



ELSEVIER

Contents lists available at ScienceDirectInternational Journal of Industrial
Organizationwww.elsevier.com/locate/ijioRecombinant innovation and the boundaries of the
firmRachel Griffith^{a,*}, Sokbae Lee^b, Bas Straathof^c^a *University of Manchester and Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE United Kingdom*^b *Columbia University and Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE United Kingdom*^c *CPB Netherlands Bureau for Economic Policy Analysis, Centraal Planbureau, van Stolkweg 14, 2585 JR Den Haag The Netherlands*

ARTICLE INFO

Article history:

Received 24 November 2015

Revised 26 October 2016

Accepted 31 October 2016

Available online 16 November 2016

JEL Classification:

F23

O32

O33

Keywords:

Recombinant innovation

Multinational firms

Patent policy

ABSTRACT

Recombinant innovation, the combination of existing ideas, is important for technological progress; we want to understand how important market frictions are in stifling the transmission of ideas from one firm to another. Although the theoretical literature emphasizes the importance of these frictions, direct empirical evidence on them is limited. We use comprehensive data on patent applications from the European Patent Office and a multiple spells duration model to provide estimates that suggest that they are substantial. It is around 30% more costly to successfully discover and utilize new ideas created in another firm than in your own. This compares to the increased costs of accessing new ideas across national borders of around 7%, and across technologies of around 20%. These result point towards substantial imperfections in the market for technology.

© 2016 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY license

[\(http://creativecommons.org/licenses/by/4.0/\)](http://creativecommons.org/licenses/by/4.0/).

* Corresponding author.

E-mail addresses: rgriffith@ifs.org.uk (R. Griffith), sl3841@columbia.edu (S. Lee), s.m.straathof@cpb.nl (B. Straathof).<http://dx.doi.org/10.1016/j.ijindorg.2016.10.005>0167-7187/© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Technological progress is reliant on recombinant innovation - the recombination of existing ideas, where firms combine prior art with their own ideas to form new patentable technologies - now more than ever (Henderson and Clark, 1990; Henderson and Cockburn, 1990; Weitzman, 1998; Wuchty et al., 2007 and Jones, 2009). There is an extensive theoretical literature that details the ways by which market frictions, such as incomplete contracts, transactions costs and intellectual property rights, might stifle the transmission and use of ideas across firms, leading to less research than is socially desirable (Scotchmer, 1991; Green and Scotchmer, 1995; Bessen, 2004; Boldrin and Levine, 2013).

Our contribution in this paper is to develop an empirically tractable model of search and innovation that we can take to data to learn about the size of these market frictions. We estimate the increase in effort that is required to successfully develop a new technology when it relies on transactions in the market, as opposed to internal to the firm. We use data on all patent applications made at the European patent office, matched to firm structures, to consider the length of time it takes a firm to build on an idea published in a patent application made by a different firm, relative to an idea published in a patent application made within the same firm.

In our data we see that on average it takes around 200 days longer for the first citation of a patent application to occur when that citation is by a different firm than when it is by the same firm, see Fig. 1 (data described below). Of course this could be due to a large number of potentially confounding factors; we use a multiple-spell duration model (Griffith et al., 2011) to try to control for these. Our empirical approach controls

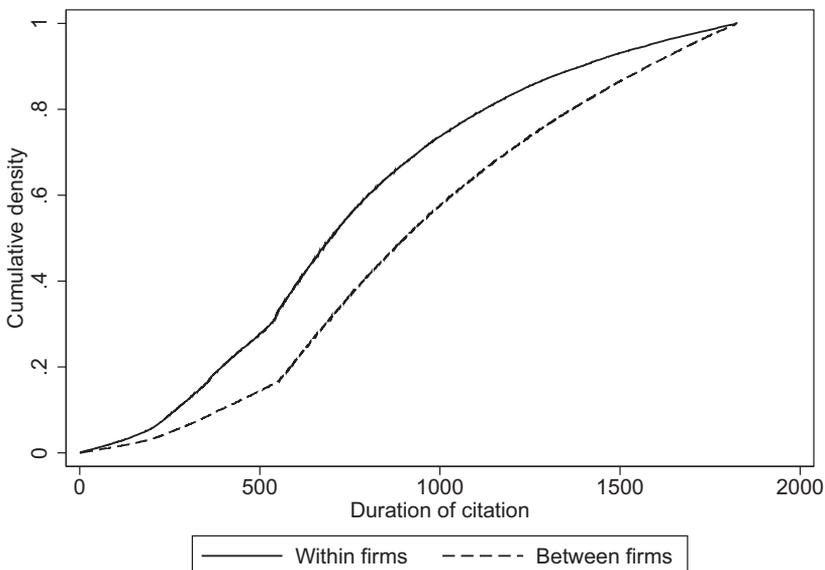


Fig. 1. Cumulative density of time to first citation in days if cited within 5 years.

for characteristics of the cited and citing patent applications; our identification strategy relies on variation between the time it takes for the first and second citation. Our central estimate suggests that successful discovery and use of information that leads to a new idea is around 30% more costly cross-firm than within-firm.

We compare this to the costs of search across geographic space, to which the empirical literature has paid a lot of attention, and to the costs of search across technological space. This compares to a 7% increase in costs when it is across national borders rather than within-country, and around 20% more costly across (4-digit) technology classes. These results suggest that the empirical magnitude of market frictions are substantial, and we interpret our evidence as pointing towards substantial imperfections in the market for technology.

Our paper is related to several relevant literatures. The patent system grants a firm monopoly rights over an idea for a fixed period, with the aim to provide monopoly rents to improve incentive to innovate (Schumpeter, 1943; Nordhaus, 1969 and Tirole, 1988), because there will be under-investment due to the public goods nature of ideas (Nelson, 1959; Arrow, 1962; Shapiro, 2007). Some argue that patents facilitate technology transfer (Acemoglu et al., 2008), while others argue that they create a barrier to technology transfer (Bessen and Maskin, 2009). When new inventions build on old ones it is not clear what impact intellectual property rights such as patents will have on total innovation (Scotchmer, 1991; Gallini and Scotchmer, 2002; Boldrin and Levine, 2008, and Boldrin and Levine, 2013).

Weitzman (1998), Wuchty et al. (2007) and Jones (2009) emphasise that recombinant innovation is growing in relevance due to the growing burden of knowledge, which requires greater investments in human capital and team work and a greater reliance on reusing existing ideas in new ways. van den Bergh (2008) and Zeppini and van den Bergh (2013) provide a model of recombinant innovation and the role it plays in firms strategies over product diversity.

A separate literature emphasises that firms have incentives to make it costly for other firms to build on their technology and effectively develop new recombinant technologies (Chang, 1995; Arundel, 2001 and Horstmann et al., 2004). Williams (2013) provides empirical evidence on the impact of intellectual property rights in stifling technology transfer by looking at one technology; she provides evidence that patents on the human genome hindered subsequent scientific research and product development.

Also related to our work is the large literature that has considered the nature and boundaries of the firm, and when technology transactions will occur in the market or within the firm (Coase, 1937; Williamson, 1971, 1975, 1979; Arrow, 1962; Mansfield and Romeo, 1980; Davidson and McFetridge, 1984; Grossman and Hart, 1986 and more recently Bessen (2004), Costinot et al. (2011), Ponzetto (2012) and Irazabal et al., 2013). In a recent paper, Alfaro et al. (2015) look at the empirical evidence regarding firm boundary choices along the global value chain and find that contractual frictions play an important role in firms' choices on integration.

