RIETI Discussion Paper Series 24-E-036

Are Firms Able to Take Advantage of Academic Advances?

OROKU, Masahiro
Recruit Co., Ltd.

Research Institute of Economy, Trade & Industry, IAA
The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 24-E-036
March 2024

Are Firms Able to Take Advantage of Academic Advances?*

Masahiro OROKU
Recruit Co., Ltd.

Abstract

This study uses patents granted by the U.S. Patent and Trademark Office (applied before 2011) to analyze the relationship among the value of patents, account information of patent-holding firms, and citations of academic papers empirically. The results are summarized as follows: First, profitable firms tend to cite papers more frequently than other firms. Second, patents that cite academic papers have more forward citations than other patents do. Third, patents that cite academic papers are cited in a wider range of technical fields than those that do not. These results imply that incorporating academic knowledge increases patent value, expands utilization range, and increases firm profitability. This situation has implications for science and technology policy. Providing public support may be important if firms with low-profit firms cannot access academic knowledge because of the lack of human and material resources.

Keywords: Basic Research, Patent Citation, Absorptive Capacity, Spillover
JEL classification: O32 and O34

*This study was conducted as a part of the Project "Innovation, Knowledge Creation and Macroeconomy" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). For the analysis in this paper, we used the patent-paper linkage database from Clarivate Analytics Japan, the PATSTAT database from the European Patent Office, and the Orbis database from Bureau van Dijk, provided by the Center for Advanced Research in Finance at the University of Tokyo. This paper is supported by Eisuke Sato (Tokyo Univ.) as a research assistant. The author is grateful for the helpful comments and suggestions provided by Makoto Nirei (Tokyo Univ.) , Koki Oikawa(Waseda Univ.) and Discussion Paper seminar participants at RIETI.

1 Introduction

What are firms searching for in academic knowledge? We aim to answer this question by analyzing the relationship between firms and academic knowledge. Using patents applied for and granted by the U.S. Patent and Trademark Office before 2011, we empirically analyzed the relationship among the value of patents, characteristics of patent-holding firms, and citations of academic articles to examine how these variables interact to shape the ways in which firms seek and use academic knowledge. Through this analysis, we hope to elucidate the mechanisms through which academic research contributes to firm innovation and market performance and provide insights that can guide firms and policymakers in fostering innovation and economic growth.

First, we explain the background of this study. Previous studies have demonstrated that the connections between academic knowledge and firms are strengthening. Recent research (e.g. Narin et al. (1997)) has shown that for the U.S. industry, relying on external sources of knowledge on public science is a problem. In Japan, the NISTEP survey[1] shows that introducing knowledge from universities or public research institutes is useful for private firms, especially when searching for R&D themes. The depletion of innovation is supposed to be the reason why firms externally search for new ideas. Bloom et al. (2020) stated that while the number of people engaged in R&D continues to grow, productivity growth decelerates. In other words, as the stock of knowledge grows, producing new knowledge becomes increasingly difficult.

Several theoretical frameworks for searching for and exploiting ideas beyond the boundaries of firms exist. Cohen and Levinthal (1989) and Cohen et al. (1990) discussed that R&D not only generates new information but also enhances the ability of firms to assimilate and exploit existing information. They named that ability "absorptive capacity". Kortum (1997) discussed this issue by constructing a search-theoretic model of knowledge growth. In this model, inventors sample from a probability distribution of potential technological breakthroughs or ideas as a Poisson process. Past research has generated a technological frontier that represents the optimal technology for producing each good in the economy. However, technological breakthroughs have become increasingly difficult to identify as the technological frontier moves forward. Olsson (2000) and Olsson (2005) mapped the set of knowledge to a bounded open set in a generalized metric space and represented technical opportunities as topological characteristics of the knowledge set.

