.035	0.025	−0.014*	0.024
	(0.025)	(0.035)	(0.008)	(0.024)
Mobility In Region	0.045*	−0.060**	−0.001	0.016
	(0.026)	(0.029)	(0.008)	(0.020)
Mobility Out Region	0.025	−0.033*	0.000	0.007
	(0.016)	(0.018)	(0.006)	(0.013)
Experience Herfindahl	−0.029	0.068	0.021**	−0.060***
	(0.033)	(0.042)	(0.010)	(0.023)
Applicant controls				
Log(Employees)	0.000	−0.003	−0.002	0.004
	(0.004)	(0.004)	(0.001)	(0.003)
R&D Intensity	−0.670**	0.838**	−0.070	−0.098
	(0.294)	(0.406)	(0.079)	(0.198)
PRI Applicant	0.010	−0.045	−0.016*	0.051*
	(0.036)	(0.040)	(0.009)	(0.030)
Individual Applicant	−0.039**	0.050**	−0.003	−0.007
	(0.017)	(0.023)	(0.006)	(0.014)
Patent controls				
Number of Inventors	0.000	−0.001	−0.001	0.003
	(0.004)	(0.006)	(0.001)	(0.003)
Commercial Exploitation	0.003	−0.006	−0.001	0.003
	(0.004)	(0.004)	(0.001)	(0.003)
Licensing	0.023***	−0.029***	0.001	0.004
	(0.004)	(0.005)	(0.002)	(0.003)
Prevent Imitation	0.007*	−0.008	0.001	0.000
	(0.004)	(0.005)	(0.001)	(0.003)
Region controls				
Log(GDPPC)	−0.040	0.060*	0.004	−0.024
	(0.027)	(0.036)	(0.008)	(0.020)
Log(Pop)	0.012	−0.008	0.006	−0.010
	(0.014)	(0.017)	(0.005)	(0.010)
Log(Area)	−0.013*	0.012	−0.004	0.004
	(0.008)	(0.010)	(0.002)	(0.006)
Log(Regional Pats)	0.002	−0.002	0.000	0.000
	(0.010)	(0.012)	(0.003)	(0.007)
TOP1% in Technology	0.012	0.009	0.024***	−0.044***
	(0.017)	(0.020)	(0.008)	(0.012)
NR Research Universities	−0.020*	0.021	−0.004	0.003
	(0.012)	(0.015)	(0.004)	(0.008)
Research University Score	0.001	−0.001	0.001	−0.001
	(0.001)	(0.001)	(0.000)	(0.001)

Robust standard errors are in parentheses, adjusted for clusters by firms' identifier. All regressions include dummies for missing values for employees, missing values for R&D intensity, missing values for MOBILITY, year of application, technological field (30 ISI-INPI-OST classes), Country of the inventor, a dummy variable for inventors with no prior patents, and one for inventors with no past coinventors. $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

that even after including this variable, the signs and statistical significance of the education variables hold for the probability to access external and distant spillovers (see table 6). Also, the marginal effect of Conferences emerges as statistically significant for the probability of accessing external spillovers and spillovers from distant sources, whereas the effect of Science declines, which suggests that the variable captures some of the inventors' scientific networking.

We also address whether Phd Degree and University Degree might measure an inventor's individual ability rather than the effect of education. We add a rough measure of ability, according to the number of patents applied for by the inventor, prior to the year of the application for the focal patent (Inventor Past Patents). In the bottom section of table 6, we show that including this variable does not reduce the importance of education in support of its role.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/449/1917213/rest_a_00259.pdf by guest on 28 September 2021

Finally, we explore the effect of regional heterogeneity using regional characteristic variables and country dummies instead of regional dummies. The results confirm the role of education and the idea that a rich regional technological endowment increases the probability of local interactions. We use Regional Patents, or the total number of patents in all sectors (average over 1994–1996) invented in the inventor's NUTS3 region, as a measure of general technological endowment. A dummy variable indicates the strength of the region in the discipline of the invention: TOP1% in Technology equals 1 for the top 1% regions in terms of distribution of the ratio $Tech_{it}/Tech_t$, which is the ratio between the patents invente in region *i* in technology *t* and the total number of patents invented in that technology in all regions.[20] From the Shanghai Academic Ranking of World Universities (ARWU), we gathered information about the number of research universities in the focal NUTS3 region (NR Research Universities) and their ranking, according to a weighted score of quality indicators (Research University Score).[21] Finally, we include country dummies and regional controls for the NUTS3 regional area (square kilometers), population (thousands of people, average 1994–1996), and general economic conditions (per capita gross domestic product in purchasing power parity, corrected for inflation, average 1994–1996). Table 7 shows the marginal effects of the covariates on the bivariate predicted probabilities of Near and Far.