The paper proceeds as follows. In the next section we lay out a simple model of search and discovery of ideas for recombination to create new patentable technologies. In [Section 3](#) we lay out our econometric approach to estimating the cost of search, discovery and use. In [Section 4](#) we describe the data, and in [Section 5](#) present our estimates. A final section summarizes and provides some final comments.

2. Theoretical framework

We develop a model in which the time and effort it takes for inventors to develop and exploit a new idea depends on the efficiency with which they search and use prior results, and the expected value of the recombination when discovered. Research projects do not typically start from scratch, but are based on earlier findings - what is termed recombinant innovation. Researchers devote time talking to other researchers, read academic journals, visit fairs and conferences, reverse-engineer a competitor's product, and search patent databases. The more effort researchers puts into finding and accessing these activities, the more likely it is that they develop a useful new technology. The model is stark in that it focuses only on recombinant innovation, and in that sense is similar in spirit to [Cohen and Levinthal \(1989\)](#).

In our model a researcher maximizes expected profits by searching for a recombination of existing technology with new ideas. The researcher decides how much effort to spend on searching for recombination of his own knowledge with knowledge of other researchers and on subsequently developing that knowledge. A successful recombination of technologies leads to a new innovation.

The expected value of the new innovation, V_{new} , depends on the value of the technology it is built on, V_{old} , and information available at time zero, Ω_0 , for example, information on future technological and demand opportunities. This can be written as

$$v \equiv E(V_{\text{new}}|\Omega_0) = V_{\text{old}} + E(\varepsilon|\Omega_0), \quad (1)$$

where ε is the difference between the values of the new and old technologies, since V_{old} is known at time zero. What matters in a researcher's search decision is the expected value of the new technology, not the realized value of it. The expected value v consists of V_{old} and $E(\varepsilon|\Omega_0)$. In the next section we show how we can control for v , which is unobserved by the econometrician, in our econometric framework.

For simplicity we start by assuming that the researcher only has one direction in which to search for technology. We then introduce a second direction for which efficiency of the search process and the value of recombinant innovation may differ. We assume that effort takes the form of expenditure by the researcher. In particular, the researcher has to spend $e_t > 0$ for all t to continue the project.

Let T be the moment when a recombination of technologies is discovered by the researcher. T is a stochastic variable with marginal density function $f(t)$ and survival function $S(t)$. The expected revenue from searching for a short time interval $[t, t + dt)$

equals the value of the innovation times the probability of a recombination during that interval, given that the recombination has not occurred at an earlier time:

$$\Pr(t \leq T < t + dt | T \geq t)v.$$

The search is stopped as soon as a suitable technology is found. The continuation cost during $[t, t + dt)$ equals $e_t dt$ if the technology has not been found before $t + dt$ and zero otherwise:

$$\Pr(T \geq t + dt | T \geq t)e_t dt.$$

The expected discounted profits π at time zero can be expressed as the integral over t of expected increments in profits:

$$E(\pi) = \int_0^\infty \exp(-\rho t) E(d\pi_t).$$

The value of the outside option is captured by the discount rate ρ . An expected increment in profits equals the expected increment in revenue minus the expected increment in cost:

$$E(d\pi_t) = \Pr(t \leq T < t + dt | T \geq t)v - \Pr(T \geq t + dt | T \geq t)e_t dt.$$

After some manipulation, the profit increments can be written as a function of the hazard $\lambda(t)$ that a recombination takes place at t . Specifically,

$$\begin{aligned} E(d\pi_t) &= \frac{\Pr(t \leq T < t + dt)v - \Pr(T \geq t + dt)e_t dt}{\Pr(T \geq t)} \\ &= \frac{f(t)v dt - S(t + dt)e_t dt}{S(t)} \\ &= \lambda(t)v dt - [1 - \lambda(t)]e_t dt \\ &= [\lambda(t)(v + e_t) - e_t]dt. \end{aligned}$$

The researcher can influence the hazard rate by varying effort at any point in time. If the response of the hazard rate to effort is instantaneous, we can write the current hazard as a function of the current effort. In particular, assume that the hazard $\lambda(t|e_t)$ is concave and monotonically increasing in e_t . The profit maximization problem for the researcher is dynamic, but does not involve a state variable, as the hazard rate is not affected by previous decisions made by the researcher.¹ The researcher maximizes expected profits by choosing an expenditure path:

$$\max_{e_t} \int_0^\infty \exp(-\rho t) [\lambda(t|e_t)(v + e_t) - e_t] dt.$$

¹ The only possible effect of previous decisions is that the search was successful at an earlier moment. If a technology is found, a new search can be started and optimal expenditure will be the same as if no technology had been found.

We now introduce a second direction for which the efficiency of searching may differ, and we allow the value of an invention to be different across directions. Let the expected value of recombination of the researcher’s knowledge and a technology found in direction i (“technology i ”) be v_i and let the expected value of recombination with a technology j be v_j . Expenditure on search in one direction does not affect the hazard rate in the second direction for given expenditure in the latter direction.² Once a suitable technology has been found in either direction, searching in both directions is stopped.

In the two-direction case the expected flow of profits depends on the expenditure on searching in both directions:

$$\frac{E(d\pi_t)}{dt} = \lambda_i(t|e_{it})v_i + \lambda_j(t|e_{jt})v_j - [1 - \lambda_i(t|e_{it}) - \lambda_j(t|e_{jt})](e_{it} + e_{jt}). \tag{2}$$

The first-order conditions for maximizing profits yield the condition that the ratio of marginal hazards should equal the inverse ratio of the values. That is, optimal expenditure paths e_{it}^* and e_{jt}^* satisfy

$$\frac{\partial \lambda_i(t|e_{it}^*)/\partial e}{\partial \lambda_j(t|e_{jt}^*)/\partial e} = \frac{v_j + e_{it}^* + e_{jt}^*}{v_i + e_{it}^* + e_{jt}^*}. \tag{3}$$

Eq. (3) is a key equation to characterize the optimal expenditures. The following propositions consider two polar cases.

Proposition 1. *Suppose that search efficiency is the same between two technologies: $\lambda_i(t|e) = \lambda_j(t|e)$ for all t and e . If $v_i \geq v_j$, then $e_{it}^* \geq e_{jt}^*$.*

Proof. Let $\lambda(t|e)$ denote the common hazard function, that is $\lambda(t|e) = \lambda_i(t|e) = \lambda_j(t|e)$ for all t and e . If $v_i \geq v_j$, it follows from (3) that

$$\frac{\partial \lambda(t|e_{it}^*)}{\partial e} \leq \frac{\partial \lambda(t|e_{jt}^*)}{\partial e}.$$

As the hazard function is a concave function of effort, $e_{it}^* \geq e_{jt}^*$. \square

Proposition 1 says that when the expected value of technology i rises relative to the expected value of technology j , then the hazard rate of finding technology i increases relative to the hazard rate of j , provided search efficiencies are symmetric. In other words, it predicts that the *ex ante* more usefulness of recombination will lead to faster patent citations.

² If the researcher faced a budget constraint, an increase in search efficiency for one direction could lead to reduction of effort in the other direction.

Proposition 2. *Suppose that values of recombination are the same between two technologies: $v_i = v_j$. If*

$$\frac{\partial \lambda_i(t|e)}{\partial e} \geq \frac{\partial \lambda_j(t|e)}{\partial e} \text{ for all } e, \quad (4)$$

then $e_{it}^* \geq e_{jt}^*$.