We now review empirical papers, particularly those focusing on the study of citation data. In a pioneering paper, Eugene (1955) described the notable uses of citation index as "association-of-ideas index" and "help the historian measure the influence of the article—that is, its 'impact factor'". Griliches (1979) described the basic production function of new knowledge with R&D inputs and suggested that the frequency with which patents from different industries cite each other could be utilized as a measure of the technological proximity of in-

[1] `http://hdl.handle.net/11035/2948` p.92

dustries. Based on Griliches' framework, Jaffe (1989) identified geographically mediated knowledge transfer from university research to commercial innovation. Regarding the relationship between economic value and citation metrics, Narin et al. (1987) showed that in publicly traded pharmaceutical companies, the average citation frequency of a patent portfolio was associated with increased corporate profits and sales. Hall et al. (2005) examined the relationship between the intensity of citation and the private value of patents by relating citation-weighted patents to patent value.

With respect to the relationship with academic knowledge, Harhoff et al. (2003) found that the number of citations in nonpatent literature predicts patent value in the pharmaceutical and chemical fields. Lemley and Sampat (2012) also found that most nonpatent prior art references in the U.S. Patent and Trademark Office (USPTO) originate from applicants, not examiners. Therefore, focusing on nonpatent articles is useful for tracking the flow of knowledge. In the context of patent citation analysis, the exploitation of knowledge takes the form of patents. The effects of new knowledge can be traced through citations between patents. Hall et al. (2005) examined the relationship between the intensity of citation and the private value of patents by relating citation-weighted patents to patent value. Harhoff et al. (2003) discovered that the number of citations in nonpatent literature predicts patent value.

The rest of this paper is organized as follows. The hypotheses of academic-industry relationships are described in Section.2. Data and variable constructions are explained in Section.3. The analysis and policy implications of the three main topics are discussed in Section.4. The conclusion and explanation of future extensions are given in Section.5. The details of the data are provided in the Appendix.

2 Academic-Industry Relationships

The reference to academic knowledge by R&D in a firm may be broken down as follows: When R&D staff in a firm plan a new R&D project, they first develop subjective beliefs about the new R&D project based on external information, including academic papers and past R&D achievements. Subsequently, they decided whether to implement the project. Given that the success or failure of a project cannot be completely foreseen in advance, the outcome is uncertain. Successful R&D results in patents and other innovations that bring monopolistic profits to the firm.

As Cohen and Levinthal (1989) and Cohen et al. (1990) described that a firm's "absorptive" capacity to recognize the value of new external information and to assimilate it is a function of the firm's level of prior-related knowledge. At the individual level, absorptive capacity includes the cognitive basis for prior-related knowledge and background diversity. At the organizational level, the absorptive capacity of the firm is different from that of its individual members. They investigated the diversity of expertise within an organization. Hence, our first hypothesis about the academic-industry relationship focuses on the characteristics of the organization. We examine how firms with low profits are less able to explore academic knowledge than other firms. If so, this inability may be attributed to a lack of diversity of expertise.

We are interested not only in the exploration of academic knowledge but also in the exploitation of knowledge. Our second hypothesis concerns the quality of the project's results, which are based on academic knowledge. We examined whether the use of academic knowledge improved the quality and the value of outcomes. Our third hypothesis concerns the range of influence of academic knowledge. As Akcigit et al. (2021) described, basic research is expected to have a broader spillover effect than applied research. We examined the effect of the use of academic knowledge across a wide range of industry sectors.

Our conceptual model of the academic-industry relationship is shown in Figure.1. Notes in brackets are proxy variables in the empirical analysis and are explained in the next section.

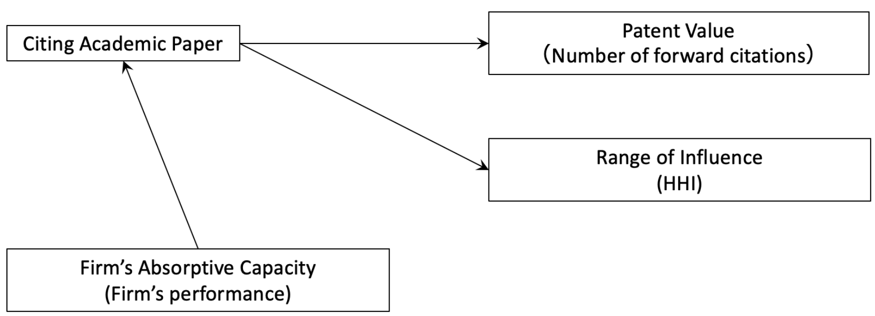

Figure 1: Conceptual model

3 Data and Variables

3.1 Data Overview

Our analysis is based on a comprehensive database of firms, patents and academic papers in the U.S. from 2004 to 2011. The firm's data, including financial information, are from Bureau van Dijk's Orbis database[2]. The patent data were obtained from the European Patent Office's Patstat database[3]. The academic paper data are from Clarivate Analytics' Web of Science database[4].