The effect of education and all other variables indicates no relevant changes with respect to the models with regional dummies. Moreover, a location that represents a top region for the technology of the invention increases the predicted probability of OnlyNear Interactions and decreases the use of distant ones. The colocation of research universities and their quality rankings have no impact on the four combinations of outcomes.[22] The results do not change if we exclude Regional Patents or *Top1% in Technology* from the model if we use a dichotomous variable for the presence of research universities, or if we adopt different ranking scores.[23]

[20] Patent data come from the Regio-Eurostast database. The dummy variables are set at the NUTS2 regional level (NUTS3 data are not available) for the years 1994–1996. The Top1% in Technology ratio is technology specific and ranges from 4% to 15%. Results are weaker when we include $Tech_{it}/Tech_t$ directly in the regressions or a dummy for regions in the top 5% of the distribution.

[21] These data are from 2003 (http://www.arwu.org/ARWU2003.jsp). Methodological details are available at http://www.arwu.org/ARWU-Methodology2003.jsp and in Aghion, Dewatripont et al. (2009). We could not find similar data for the years of the surveyed patents, but dramatic changes in the ranking and number of research universities are unlikely in the short run. If more than one university appears in a region, we use the highest ranking. We also gathered data from the Leiden ranking of European universities in 2008 (http://www.cwts.nl/ranking/), but these data produce no relevant changes in the empirical results.

[22] When we reestimate the model separately for the three subsamples of inventors with different education levels, Research University Score is positive and statistically significant for the probability of using local spillovers only for inventors with PhD degrees.

[23] To control for whether the results hold with a lower presence of large companies that typically employ more PhDs than smaller firms, we reestimate the regressions for firms in the bottom 75% and 50% of the distribution in size. The signs and statistical significance of the education variables are similar to those in the full sample.

VI. Conclusion

We use data from a large-scale survey on inventors (PatVal-EU survey) to explore whether individual factors, and in particular the level of education, contribute to a person's capacity to scout for and absorb knowledge spillovers regardless of the geographic location of the spillovers, as well as whether it gives better access to knowledge spillovers from distant sources. We estimate the net effect of education by controlling for several other means that might shape access to knowledge spillovers with varying geographical coverage, such as past collaborations, inventors' mobility, open epistemic communities, the features of the applicant organization, and location effects.

Even after controlling for these factors, the level of education of the inventors benefits their openness when it comes to accessing knowledge spillovers, both in general and with respect to the probability of accessing geographically distant spillovers. This finding does not contradict existing evidence about the importance of location or other mechanisms to absorb knowledge spillovers, but it provides new insights into the drivers of knowledge diffusion. The ability to absorb knowledge spillovers and be "cosmopolitan" in accessing them appears to represent an individual factor, to which a high level of education contributes, because it provides people the ability to recognize the value of, assimilate, and exploit external knowledge, with less need for physical proximity or face-to-face contacts. Reliance on local spillovers offers a better option for inventors who lack the capacity to move beyond their regional setting, therefore reducing the potential pool of knowledge that an inventor can access and use to develop new ideas.

Similar effects of academic training might appear in other professions as well (for example, managers, policymakers). If so, education and other interventions directed at stimulating individual openness might foster the diffusion of knowledge and enlarge the potential pool of information that these people can also absorb and exploit. Our work also suggests that the "death of distance" through modern information and communication technologies, which have diminished obstacles to knowledge interactions, is not applicable for everyone; to make distance "disappear," people need some complementary competencies, which education helps build.

Several contributions indicate a positive correlation between education and the economic performance of firms and countries. Research also notes the direction of causation between the two variables and the mechanisms that explain the transition from education, through spillovers, to economic growth and productivity (see Aghion, Bouston et al., 2009; Moretti, 2004). The microlevel evidence we present might help clarify this link. That is, among other mechanisms, education could contribute to individual cognitive capabilities to scout for and absorb knowledge spillovers, across distance and contexts, which increases the efficiency of the search for knowledge inputs and enlarges

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/449/1917213/rest_a_00259.pdf by guest on 28 September 2021

the pool and expected quality of inputs people receive to develop new ideas.