Proof. Since $v_i = v_j$, it follows from (3) that

$$\frac{\partial \lambda_i(t|e_{it}^*)}{\partial e} = \frac{\partial \lambda_j(t|e_{jt}^*)}{\partial e}.$$

Hence, if (4) holds, then it must be that $e_{it}^* \geq e_{jt}^*$. \square

Proposition 2 says that when the expected value of recombination is the same for both directions, then higher search efficiency (expressed as a larger marginal increase in the hazard rate with respect to the expenditure) will result in faster patent citations.

The two propositions together imply that any difference in hazard rates between the two directions can be due to differences in search efficiency, differences in the expected value of technology, or both. In particular, Propositions 1 and 2 suggest that if we want to investigate the effect of barriers to knowledge diffusion we need to condition on expected values of recombination.

The expected values of recombination are intrinsically difficult to observe. Recall that $v_j = V_{\text{old}, j} + E(\varepsilon_j|\Omega_0)$. In other words, $E(\varepsilon_j|\Omega_0)$ represents the *ex ante* average difference between values of new and old technologies in direction j . Suppose that the second term $E(\varepsilon_j|\Omega_0)$ can be well approximated by observed characteristics of the citing patent (that patent that makes the citation). This is a reasonable first-order approximation since the *ex ante* average difference is determined at time zero. Therefore, we will use an estimation method that is robust to *unobserved* quality differences in the cited patents $V_{\text{old}, j}$, while controlling for $E(\varepsilon_j|\Omega_0)$ by including characteristics of the citing patent. Once we control for the expected value of recombination, then we can isolate differences in (marginal) search efficiency $(\partial \lambda_i(t|e)/\partial e)$. In this paper, our focus is whether and to what extent firm boundaries are a barrier to the diffusion of knowledge. Since other factors such as geographic proximity are likely to affect search costs, our estimation strategy controls for these other factors as additional explanatory variables.

3. Econometric approach

We are interested in investigating how firm boundaries affect the cost of searching for a new recombination. To isolate this effect we use a multiple-spell duration model similar to that of Griffith et al. (2011).

Cross-firm citations are likely to differ from within-firm citations for a number of reasons. Patents belonging to the same firm are more likely to have common inventors,

draw from inventions originating in the same geographical area, and to have similar technological characteristics. All these factors can potentially explain why cross-firm citations are slower. We control for cross-inventor citations, cross-border citations, and cross-technology citations. If a substantial part of the cross-firm delay remains unexplained after controlling for these factors, we interpret this as suggesting that market frictions affect the citation duration.

A fourth reason why citations between firms might differ from citations within firms, is that patents cited between firms are more valuable than patents cited within the firm. Inventions are more likely to lead to follow-up research when they are valuable, such that valuable inventions have larger probability of being cited. If, in addition, follow-up research is more profitable if it is based on a patent owned by the same firm, then the expected value of cited patents will be smaller for within-firm citations than for cross-firm citation. Ignoring heterogeneity in the value of cited patents could lead to underestimation of the cross-firm citation delay. It is therefore crucial that we control for (unobserved) characteristics of cited patents.

We follow the literature that uses patent citation information as a direct measure of the recombination of knowledge. The citation of one patent application by another strongly suggests that the first patent application contained useful knowledge which helped the second innovation.³ As in [Griffith et al. \(2011\)](#), we consider the intensity with which a patent is cited, and adopt a duration modelling framework that explicitly deals with the problem of unobserved patent characteristics that may be correlated with location or other characteristics.

Consider a patent application that will be cited by other patent applications. The time it takes for a patent to be cited is expected to be shorter if efficiency of searching is higher. This is especially relevant for the first few citations a patent receives as they are more likely to reflect newly acquired knowledge.

As emphasised in [Section 2](#), firms will also search harder if the expected value of finding new recombination is larger. The *ex post* value of a recombination is not known prior to invention; we assume that the expectation that researchers have about the value of a recombination is a function of the value of the patent application on which they intend to build. High-quality patents are more likely to lead to a new recombination than patents with a low quality. We treat the quality of cited patents as an unobserved (to us) characteristic. We estimate how the hazard of being cited differs for citations within firms and citations between firms. We control for unobserved characteristics of the cited patent by using a method that is analogous to first-differencing a linear model - we compare citations that are adjacent in time for each cited patent (in particular, the first and second citations).

Let citing patents be indexed by k and let cited patents be indexed by ℓ . A citation of ℓ by k has a duration $T_{\ell k}$, which is defined as the number of days between the application

³ A classic paper in this literature is [Jaffe et al. \(1993\)](#). Also, see the monograph by [Jaffe and Trajtenberg \(2002\)](#) and other recent work ([Thompson and Fox Kean, 2005](#); [Thompson, 2006](#)).

date of the cited patent and the application date of the citing patent. Let $X_{\ell k}$ be a vector of other observed attributes of citation $T_{\ell k}$ (including attributes of k that do not vary with ℓ) and let V_ℓ be the unobserved characteristics of the cited patent.

We model the hazard λ that patent ℓ is cited by k after $t_{\ell k}$ days, conditional on $X_{\ell k} = x_{\ell k}$ and $V_\ell = v_\ell$, as

$$\lambda(t_{\ell k}|x_{\ell k}, v_\ell) = \theta_\ell(t_{\ell k}|v_\ell) \exp(x'_{\ell k}\beta), \quad (5)$$

with $\theta_\ell(\cdot)$ being a cited-patent specific baseline hazard function and β a vector of coefficients to be estimated. The empirical specification in (5) is motivated by the theoretical framework in Section 2, taking the particular shape of a mixed proportional hazard model. We aim to estimate effects of search costs using characteristics of citing patents ($X_{\ell k}$), while controlling for differences in values of cited patents with fixed effects V_ℓ .

A patent can be cited by more than one other patent, so there can be multiple spells for a cited patent. We allow the attributes $X_{\ell k}$ and the unobserved characteristics V_ℓ to be correlated arbitrarily, but are assumed to be constant over time. We assume that V_ℓ is constant across citing firms, implying that a technology has the same *ex ante* value to everyone and is not more valuable to one firm than another. The baseline hazard function $\theta_\ell(\cdot)$ depends on the unobserved characteristics of the cited patent and is left unspecified. That is, the baseline hazard function is allowed to differ across cited patents in a general, unspecified way. The attributes $X_{\ell k}$ include 0–1 indicator variables for, amongst others, between-firm citations, cross-border citations, and cross-technology citations. The hazard rate will be smaller for citations that are more costly.

We can estimate the coefficients β without knowing the cited-patent specific baseline hazard by exploiting observations on cited patents that receive multiple citations. Assuming that $X_{\ell k}$ and $X_{\ell k'}$ ($k \neq k'$) are independent of each other conditional on $X_{\ell k}$, $X_{\ell k'}$, and V_ℓ , we can follow the conditional likelihood approach of Griffith et al. (2011).⁴ The intuition behind this approach is the following. Suppose a patent receives two citations. The probability that the observed first citation is first conditional on the duration of the first citation and conditional on the characteristics of both citations, is independent of the baseline hazard:

$$\begin{aligned} & \Pr [T_{\ell 1} \leq T_{\ell 2} | T_{\ell 1} = t_{\ell 1}, X_{\ell 1} = x_{\ell 1}, X_{\ell 2} = x_{\ell 2}, V_\ell = v_\ell] \\ &= \frac{\theta_\ell(t_{\ell 1}|v_\ell) \exp(x'_{\ell 1}\beta)}{\theta_\ell(t_{\ell 1}|v_\ell) \exp(x'_{\ell 1}\beta) + \theta_\ell(t_{\ell 2}|v_\ell) \exp(x'_{\ell 2}\beta)} \\ &= \frac{\exp(x'_{\ell 1}\beta)}{\exp(x'_{\ell 1}\beta) + \exp(x'_{\ell 2}\beta)}, \end{aligned} \quad (6)$$

which does not depend on v_ℓ or $\theta_\ell(\cdot)$. For patents that are cited twice or more it is possible to estimate the coefficients β without the incidental parameters problem.