We constructed a combined database by linking the above three datasets. Citation linkages from patents to academic papers are provided by Clarivate Analytics. Patent ownership information for firms is included in the Orbis database. Cases wherein one patent was applied for by multiple firms were excluded. An overview of the database is shown in Figure.2. Underlined columns represent the primary key of each database. Details of each column are summarized in Appendix.A.

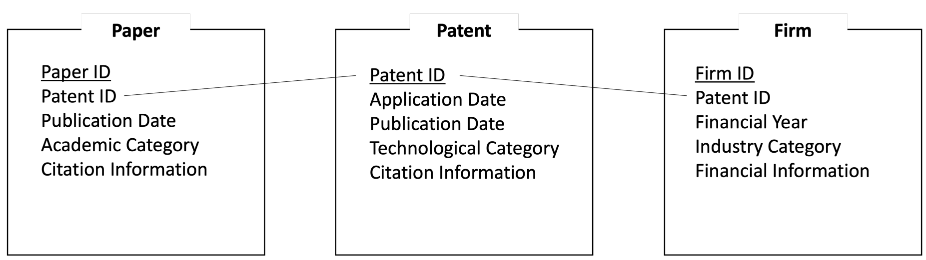

Figure 2: Overview of the database

The data extraction conditions are summarized in Table.1. We focused on patents that were applied for and granted in the U.S. We excluded patents with less than five forward citations to ensure the quality of patents. Given the availability of data, patents filed from 2004 to 2011 and financial information from 2004 to 2011 were used. Considering that financial information includes missing values, the data constructed an imbalanced panel data of firms.

Table 1: Data extraction condition

Variable	Condition
Country of patent application	US
Granted	True
Minimum number of patent forward citations	5
Year of the patent application	2004 - 2011
Year of the financial information	2004 - 2011

[2] https://www.bvdinfo.com/ja-jp/our-products/data/international/orbis
[3] https://www.epo.org/searching-for-patents/business/patstat.html
[4] https://clarivate.com/ja/solutions/web-of-science/

The sample size of the data analyzed in this paper is summarized in Table.2.

Table 2: Sample size

	Number of patents	Number of firms
Total	118891	1728
With paper citations	14267	767

3.2 Variables and Summary Statistics

We constructed the variables by using the characteristics of the citation network. Let i, j, f and t be the indices of paper, patent, firm and year, respectively. $P_{f,t}$ is the set of patents applied by firm f in year t. Paper citation dummy d_j^p indicates whether patent j cites academic papers or not:

$$d_j^p = \begin{cases} 1 & \text{if patent } j \text{ cites academic papers} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Paper citation ratio $r_{f,t}$ represents the proportion of patents applied by firm i that cite academic papers and is calculated as follows:

$$r_{f,t} = \frac{\sum_{j \in P_{f,t}} d_j^p}{|P_{f,t}|} \tag{2}$$

In accordance with previous studies and as presented by Hall et al. (2005) and Kogan et al. (2017), we used the logarithm of the number of forward citations fc_j as the value of patents and papers: $\log 1 + fc_j$. We apply the Harfindahl-Hirschman Index (HHI) of forward citations as an indicator of the technological variety v_j of the patent : $v_j = 1 - HHI_j$. The details and definition of HHI are described in Section.4.3.

The technological field dummy shows whether a patent is applied in a technological field. The industrial field dummy identifies whether the patent-owner firm is in the industrial field or not. Sales were used as an indicator of firm size. The gross margin ratio (the ratio of gross profit to sales) was used as an indicator of firm profitability. R&D ratio (ratio of R&D expenses to operating revenue) is employed as an indicator of a firm's R&D intensity.