REFERENCES

Aghion, Philippe, Leah Boustan, Caroline M. Hoxby, and Jerome Vandenbussche, "Exploiting States' Mistakes to Identify the Causal Impact of Higher Education on Growth," unpublished manuscript, Harvard University (2009).

Aghion, Philippe, Mathias Dewatripont, Caroline M. Hoxby, Andreu Mas-Colell, and André Sapir, "The Governance and Performance of Universities: Evidence from Europe and the US," *Economic Policy* 25 (2009), 7–59.

Agrawal, Ajay K., Iain Cockburn, and John McHale, "Gone But Not Forgotten: Labor Flows, Knowledge Spillovers and Enduring Social Capital," *Journal of Economic Geography* 6 (2006), 571–591.

Agrawal, Ajay K., Devesh Kapur, and John McHale, "How Do Spatial and Social Proximity Influence Knowledge Flows? Evidence from Patent Data," *Journal of Urban Economics* 64 (2008), 258–269.

Alcacer, Juan, and Michelle Gittelman, "Patent Citations as a Measure of Knowledge Flows: The Influence of Examiner Citations," this REVIEW 88 (2006), 774–779.

Allen, Thomas. J., *Managing the Flow of Technology: Technology Transfer and the Dissemination of Technological Information within the R&D Organization* (Cambridge, MA: MIT Press, 1977).

Allen, Thomas J., and Steven Cohen, "Information Flow in R&D Laboratories," *Administrative Science Quarterly* 14 (1969), 12–19.

Almeida, Paul, and Bruce Kogut, "Localization of Knowledge and the Mobility of Engineers in Regional Networks," *Management Science* 45 (1999), 905–917.

Audretsch, David B., and Paula E. Stephan, "Company-Scientist Locational Links: The Case of Biotechnology," *American Economic Review* 86 (1996), 641–652.

Bercovitz, Janet, and Maryann P. Feldman, "Academic Entrepreneurs: Organizational Change at the Individual Level," *Organization Science* 19 (2008), 69–89.

Breschi, Stefano, and Francesco Lissoni, "Mobility of Skilled Workers and Co-Invention Networks: An Anatomy of Localized Knowledge Flows," *Journal of Economic Geography* 9 (2009), 439–468.

Cohen, Wesley M., and Daniel A. Levinthal, "Absorptive Capacity: A New Perspective on Learning and Innovation," *Administrative Science Quarterly* 35 (1990), 128–152.

Cole, Stephen, "Age and Scientific Performance," *American Journal of Sociology* 84 (1979), 958–977.

Dasgupta, Partha, and Paul A. David, "Towards a New Economics of Science," *Research Policy* 23 (1994), 487–521.

Ding, Waverly W., Fiona Murray, and Toby E. Stuart, "Gender Differences in Patenting in the Academic Life Sciences," *Science* 313 (2006), 665–667.

Doring, Thomas, and Jan Schnellenbach, "What Do We Know about Geographical Knowledge Spillovers and Regional Growth? A Survey of the Literature," *Regional Studies* 40 (2006), 375–395.

Duguet Emmanuel, and Megan MacGarvie, "How Well Do Patent Citations Measure Flows of Technology? Evidence from French Innovation Surveys," *Economics of Innovation and New Technology* 14 (2005), 375–393.

Feldman, Maryann P., "The New Economics of Innovation, Spillovers and Agglomeration: A Review of Empirical Studies," *Economics of Innovation and New Technology* 8 (1999), 5–25.

Fleming, Lee, Charles King, and Adam Juda, "Small Worlds and Regional Innovation," *Organization Science* 18 (2007), 938–954.

Gittelman, Michelle, "Does Geography Matter for Science-Based Firms? Epistemic Communities and the Geography of Research and Patenting in Biotechnology," *Organization Science* 18 (2007), 724–741.

Gittelman, Michelle, and Bruce Kogut, "Does Good Science Lead to Valuable Knowledge? Biotechnology Firms and the Evolutionary Logic of Citation Patterns," *Management Science* 49 (2003), 366–382.

Giuri, Paola, Myriam Mariani, Stefano Brusoni, Gustavo Crespi, Dominique Francoz, Alfonso Gambardella, Walter Garcia-Fontes, Aldo Geuna, Raul Gonzales, Dietmar Harhoff, Karin Hoisl, et al. "Inventors and Invention Processes in Europe. Results from the PatVal-EU Survey," *Research Policy* 36 (2007), 1107–1127.

Grossman, Gene M., and Elhanan Helpman, *Innovation and Growth in the Global Economy* (Cambridge, MA: MIT Press, 1991).