⁴ See, also, Chamberlain (1985), Ridder and Tunali (1999), Horowitz and Lee (2004) and Lee (2008), among others.

There remains a problem of censoring. We need to observe at least two citations to implement this estimator, but for a large number of patents we do not (yet) observe two citations. The exclusion of these censored observations can lead to sample selection bias in the estimation results. For example, a patent application in 1999 is more likely to have received two citations than a patent application in 2004. Older patents are therefore more likely to be included in our sample than young patents. We correct for bias due to censoring in two stages. First, we exclude all patents that did not receive at least one citation within five years of the application date. We view these as a different type of patents - either low quality, or for other reasons not relevant for recombinant innovation. We restrict our analysis to the population of patents that receive at least one citation within five years. Second, to correct for those patents for which we do not observe a second citation we weight all observations with the inverse censoring probability, as in [Griffith et al. \(2011\)](#). This is valid if the censoring is independent of citation durations and covariates. See the appendix of [Griffith et al. \(2011\)](#) for further details of the conditional maximum likelihood estimation method we have adopted.

4. Data

We use data on European patent applications made between 1985 and 2004 from the European Patent Office's (EPO's) Worldwide Statistical Patent Database (PATSTAT). The number of days between the priority date of a patent application, which is the application date of the first patent application (anywhere in the world) for that invention, and the priority date of patents that are cited on that application provides information on the time it takes for one invention to lead to another.

Any legal person can apply for a patent, such that the applicant(s) listed on a patent application can be subsidiaries of a firm, parent companies or natural persons. The legal person(s) listed as applicants might not be the ultimate owners of the patent. For our purpose we are interested in the ultimate owner, as this is what is relevant for considering the search costs and market frictions; therefore, we need information on the ownership structure of firms. Without this information, a subsidiary citing a patent from another subsidiary belonging to the same parent firm would be incorrectly identified as an inter-firm citation.

We use commercially available data on ownership structure for European firms (Amadeus) and US firms with European subsidiaries (Icarus) from Bureau Van Dijk (BVD). Names of corporate applicants in PATSTAT have been matched to BVD company names in order to identify the parent firms of applicants. The matching procedure as well as the resulting database are described in [Abramovsky et al. \(2008\)](#). Applicant names have been matched to company names for 15 European countries, four of which have been dropped because the number of patent applications from these countries is small.⁵ The remaining eleven European countries are: Belgium, Denmark, Finland, France, Germany,

⁵ The following countries are dropped: Czech Republic, Greece, Poland and Portugal.

Table 1

Number of patent applications by country and industry of parent firm.

	BE, NL	DE	FR	GB	IT, ES	SCA	US	Total
Chemical	11,216	42,099	11,878	8101	5756	4816	68,696	152,562
Electric	6704	30,914	13,198	6043	5715	4267	41,700	108,541
Engineer	7887	52,164	18,365	10,145	12,126	9967	47,338	157,992
ICT	10,645	18,744	11,623	5624	3580	7164	65,035	122,415
Pharma	5455	11,051	5458	6617	3234	3860	37,639	73314
Other	1837	6893	2404	1917	1203	1335	12,452	28041
Total	43,744	161,865	62,926	38,447	31,614	31,409	272,860	642,865

Notes: Number of applications for a European Patent between 1985 and 1999. Excludes applications by non-corporate applicants (individuals, universities, etc.). Includes patents that are not cited. The row “Other” refers to applications without a Derwent Code. The country or country group of an application refers to the location of the parent company of the applicant. “BE” is Belgium, “NL” is Netherlands, “DE” is Germany, “FR” is France, “GB” is Great Britain, “IT” is Italy, “ES” is Spain, and “SCA” comprises Denmark, Finland, Norway, and Sweden.

Italy, Netherlands, Norway, Spain, Sweden and UK. We use information on parent firms in these European countries plus the US and EPO patent applications made by their subsidiaries that are located in these countries. The sample period of our data spans from 1985 through 2004 and we focus on applications for a European patent between 1985 and 1999 to have a five-year window of forward citations.

We identify the industries in which the technology developed in each patent is applied using the Derwent Innovation Index Manual Codes. [Abramovsky et al. \(2008\)](#) provide a comparison of Derwent Codes with NACE and International Patent Classification (IPC) codes. [Table 1](#) shows the number of cited patents by industry and country of the parent firm.⁶ Engineering, chemical, and ICT are the industries with the largest number of patent applications. The pharmaceutical industry has the smallest number of applications. The United States has by far the most patent applications, followed by Germany. The three largest cells in the table are ICT and chemical in the US and engineering in Germany.

Identification of the cited-patent fixed effects requires that we observe at least two citations per patent. [Table 2](#) displays summary statistics for our data. The first three rows are proportions of all patent applications made by corporations. More than eighty percent of all patent applications since 1985 were not cited within five years. In pharmaceuticals this proportion is even higher. Nine percent of all applications received exactly one citation, another nine percent of all applications received two citations or more. Patent citations are censored at December 31, 2004. The patents that did not receive a citation prior to that date may still receive a citation at a future date. As this applies *a fortiori* for patents with a priority date approaching the end of 2004, citations to these young patents are likely to be underrepresented compared to citations of older patents. We take censoring into account when estimating ([Section 3](#)).

⁶ The industries are aggregates of Derwent Sections: “Chemical” includes the sections A, C, E, F, G, H, J, K, L, M, N; “Electric” includes S, V, X; “Engineer” includes P, Q; “ICT” includes W, T, U; “Pharma” includes B, D.

Table 2
Descriptive statistics by industry.

Chemical		Electric	Engineer	ICT	Pharma	Other	Total
<i>Proportion of all patents (1985–1999)</i>							
with ≥ 2 citations	0.102	0.080	0.070	0.095	0.081	0.097	0.086
with 1 citation	0.093	0.087	0.083	0.090	0.062	0.084	0.085
without citations	0.805	0.833	0.848	0.815	0.857	0.819	0.829
<i>Proportion of second citations</i>							
Cross-firm	0.773	0.839	0.848	0.851	0.764	0.798	0.812
Cross-inventor	0.939	0.971	0.966	0.974	0.936	0.959	0.958
Cross-border	0.393	0.478	0.457	0.470	0.378	0.428	0.435
Cross-tech. (3 digit)	0.080	0.152	0.806	0.090	0.055	N.A.	0.231
Cross-tech. (4 digit)	0.125	0.211	0.825	0.158	0.107	N.A.	0.276

Notes: Data consist of corporate applications for a European Patent between 1985 and 1999. The first three rows are proportions of all patent applications in the sample; the other rows are proportions of all second citations of patent applications that have received their first citation within five years. Citations may occur between 1985 and 2004. “Cross-firm” requires that the cited patent application and the citing patent application have different ultimate owners. “Cross-border” requires that none of the inventors on the cited patent application is located in the same country as any of the inventors on the citing patent. “Cross tech.” refers to a match between the Derwent Code of the cited and citing patent application (match at three- and four-digit level; a patent can be assigned several Derwent Codes).