The list of variables is summarized in Table.3. The variables aggregated by the firm are applied to the analysis described in Section.4.1. The variables aggregated by patent are used for the analysis described in Section.4.2 and Section.4.3.

The summary statistics for the variables aggregated at the patent level are presented in Table.4. The number of observations was 118891. Log patent forward citations and backward citations are the logarithm of the number of forward citations and backward citations for each patent, respectively. is the logarithm of the number of forward citations for each patent.

Table 3: List of variables

Variable	Definition
i	index for paper
j	index for patent
f	index for firm
t	index for year
d_j^p	paper citation dummy of the patent
fc_j	# of patent forward citations
bc_j	# of patent backward citations
v_j	technological variety of the patent
d_j^{tech}	technological field dummy of the patent
d_f^{ind}	industrial field dummy of the firm
$sales_f^t$	sales of the firm
GMR_f^t	gross margin ratio of the firm
$R\&D_f^t$	R&D ratio of the firm

Table 4: Summary statistics

Name	Mean	Std. Dev.	Min	Max
paper citation dummy	0.120	0.325	0.000	1.000
log patent forward citations	0.269	0.191	0.074	2.215
log patent backward citations	1.231	0.435	0.301	3.568
technological variety	0.650	0.227	0.000	0.965

The summary statistics for the variables aggregated at the firm level are shown in Table.5. The number of observations (firm × year) is 12810. The sales values are the logarithm of real sales (in thousands of U.S. dollars) for each industry deflated by the producer price index[5] for each industry.

Table 5: Summary statistics

Name	Mean	Std. Dev.	Min	Max
paper citation ratio	0.103	0.233	0.000	1.000
log sales	3.238	1.190	0.000	6.457
gross margin ratio	0.431	0.225	-0.982	1.000
R&D ratio	0.115	0.062	0.000	1.000

[5] https://www.bls.gov/ppi/

4 Empirical Analysis

We explain the main empirical analyses in this section. We investigate how the citation of academic papers relates to firm performance, patent value, and technological variety. Subsequently, we discuss the interpretations and possible policy implications of the results.

4.1 Firm Performance

We focus on two characteristics of patent-holder firm performance: size and gross margin.

Sales

First, we observe the relationship between paper citation and firm size, which is measured by sales. Figure.3 shows the histogram of firm-level average sales in log form during the sample period with the average paper citation ratio among firms in each histogram bin (the vertical bar for each dot represents the 95% confidence interval). The relationship is not monotonic.