Haeussler, Carolin, "Information-Sharing in Academia and the Industry: A Comparative Study," *Research Policy* 40 (2011), 105–122.

Hall, Bronwyn H., Jaques Mairesse, and Laure Turner, "Identifying Age, Cohort and Period Effects in Scientific Research Productivity: Discussion and Illustration Using Simulated and Actual Data on French Physicists," *Economics of Innovation and New Technology* 16 (2007), 159–177.

Harhoff, Dietmar, Karin Hoisl, and Colin Webb, "European Patent Citations. How to Count and How to Interpret Them," unpublished manuscript, University of Munich (2008).

Haskell, Robert E., *Transfer of Learning. Cognition, Instruction, and Reasoning* (San Diego, CA: Academic Press, 2001).

Jaffe, Adam B., Manuel Trajtenberg, and Michael S. Fogarty, "Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors," *American Economic Review* 90 (2000), 215–218.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson, "Geographic Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics* 108 (1993), 577–598.

Jones, Benjamin F., "The Burden of Knowledge and the Death of the Renaissance Man: Is Innovation Getting Harder?" *Review of Economic Studies* 76 (2009), 283–317.

Jones, Benjamin F., Stefan Wuchty, and Brian Uzzi, "Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science," *Science* 322 (2008), 1259–1262.

Kerr, William, "Ethnic Scientific Communities and International Technology Diffusion," this REVIEW 90 (2008), 518–537.

Knorr Cetina, Karin, *Epistemic Cultures: How the Sciences Make Knowledge* (Cambridge, MA: Harvard University Press, 1999).

Krugman, Paul R., *Geography and Trade* (Cambridge, MA: MIT Press, 1991).

Lazear, Edward, "Balanced Skills and Entrepreneurship," *American Economic Review* 94 (2004), 208–211.

Levin, Sharon G., and Paula E. Stephan, "Research Productivity over the Life Cycle: Evidence for Academic Scientist," *American Economic Review* 81 (1991), 114–132.

Marshall, Alfred, *Principles of Economics* (New York: Macmillan, 1890).

Merton, Robert K., *The Sociology of Science: Theoretical and Empirical Investigations* (Chicago: University of Chicago Press, 1973).

Miyake, Naomi, and Donal A., Norman, "To Ask a Question, One Must Know Enough to Know What Is Not Known," *Journal of Verbal Learning and Verbal Behavior* 18 (1979), 357–364.

Moretti, Enrico, "Workers' Education, Spillovers and Productivity: Evidence from Plant-Level Production Functions," *American Economic Review* 94 (2004), 656–690.

Romer, Paul M., "Endogenous Technological Change," *Journal of Political Economy* 98 (1990), 71–102.

Saxenian, AnnaLee, *Regional Advantage, Culture and Competition in Silicon Valley and Route* 128 (Cambridge, MA: Harvard University Press, 1994).

Singh, Jasjit, "Collaborative Networks as Determinants of Knowledge Diffusion Patterns," *Management Science* 51 (2005), 756–770.

Singh Jasjit, and Lee Fleming, "Lone Inventors as Sources of Breakthroughs: Myth or Reality?" *Management Science* 56 (2010), 41–56.

Sorenson, Olav, and Jasjit Singh, "Science, Social Networks and Spillovers," *Industry and Innovation* 14 (2007), 219–238.

Sumell, Albert J., Paula E. Stephan, and James D. Adams, "Capturing Knowledge: The Location Decision of New PhDs Working in Industry," in R. B. Freeman, and D. F. Goroff (eds.), *Science and Engineering Careers in the United States: An Analysis of Markets and Employment* (Chicago: University of Chicago Press, 2009).

Thompson, Peter, "Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-Added Citations," this REVIEW 88 (2006), 383–388.

Thompson, Peter, and Melanie Fox-Kean, "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment," *American Economic Review* 95 (2005), 450–460.

Verspagen, Bart, "European Regional Clubs: Do They Exist and Where Are They Heading? On Economic and Technological Differences between European Regions," In J. Adams and F. Pigliaru (eds.), *Economic Growth and Change: National and Regional Patterns of Convergence and Divergence* (Cheltenham: Edward Elgar, 1997).

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/449/1917213/rest_a_00259.pdf by guest on 28 September 2021

Zucker, Lynne G., Michael R. Darby, and Jeff Armstrong, "Geographically Localized Knowledge: Spillovers or Markets?" *Economic Inquiry* 36 (1998), 65–86.