The remaining rows of [Table 2](#) show the proportion of second citations for our main variables of interest for patents that have received their first citation within five years. A citation is classified as a cross-firm citation if none of the applicants on the cited patent belong to the same parent firm as any of the applicants on the citing patent. A citation is considered to be a cross-inventor citation if none of the inventors on the citing patent are also listed as inventors on the cited patent. A cross-border citation requires that none of the inventors on the cited patent are located in the same country as any of the inventors on the citing patent. A cross-technology citation means that none of the Derwent codes assigned to the cited patent match any of the Derwent codes assigned to the citing patent. We consider both three- and four-digit aggregation levels. Derwent manual codes, which are published by Thomson, classify patents according to the industries in which they are used. We use Derwent codes rather than the standard International Patent Classification (IPC) codes, as the IPC identifies patents according to their technological similarity and not their economic relatedness. Using Derwent codes we can construct a measure of technological distance that better reflects the perspective of inventors than examiners. About 81 percent of citations are across firms, 96% are across inventors, 44% are across countries, and 23 (28)% of citations are across technology at the three-digit (four-digit) level.

Stylized facts suggest that differences in ownership are associated with substantial delays in the recombination of knowledge. The top panel of [Table 3](#) shows the time it takes for a patent to be cited for the first time. It shows the number of days between the priority date of the cited patent and the priority date of the citing patent. The first line has the average number of days for citations belonging to different firms and citations of patents belonging to the same firm. The average duration of citations between firms is

Table 3

Time to first and second citation in days if cited within 5 years.

	Between firms	Within firms	Difference	
	(days)	(days)	(days)	(% of within firm)
All first citations	937	736	200	27
<i>Industry</i>				
Chemical	903	705	198	28
Electric	967	798	168	21
Engineer	981	806	175	22
ICT	953	745	207	28
Pharma	810	625	185	30
Other	937	728	209	29
All second citations	1734	1179	555	47
<i>Industry</i>				
Chemical	1703	1140	563	49
Electric	1819	1283	536	42
Engineer	1909	1366	543	40
ICT	1630	1165	466	40
Pharma	1516	957	559	58
Other	1782	1190	591	50

Notes: The data in the table are the average number of days between the priority date of the cited patent and the priority date of the patent that contains the first (second) citation if the first citation occurs within 5 years. Industry “Other” refers to patents that were not assigned a Derwent code.

937 days, while the average duration of citations within firms is 736 days. Cross-firm citations are on average 27% slower than within-firm citations. The remaining part of the top panel of [Table 3](#) displays the time to first citation for five industries. Citations take the least time in the pharmaceutical industry and take the most time in engineering. This holds both for citations between firms and citations within firms. The relative lag of citations between firm is smallest in electrical engineering and longest in pharmaceuticals. The bottom panel of [Table 3](#) shows the time it takes for a patent to be cited for the second time. Cross-firm citations are on average 47% slower than within-firm citations. Citation patterns across industries are similar between the first and second citation.

That citations between firms are slower than citations within firms is not just a property of the mean citation time, but applies to the distribution of citation times as well. Recall that [Fig. 1](#) shows the cumulative density of citation time for between-firm citations and within-firm citations.

The number of observations available for our analysis might appear to be small in comparison with studies on patent application at the United States Patent and Trademark Office (USPTO). This is because of a number of differences in the patent application and examination process in the EPO and the USPTO.

One main reason is that there is more innovation in the US than in Europe, and the EPO is a younger organization than the USPTO. In addition, there are some institutional reasons. The novelty requirements of the EPO are more strict than those of the USPTO. The EPO does not grant a patent when it is based on an idea that has been described

before, or used for the same purpose before *anywhere in the world*. The USPTO requires that the invention was not known by others *in the United States*, and that the invention was not patented or described in a printed publication anywhere in the world.⁷ As a consequence the number of patents granted by the USPTO is larger than the number of patents granted by the EPO.⁸

In addition, EPO applicants are not subject to a ‘duty of candor’ (Alcácer et al., 2008). The incentive for applicants to cite other patents is therefore much smaller for EPO patents than it is for USPTO patents. Not only does this imply that European Patents on average receive fewer citations than their US counterparts, it also means that the proportion of EPO citations added by examiners is larger (Criscuolo and Verspagen, 2008). For this reason our main results are based on both inventor-added and examiner-added citations. Thompson (2006) shows that citations added by inventors are 20% more likely to match the country of the cited patent than examiner-added citations. This could lead us to underestimated border effects.

Tables 4 and 5 give a description of variables used in the next section and presents descriptive statistics for those variables. There are more cross-firm, more cross-inventor, more cross-border, more cross-tech, more cross-Atlantic citations in the second citations than in the first citations. 65% of citations are by firms with more than 100 patents and about 40% of citations are by triadic applications.

5. Empirical results

5.1. Main results

We start by estimating the hazard function (5). Column (1) of Table 6 presents the estimates using the Cox estimator with an indicator for whether the citation is made across firms, including controls for the industry and country of the citing patents. This estimator does not control for unobserved cited patent characteristics or censoring, and uses only first citations. The negative and significant coefficient suggests that inventors are quicker to cite patent applications within the firm than from a different firm. Column (2) adds controls for whether the citation is cross-inventor, cross-border and cross-technology, still using the Cox estimator with only first citations.

If citations within a firm tend to be quicker, this could simply be a reflection of the fact that inventor self-citations are more prevalent among citations within firms than among citations across firms. Unsurprisingly, cross-inventor citations are notably slower than citations to patents that share at least one inventor. The magnitude of the coefficient for cross-firm citations declines from -0.37 to -0.24 .

Searching for a patent that was applied for in another country can be more costly than searching domestically, because of differences in language and terminology, and

⁷ 35 U.S.C. 102 Conditions for patentability; novelty and loss of right to patent.

⁸ The difference in novelty requirements makes it more attractive for firms to apply for a patent in the US, as this already excludes the possibility of somebody else being granted a patent for the same invention.

Table 4
Description of variables.

Citation duration	Number of days between the priority date of the cited patent and the priority date of the citing patent
Cross-firm	Equals one if none of the applicants on the citing patent have the same ultimate owner as any of the applicants on the cited patent
Cross-inventor	Equals one if none of the inventors on the citing patent are listed as an inventor on the cited patent
Cross-border	Equals one if none of the countries of the inventors on the citing patent match any of the inventor countries of the cited patent
Cross-technology	Equals one if none of the Derwent Codes at 4 (3) digit level on the citing patent match any of the 4 (3) digit Derwent Codes of the cited patent
Citing firm has > 100 patents	Equals one if the number of patents applied for by the citing firm (ultimate owner) during the sample period is larger than 100
Citations per patent of citing firm	Equals one if the average number of citations received by patents applied for by the citing firm (ultimate owner) during the sample period exceeds one
Log patents per industry/country/year	The natural logarithm of yearly number of applications per Derwent Section (2 digit) and country of citing firm (ultimate owner)
Cross-Atlantic	Indicator variable for citations on applications by US ultimate owners to applications by European ultimate owners, and vice versa
Citing patent is triadic	Equals one if a citing application is for an invention that also has patent applications with the USPTO and the JPO
Examiner-added citation	Equals one if the citation was added by the examiner (rather than the applicant)
Industry dummies	Indicator variables for Derwent Sections (2 digit) of the citing patent

because the probability of learning about a patent through other channels than searching a patent database will be smaller for foreign patents. Likewise, searching for a patent in an unfamiliar technological field is more costly than searching for patents in a related technological field. The estimated coefficients for cross-border citations and cross-technology citations are less than half of the size of the coefficient for cross-firm citations. This suggests that the effect of firm boundaries is more than two times as strong as the effects of national borders or technology, though it is less than half the size of the cross-inventor effect.