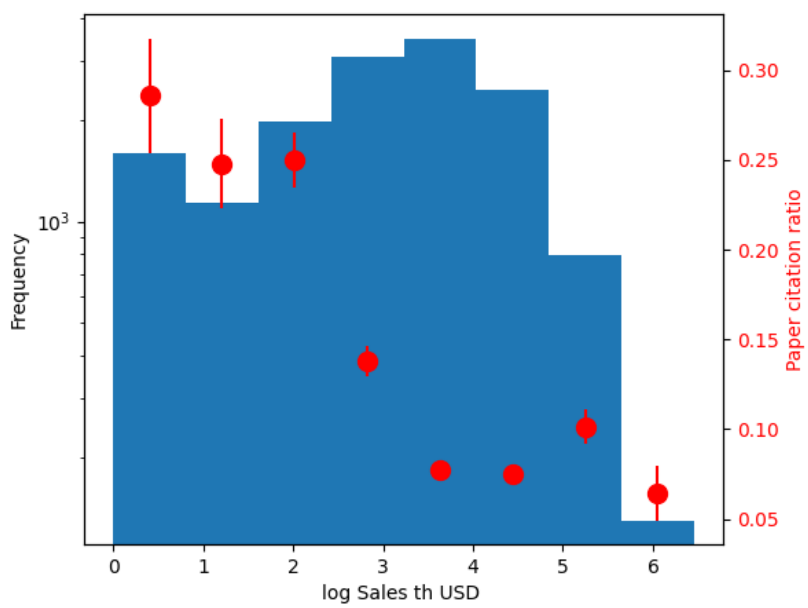

Figure 3: Sales size and paper citation ratio

Figure.4 shows the histogram of paper citation ratios for the size groups divided by the median of sales. We can see that numerous small firms engage in basic research. However, there is considerable heterogeneity among these firms. The results of the Mann-Whitney U test are presented in Table.6. The U test examines whether a particular population tends to have a higher value than another population based on the null hypothesis that the two populations are the same. $p = 0.08$ shows that firms with low sales appear to have a higher paper citation ratio than those with high sales. However, the evidence for rejecting the null hypothesis is insufficient.

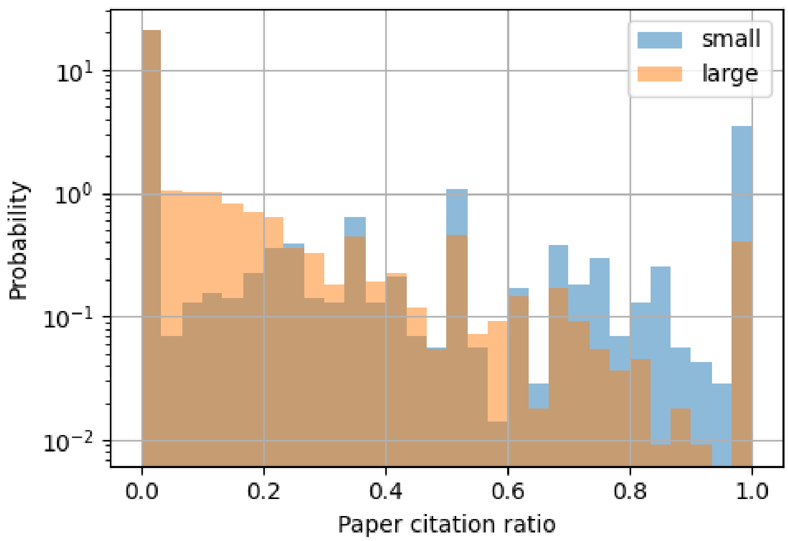

Figure 4: Histograms of paper citation ratios for firms with small and large sales

Table 6: Distribution of paper citation ratio by sales size

	Mean	Std. Dev.
small sales firms	0.18	0.33
large sales firms	0.08	0.17
		$U = 3.75 \times 10^6$
		$p = 0.08$

Gross Margin Ratio

Second, we observed the relationship between paper citation ratio and gross margin ratios in the same way. Figure.5 depicts the relationship between the gross margin ratio (only for firms with positive margins) during the sample period and the average paper citation ratio among firms in each bin of the histogram. This figure shows a positive correlation between the two variables: Successful firms, in terms of gross margin, tend to have higher paper citation ratios than other firms. In other words, firms that invest in basic research tend to earn greater profits than those that do not.