APPENDIX A

A. *The PatValEU Survey: Sample Selection*

We selected the target sample of patents from the population of all EPO-granted patents with priority in 1993 to 1997. We first assigned patents according to the country of the first inventor listed. Patents from France, Germany, Italy, the Netherlands, Spain, and the United Kingdom represented 42.2% of total EPO patents and 88.0% of the EU-15 EPO patents.[24] To address the skewed distribution of "important" patents, we oversampled patents that had been opposed or received at least one citation and attained 15% additional observations at the EU6 level, compared with their initial share (28.5%). Our goal was to obtain 10,000 usable questionnaires: 3,500 from Germany, 1,750 from France, 1,750 from the United Kingdom, 1,250 from Italy, 1,250 from the Netherlands, and 500 from Spain. The country share of patents in the population and the country response rate in the pilot surveys determined the number of questionnaires submitted in each country. In the end, we used a stratified sample of 27,531 EPO patents, composed of all opposed or cited patents from 1993 to 1997, as well as a random sample of uncited and unopposed patents from the same period. The interviews began in May 2003 and ended in January 2004.

B. *Family Issues*

In patent data, groups of patents might be part of the same "family." Our patents date back to a period in which filing multiple related patents was not a common practice, but we confirmed this likelihood with two tests. We investigated whether the publication numbers in our sample are listed in TLS205 (Technical Relationships), a database built by EPO examiners to find patent filings that, though unrelated by priority, actually are closely connected. We found no matches.[25] In addition, we checked for equivalents, that is, whether an applicant produced multiple equivalent patents from one priority. We found 28 $\times$ 2 cases of this sort. By manually checking the extent of overlap in these 28 cases (Gittelman & Kogut, 2003), we found that 14 $\times$ 2 represent potential family concerns, because they have same inventors, locations, and main IPC class.

C. *Inventors' Past Patents*

We collected patents that PatVal-EU inventors contributed to in the ten years before the focal patent was conducted, in collaboration with Karin Hoisl (University of Munich). The EPOLINE patent database of the EPO covers more than 1.26 million patent files with application dates from 1978. The search relied on matching criteria (inventor's last name, inventor's first name or part of it, street or city address (or both), IPC code, name of the applicant), which we combined into 38 subsets, each with three or four criteria. The procedure matched the PatVal-EU patent information with data displayed in the EPOLINE patents. The queries used MYSQL version 4. The MYSQL control center was applied as SQL-Interface. All Java classes were constructed with ECLIPSE. The search resulted in 38 data sets containing potential matches, each with an expected match quality score, assigned according to the specific subset of criteria that they met. We merged the 38 data sets in one master database and checked the records manually to remove duplicate patent applications and incorrect matches. The final data set consists of 64,000 EPO patents that the PatVal-EU inventors contributed to invent.

APPENDIX B

ISI-INPI-OST Technological Classes Used in the Study and Descriptive Statistics.

	Mean	SD
Electrical devices, engineering, energy	0.078	0.268
Audiovisual technology	0.021	0.143
Telecommunications	0.034	0.181
Information technology	0.023	0.150
Semiconductors	0.011	0.105
Optics	0.021	0.143
Analysis, measurement, control technology	0.059	0.235
Medical technology	0.022	0.147
Organic fine chemistry	0.066	0.248
Macromolecular chemistry, polymers	0.060	0.237
Pharmaceuticals, cosmetics	0.016	0.124
Biotechnology	0.007	0.085
Materials, metallurgy	0.033	0.178
Agriculture, food chemistry	0.013	0.111
Chemical and petroleum, basic materials chemicals	0.038	0.191
Chemical engineering	0.029	0.167
Surface technology, coating	0.016	0.124
Materials processing, textiles, paper	0.055	0.227
Thermal processes and apparatus	0.022	0.146
Environmental technology	0.018	0.133
Machine tools	0.036	0.185
Engines, pumps, turbines	0.031	0.174
Mechanical elements	0.042	0.201
Handling, printing	0.076	0.265
Agricultural and food processing, machine apparatus	0.019	0.135
Transport	0.068	0.252
Nuclear engineering	0.003	0.056
Space technology weapons	0.004	0.065
Consumer goods and equipment	0.046	0.209
Civil engineering, building, mining	0.036	0.187

N = 6,051.

[24] We undersampled the share of German and French patents and oversampled patents from other countries to obtain sufficiently large samples for all countries.

[25] We thank Dietmar Harhoff for helping us address the family issue and for performing the tests.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/449/1917213/rest_a_00259.pdf by guest on 28 September 2021