If an existing patent is particularly valuable, then this will induce researchers to seek to build on this patent, rather than on a patent that is less valuable. If within-firm recombinations are less costly or more rewarding than cross-firm inventions, this can make modest inventions more likely to be the result of within-firm recombinations. The average quality of patents cited across firms could be larger than the average quality of patents receiving citations from the same firm. This could lead to underestimation of the delay in citations to the patents of other firms.

In column (3) we correct for biases due to heterogeneity in patent quality by using the fixed effects estimator (6), also controlling for right censoring. It is now necessary to

Table 5
Descriptive statistics.

Variable	First citation				Second citation			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Citation duration	809	427	0	1825	1600	1036	0	7230
Cross-firm	0.68	0.47	0.00	1.00	0.75	0.43	0.00	1.00
Cross-inventor	0.92	0.27	0.00	1.00	0.96	0.20	0.00	1.00
Cross-border	0.39	0.49	0.00	1.00	0.44	0.50	0.00	1.00
Cross-tech. (3 digit)	0.24	0.43	0.00	1.00	0.25	0.43	0.00	1.00
Cross-tech. (4 digit)	0.29	0.45	0.00	1.00	0.30	0.46	0.00	1.00
Citing firm has > 100 patents	0.65	0.48	0.00	1.00	0.65	0.48	0.00	1.00
Citations per patent of citing firm	0.37	0.48	0.00	1.00	0.28	0.45	0.00	1.00
Log patents per industry/country/year	6.43	1.13	0.00	8.48	6.53	1.15	0.00	8.53
Cross-Atlantic	0.23	0.42	0.00	1.00	0.27	0.44	0.00	1.00
Citing patent is triadic	0.41	0.49	0.00	1.00	0.39	0.49	0.00	1.00
Examiner-added citation	0.98	0.14	0.00	1.00	0.97	0.18	0.00	1.00

Table 6
Estimation results.

Coefficient	(1)	(2)	(3)	(4)	(5)
Cross-firm	-0.374 (0.01)	-0.243 (0.01)	-0.428 (0.02)	-0.427 (0.02)	-0.383 (0.02)
Cross-inventor		-0.677 (0.02)	-0.773 (0.04)	-0.746 (0.04)	-0.716 (0.04)
Cross-border	(0.01)	-0.075 (0.02)	-0.055 (0.02)	-0.054 (0.02)	-0.069
Cross-technology (3-digit)		-0.005 (0.03)	0.085 (0.04)	0.087 (0.04)	0.084 (0.05)
Cross-technology (4-digit)		-0.113 (0.02)	-0.240 (0.04)	-0.223 (0.04)	-0.219 (0.04)
Citing firm has > 100 patents				-0.069 (0.02)	-0.062 (0.02)
Citations per patent of citing firm				0.429 (0.02)	0.420 (0.02)
Patents per industry/country/year of citing patent					-1.123 (0.04)
Cross-Atlantic citation					-0.083 (0.02)
Citing patent is triadic					0.099 (0.02)
Examiner-added citation					0.598 (0.05)
Estimator	Cox	Cox	Fixed effects (censoring)	Fixed effects (censoring)	Fixed effects (censoring)

Notes: Standard errors are reported in parentheses. The duration is the time between the priority dates of the cited and citing patent. All patents are included that have been cited at least twice and with the first citation occurring within five years since the application date. The number of observations is 48,688. In each specification, dummies are included for the countries and Derwent Sections of citing patents.

use both first and second citations to estimate the fixed effects estimator. Note that we control for unobserved cited patent quality by using only within-cited-patent variations. The cross-firm coefficient increases in absolute terms from -0.24 to -0.43 , suggesting that the average expected value of recombinations based on a patent from within your own firm is smaller than for recombinations based on patents of another firm. This is consistent with Arrow's replacement effect (Arrow, 1962).

Controlling for the quality of inventions also changes the other coefficients. The estimated effects of cross-inventor and cross-technology increase, which is consistent with the hypotheses that it is more difficult to follow-up on other people's inventions and that it is more difficult to combine distinct technologies. The decrease in the effect of national borders is statistically insignificant.

Although we control for unobserved characteristics of the cited patents, it might be that our results are partly driven by characteristics particular to citing patents or citations. We add observable characteristics of citing patents in columns (4) and (5). In column (4) two controls are added. First, firms that have many patents (> 100) could be more likely to cite one of their own patents, even if cross-firm citations would not be slower than within-firm citations. Second, firms that own patents that receive many citations are likely to produce better patents and would therefore be more likely to cite one of their own patents.⁹ We find that citations by large firms tend to be somewhat slower and that citations by highly cited firms are considerably faster. The estimated cross-firm effect, as well as the other estimates are not affected.

In column (5) we add four more regressors. The natural logarithm of the number of patents per industry-country-year of the citing patent is included to take into account that patents are more prevalent in some industries, countries and years. For a 10% increase in the number of patents per industry-country-year of the citing patent, the coefficient decreases by about 0.11 ($\approx 1.12 \times \ln(1.1)$). The negative sign of this coefficient is consistent with the estimated effect of the "base" variable in Griffith et al. (2011). Second, an indicator for cross-Atlantic citations is included as citations of or by patents owned by American firms might take longer than citations between European firms. As expected from the coefficient of the cross-border effect, cross-Atlantic citations have a negative effect (are slower). Third, an indicator for citing patents that are also applied for at USPTO and JPO (triadic patents) is included. Triadic patents are more valuable, the expected value of research leading to these patents will also have been larger. This increases the hazard rate of citations from a triadic patent. Fourth, an indicator for whether the citation was added by the examiner (as opposed to the applicant), which increases the hazard of being cited. Adding these four controls increases the absolute size of the cross-border coefficient but reduces that of the other coefficients somewhat.

⁹ Recall that the variable labeled as "Citations per patent of citing firm" equals one if the average number of citations received by patents applied for by the citing firm (ultimate owner) during the sample period exceeds one.

Table 7
Estimation results by industry of ultimate owner of cited patent.

	Chemical	Electric	Engineering	ICT	Pharma
Cross-firm	-0.471 (0.04)	-0.428 (0.06)	-0.419 (0.05)	-0.311 (0.05)	-0.321 (0.07)
Cross-inventor	-0.744 (0.06)	-0.975 (0.11)	-0.625 (0.09)	-0.609 (0.10)	-0.701 (0.10)
Cross-border	-0.045 (0.04)	-0.016 (0.05)	-0.053 (0.04)	-0.058 (0.05)	-0.077 (0.07)
Cross-technology (3-digit)	0.130 (0.09)	0.121 (0.10)	0.115 (0.15)	0.036 (0.09)	0.064 (0.15)
Cross-technology (4-digit)	-0.235 (0.07)	-0.224 (0.08)	-0.316 (0.15)	-0.109 (0.06)	-0.304 (0.11)
Citing firm has > 100 patents	-0.107 (0.04)	-0.039 (0.04)	-0.128 (0.04)	-0.036 (0.04)	0.115 (0.06)
Citations per patent of citing firm	0.470 (0.03)	0.416 (0.04)	0.457 (0.03)	0.365 (0.03)	0.360 (0.05)
Patents per industry/country/year of citing patent	-0.645 (0.09)	-0.834 (0.06)	-0.829 (0.06)	-1.355 (0.06)	-0.586 (0.08)
Cross-Atlantic citation	0.024 (0.05)	0.003 (0.05)	-0.078 (0.05)	-0.119 (0.05)	-0.036 (0.07)
Citing patent is triadic	0.071 (0.03)	0.183 (0.04)	0.127 (0.04)	0.070 (0.03)	0.078 (0.05)
Examiner-added citation	0.610 (0.10)	0.772 (0.12)	0.629 (0.08)	0.205 (0.17)	0.674 (0.18)
Observations	14,234	8,116	10,308	10,850	5,180

Notes: Estimation results are for fixed effects estimation with censoring correction. Standard errors are reported in parentheses. For each industry of the parent firm all patents are included that have been cited at least twice and with the first citation occurring within five years since the priority date. Dummies are included for the countries and Derwent Sections of citing patents. Country-industry combinations with a small number of observations were grouped together.