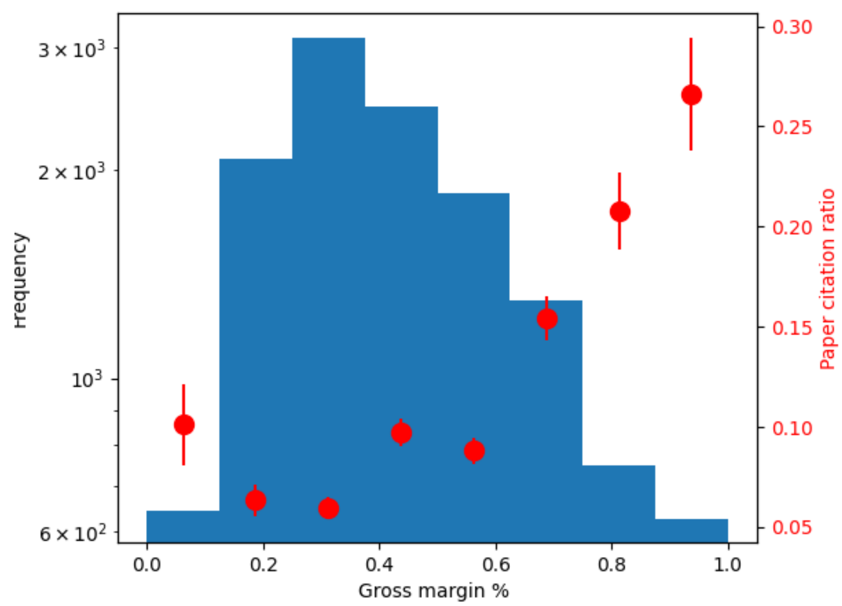

Figure 5: Gross margin ratio and paper citation ratios

Figure.6 shows the histogram of paper citation ratios for the groups divided by the median of gross margin. This figure also shows a positive correlation between the gross margin and paper citation ratio. The results of the Mann-Whitney U test are presented in Table.7. $p = 0.00$ showing that firms with large gross margin tend to have higher paper citation ratios than those with small gross margins.

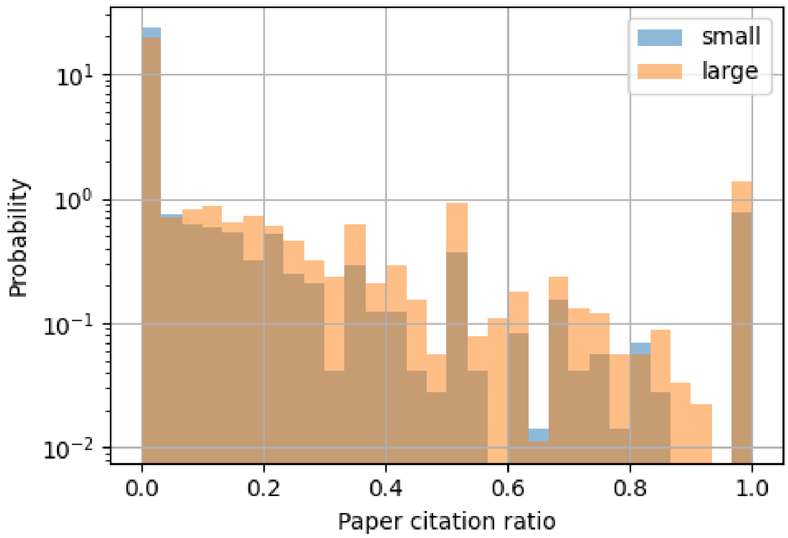

Figure 6: Histograms of gross margins with and without paper citations

Table 7: Distribution of paper citation ratio by gross margin

	Mean	Std. Dev.
small gross margin firms	0.07	0.19
large gross margin firms	0.13	0.26
		$U = 2.50 \times 10^6$
		$p = 0.00$

4.2 Patent Value

Generally, two types of patent value exist: business measures, such as stock prices and sales, and technical measures, such as technological quality and effect. Previous studies by Hall et al. (2005) and others have revealed that the business value measured by stock price and technical value measured by the number of citations are positively correlated, and measuring the value of a patent by the number of forward citations is generally considered reasonable. Our research also followed this approach.

To investigate the relationship between paper citation and patent value, We regress the value of patent value on the paper citation ratio and other explanatory variables using the ordinary least squares (OLS) method. We used the to investigate the relationship between paper citations and patent value. We use the information of the firm that applied for the patent as an explanatory variables. First, we used the paper citation dummy, the log of sales, the log of the number of backward citations, and R&D expenditure, year dummy, and technological field dummy as the explanatory variables. Second, we added the interaction term of the year dummy and paper citation dummies to the explanatory variables. Third, we added the interaction term of the year dummy and technological field dummies.

The results are summarized in Table.8. These results illustrate that citations of academic papers increase the patent value. The results are robust to application duration, technical field, and firm size. We found that that the number of backward citations is correlated with the number of forward citations. This result was consistent with the findings of Harhoff et al. (2003). The negative coefficients of sales and R&D indicate that the value of patents decreases as these variables increase. This effect could be due to the tendency of firms to obtain broad, general patents to increase sales. Given that these patents provide protection for specific products or technologies, their value may be limited. In addition, as the R&D ratio (ratio of R&D expenses to operating revenue) increases, companies may become increasingly likely to acquire patents that are not directly related to their core business. In other words, R&D has a diminishing effect on company earnings relative to input. In such cases of overinvestment in R&D, the value of patents may also be limited.