Our final result suggests that the mean hazard rate of a citation across firms is 32% smaller ($1 - \exp(-0.38)$) than the hazard rate for a citation within the firm. The effect of ownership on the speed of recombination is substantial, much larger than the delay of recombination across geographic space (7%) or technological fields (13%, 3-digit and 4-digit combined).

5.2. Robustness

We consider the robustness of our results to four potential concerns. First, we consider whether the effects of ownership on the speed of recombination vary across industries. Second, we consider only inventor added citations. Third, we split the sample into two ten year time periods. Fourth, we use publication date rather than priority date.

Table 7 shows the results for our final specification for subsamples of patent applications by industry of use (based on Derwent classification). The differences across industries are generally modest. Cross-firm citations are fastest in ICT (a delay of 27%) and slowest in chemical engineering 37%. These results only slightly different from the delay

Table 8
Robustness.

	Applicant-added citations only	1985–1994	1995–2004	Corrected publication dates	Cross-inventor citations only
Coefficient	(1)	(2)	(3)	(4)	(5)
Cross-firm	−0.133 (0.22)	−0.267 (0.04)	−0.321 (0.05)	−0.396 (0.02)	−0.372 (0.02)
Cross-inventor	−0.450 (0.32)	−0.865 (0.07)	−0.409 (0.07)	−0.745 (0.04)	
Cross-border	−0.639 (0.27)	−0.073 (0.04)	−0.041 (0.05)	−0.050 (0.02)	−0.079 (0.02)
Cross-technology (3-digit)	0.122 (0.43)	−0.009 (0.09)	0.058 (0.09)	0.073 (0.04)	0.095 (0.05)
Cross-technology (4-digit)	0.194 (0.35)	−0.189 (0.07)	−0.219 (0.08)	−0.206 (0.04)	−0.232 (0.04)
Citing firm has > 100 patents		−0.017 (0.03)	0.005 (0.04)	−0.047 (0.02)	−0.067 (0.02)
Citations per patent of citing firm		0.632 (0.03)	0.541 (0.03)	0.468 (0.01)	0.418 (0.02)
Patents per industry/country/year of citing patent		−0.663 (0.10)	−0.721 (0.11)	−1.131 (0.03)	−1.148 (0.16)
Cross-Atlantic citation		−0.010 (0.04)	−0.116 (0.05)	−0.079 (0.02)	−0.072 (0.03)
Citing patent is triadic		0.032 (0.03)	0.055 (0.04)	0.145 (0.02)	0.103 (0.02)
Examiner-added citation		0.285 (0.29)	0.287 (0.08)	0.610 (0.05)	0.656 (0.05)
Observations	663	15,952	10,994	51,561	43,822

Notes: Estimation results are for fixed effects estimation with censoring correction. Standard errors are reported in parentheses. All patents are included that have been cited at least twice and with the first citation occurring within five years since the priority date. In each specification, dummies are included for the countries of citing patents. In columns (1)–(3) and (5), the duration is the time between the priority dates of the cited and citing patent. In column (4), the duration is the time between the publication dates of the cited and citing patent, where 540 days are subtracted for cross-inventor citations.

of 32% found for the full sample. In the industry-by-industry estimates cross-border citations are not significantly slower than citations within countries (point estimates range from 2% to 7%). This is consistent with findings in [Griffith et al. \(2011\)](#), [Keller \(2002\)](#) and [Thompson \(2006\)](#). At the 4 digit level, cross-technology citations are 10% (ICT) and 20–27% (other industries) slower. This is consistent with the delay of 20% found for the full sample.

[Table 8](#) show the results for the other robustness concerns. In column (1) we consider only applicant added citations. In our main results we include an indicator variable for whether a citation is added by the examiner. It enters significantly and positively, suggesting that examiner added citations are cited faster than applicant added ones, but has relatively little impact on the other coefficients. We show here that our results are robust to using only those citations that were added by applicants. This cuts the sample

size dramatically (from 48,688 to 663). While the coefficient estimates are unsurprisingly much less precise, and we are not able to estimate the full specification (col (5) in [Table 6](#)), we do still see an indication that our main results hold in this much smaller sub-sample.

In columns (2)–(3) we show results for two ten year time periods. There is little suggestion of parameter instability over these two time periods, expect perhaps for cross-inventor and cross-Atlantic citations. In the earlier period cross-inventor citations were much slower than self-citations, this effect is roughly halved in the later period. Cross-Atlantic citations on the other hand become slower (compared to within Europe citations) in the later period.

In column (4) we use publication date rather than priority date. A potential concern is that our finding, that citation duration within-firm is faster than between-firm, could be due to an obvious and trivial explanation that arises due to the institutions governing the patent system. Patent applications are not published immediately, but are normally published 18 months (540 days) after the priority date. It could simply be that this delay leads to a delay in information transmission between firms. While this is a market friction, it is perhaps not the most interesting friction. In column (4), we show that this is not the key factor driving our results, and that the evidence suggests that there are non-trivial market imperfections that drive this delay. We do this by using the publication date (rather than the priority) date of the cited application. Specifically, in column (4), the citation duration is defined to be the time between the publication dates of the cited and citing patent, where 540 days are subtracted for cross-inventor citations. Inventors outside the firm should not have information on a patent application until it is published (except perhaps through tacit knowledge gained through interaction of investors at conferences etc.). Indeed we do not see any citations occurring before the publication date of cited patent. The results in column (4) are similar to those in column (5) in [Table 6](#). In addition to using the corrected publication date, we also consider the subsample of cross-inventor citations only in column (6). The results in column (6) are again similar to those in column (5) in [Table 6](#).

6. Summary and discussion

We find robust empirical evidence that market frictions exist and that they are substantial in the market for technology. Firm boundaries lead to slower citation times. Our baseline regression results imply that the delay caused by firm boundaries is about 236 days (0.32×736) on average, much larger than the delay due to national borders or technological distance. These results speak to the debate on how well technology markets work. The theoretical literature emphasizes the importance of these frictions, but there has been relatively less empirical work. Our results suggest that the empirical magnitude of market frictions are substantial, and we interpret our evidence as pointing towards substantial imperfections in the market for technology.

Our paper has some limitations and we would like to end with a number of caveats. Our theoretical model is highly stylized and therefore we cannot exclude other possible

reasons for the observed cross-firm citation lags. For example, the citation lags may result from the delay due to learning or firms' strategic behaviors regarding firm-specific specializations in R&D and strategic choices in investments. In this paper, we establish that our simple model's predictions are consistent with the empirical findings and leave it a topic for future research to establish the tight links between the observed citation lags and firms' behaviors. In addition, as with all empirical work, we are only able to identify the parameters relevant for the time period that our data covers; it is possible that the world has changed since then. New information and communication technologies have reduced search costs, and it could be that these frictions have reduced in recent years. In the robustness section we do not see evidence of any major changes between 1985–1994 and 1995–2004, but it is of course possible that more recent years would look more different; we leave this to future work with more recent data.

Finally, while we have shown robust evidence that some sort of frictions are likely to be present, we have not addressed the sources of these frictions and why they might arise, which is important for policy formulation. There are a number of possible reasons, for example, frictions arising due to difficulties in writing or comprehending a new technology, intentional and strategic hiding by rivals, or uncertainty over property right. This is an important avenue for future research.

Acknowledgments

The authors thank an editor and two anonymous referees for helpful comments. The authors gratefully acknowledge financial support from the [Economic and Social Research Council](#) (ESRC) under the Centre for the Microeconomic Analysis of Public Policy (CPP), grant number [RES-544-28-0001](#), from the [European Research Council](#) (ERC) grant [ERC-2009-AdG-249529-MAPFAN](#), and from the European Research Council (ERC) grant [ERC-2014-CoG-646917-ROMIA](#).

References

- Abramovsky, L., Griffith, R., Macartney, G., Miller, H., 2008. The location of innovative activity in Europe. IFS Working Paper, WP08/10.
- Acemoglu, D., Bimpikis, K., Ozdaglar, A., 2008. Experimentation, patents, and innovation. NBER Working Paper, 14408.
- Alcácer, J., Gittelman, M., Samplatt, B., 2008. Applicant and examiner citations in US patents: an overview and analysis. Harvard Business School Working Paper, 09–016.
- Alfaro, L., Antràs, P., Chor, D., Conconi, P., 2015. Internalizing global value chains: a firm-level analysis. Working Paper.
- Arrow, K., 1962. Economic welfare and the allocation of resources for invention. In: *The Rate and Direction of Inventive Activity: Economic and Social Factors*. NBER, Princeton University Press, Princeton, pp. 609–626.
- Arundel, A., 2001. The relative effectiveness of patents and secrecy for appropriation. *Research Policy* 30, 611–624.
- Bessen, J., 2004. Holdup and licensing of cumulative innovations with private information. *Economic Letters* 82 (3), 321–326.
- Bessen, J., Maskin, E., 2009. Sequential innovation, patents and imitation. *RAND Journal of Economic* 40 (4), 611–635.

- Boldrin, M., Levine, D., 2008. Perfectly competitive innovation. *Journal of Monetary Economics* 55, 435–453.
- Boldrin, M., Levine, D., 2013. The case against patents. *Journal of Economic Perspectives* 27 (1), 3–22.
- Chamberlain, G., 1985. Heterogeneity, omitted variable bias, and duration dependence. In: Heckman, J.J., Singer, B. (Eds.), *Longitudinal Analysis of Labor Market Data*. Cambridge University Press, Cambridge.
- Chang, H., 1995. Patent scope, antitrust policy, and cumulative innovation. *RAND Journal of Economic* 26 (1), 34–57.
- Coase, R.H., 1937. The nature of the firm. *Economica* 16 (4), 386–405.
- Cohen, W.M., Levinthal, D.A., 1989. Innovation and learning: the two faces of R&D. *Economic Journal* 99 (397), 569–596.
- Costinot, A., Oldenski, L., Rauch, J., 2011. Adaptation and the boundary of multinational firms. *Review of Economics and Statistics* 93, 298–308.
- Criscuolo, P., Verspagen, B., 2008. Does it matter where patent citations come from? inventor vs. examiner citations in European patents. *Research Policy* 37, 1892–1908.
- Davidson, W., McFetridge, D., 1984. International technology transactions and the theory of the firm. *Journal of Industrial Economics* 32, 253–264.
- Gallini, N., Scotchmer, S., 2002. Intellectual property: when is it the best incentive system? *Innovation Policy and the Economy* 2, 51–77.
- Green, J., Scotchmer, S., 1995. On the division of profit in sequential innovation. *RAND Journal of Economics* 21 (1), 147–160.
- Griffith, R., Lee, S., Van Reenen, J., 2011. Is distance dying at last? Falling home bias in fixed effects models of patent citations. *Quantitative Economics* 2, 211–249.
- Grossman, G., Hart, O., 1986. The costs and benefits of ownership: a theory of vertical and lateral integration. *Quantitative Economics* 94 (4), 691–719.
- Henderson, R., Clark, K., 1990. Architectural innovation: the reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly* 35, 9–30.
- Henderson, R., Cockburn, I., 1990. Measuring competence? Exploring firm effects in pharmaceutical research. *Strategic Management Journal* 15, 63–84.
- Horowitz, J.L., Lee, S., 2004. Semiparametric estimation of a panel data proportional hazards model with fixed effects. *Journal of Econometrics* 119 (1), 155–198.
- Horstmann, I., MacDonald, G., Slivinski, A., 2004. Patents as information transfer mechanisms: to patent or (maybe) not to patent. *Journal of Political Economy* 93, 837–858.
- Irrazabal, A., Moxnes, A., Oromolla, L., 2013. The margins of multinational production and the role of intrafirm trade. *Journal of Political Economy* 121 (1), 74–126.
- Jaffe, A.B., Trajtenberg, M., 2002. *Patents, Citations and Innovation: A Window on the Knowledge Economy*. MIT Press, Cambridge.
- Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics* 108 (3), 577–598.
- Jones, B., 2009. The burden of knowledge and the death of the renaissance man: is innovation getting harder? *Review of Economic Studies* 76, 283–317.
- Keller, W., 2002. Geographic localization of international technology diffusion. *American Economic Review* 92 (1), 120–142.
- Lee, S., 2008. Estimating panel data duration models with censored data. *Econometric Theory* 24, 1254–1276.
- Mansfield, E., Romeo, A., 1980. Technology transfer to overseas subsidiaries by U.S.-based firms. *Quarterly Journal of Economic* 95, 737–750.
- Nelson, R., 1959. The simple economics of basic scientific research. *Journal of Political Economy* 67 (3), 297–306.
- Nordhaus, W., 1969. *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*. MIT Press, Cambridge USA.
- Ponzetto, G., 2012. Intellectual property rights and efficient firm organization. CEPR Discussion Paper 9212.
- Ridder, G., Tunali, I., 1999. Stratified partial likelihood estimation. *Journal of Econometrics* 92 (2), 193–232.
- Schumpeter, J.A., 1943. *Capitalism, Socialism and Democracy*. George Allen & Unwin, London.
- Scotchmer, S., 1991. Standing on the shoulders of giants: cumulative research and the patent law. *Journal of Economic Perspectives* 5 (1), 29–41.

- Shapiro, C., 2007. Patent reform: aligning reward and contribution 13141. NBER Working Paper
- 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.
- Thompson, P., Fox Kean, M., 2005. Patent citations and the geography of knowledge spillovers: a reassessment. *American Economic Review* 95 (1), 450–460.
- Tirole, J., 1988. *The Theory of Industrial Organization*. MIT Press, Cambridge, Mass. and London.
- van den Bergh, J., 2008. Optimal diversity: increasing returns versus recombinant innovation. *Journal of Economic Behavior and Organization* 68 (3–4), 565–580.
- Weitzman, M.L., 1998. Recombinant growth. *Quarterly Journal of Economics* 113 (2), 331–360.
- Williams, H., 2013. Intellectual property rights and innovation: evidence from the human genome. *Journal of Political Economy* 121 (1), 1–27.
- Williamson, O., 1971. Vertical integration of production: market failure considerations. *American Economic Review* 61 (2), 112–123.
- Williamson, O., 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. The Free Press, New York.
- Williamson, O., 1979. Transaction-cost economics: the government of contractual relations. *The Journal of Law and Economics* 22 (2), 233–261.
- Wuchty, S., Jones, B., Uzzi, B., 2007. The increasing dominance of teams in production of knowledge. *Science* 316 (18 May 2007), 1036–1039.
- Zeppini, P., van den Bergh, J., 2013. Optimal diversity in investments with recombinant innovation. *Structural Change and Economic Dynamics* 24 (March), 141–156.