Table 8: Regression Result(OLS)

	(1)	(2)	(3)
paper citation	0.015***		
	(0.002)		
paper citation (applied in 2004)		0.017***	0.019***
		(0.003)	(0.004)
paper citation (applied in 2005)		0.013***	0.014***
		(0.004)	(0.004)
paper citation (applied in 2006)		0.020***	0.023***
		(0.004)	(0.004)
paper citation (applied in 2007)		0.011**	0.011***
		(0.005)	(0.004)
paper citation (applied in 2008)		0.001	-0.001
		(0.006)	(0.005)
paper citation (applied in 2009)		0.008	0.006
		(0.010)	(0.007)
paper citation (applied in 2010)		0.057***	0.057**
		(0.019)	(0.010)
backward citation (log)	0.054***	0.054***	0.054***
	(0.002)	(0.002)	(0.002)
sales (log)	-0.012***	-0.012***	-0.012***
	(0.001)	(0.001)	(0.001)
R&D ratio	-0.072***	-0.072***	-0.072***
	(0.007)	(0.002)	(0.002)
application year FE	Yes	Yes	Yes
technology field FE	Yes	Yes	Yes
application year $\times$ technology field FE	No	No	Yes
Observation	96279	96279	96279
Adj. R-squared	0.04	0.04	0.04

* $p < 0.10$
** $p < 0.05$
*** $p < 0.01$

4.3 Technological Variety

We investigated whether patents citing academic papers are cited in a wider range of technical fields than patents that do not. As an indicator of broadness, we use the Herfindahl-Hirschman Index (HHI) calculated for each patent in the form of Equation.3. In this equation, s_j represents the share of the technological field $j \in \{1, \dots, J\}$ within the set of forward citations of a patent.

$$HHI = s_1^2 + s_2^2 + \ldots + s_J^2 \tag{3}$$

HHI takes a continuous value from 0 to 1. When HHI is small (or $1 - HHI$ is large), the patent is cited in a wide range of field.

We compared distributions between patents with and without academic paper citations. The result is shown in Figure.7.

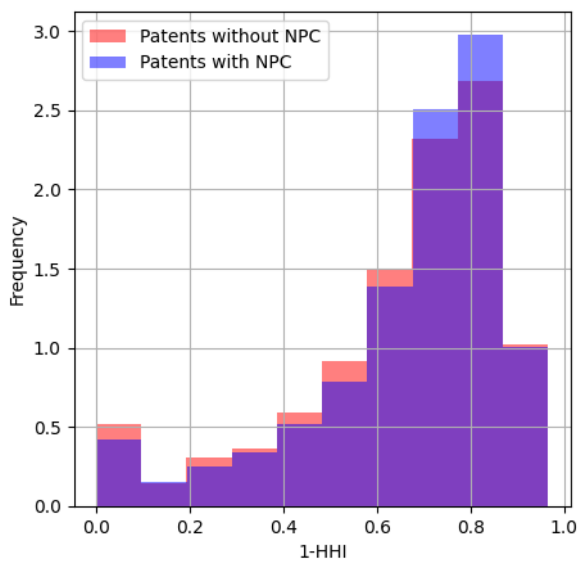

Figure 7: Patent distribution of HHI

The distribution of $1 - HHI$ and the results of the Mann-Whitney U test are summarized in Table.9. $p = 0.00$ shows that patents with paper citations tend to be cited in a wider range of technological fields than those without.

Robustness was checked over the technological field of the cited patent, as shown in Figure.8. A general tendency across different technological fields is observed. This tendency is particularly strong in the fields of human necessities (A) and textiles (D).

Table 9: Distribution of 1 - HHI by paper citation

	Obs.	Mean	Std. Dev.
without paper citation	104607	0.65	0.23
with paper citation	14284	0.67	0.22
			$U = 7.11 \times 10^8$
			$p = 0.00$

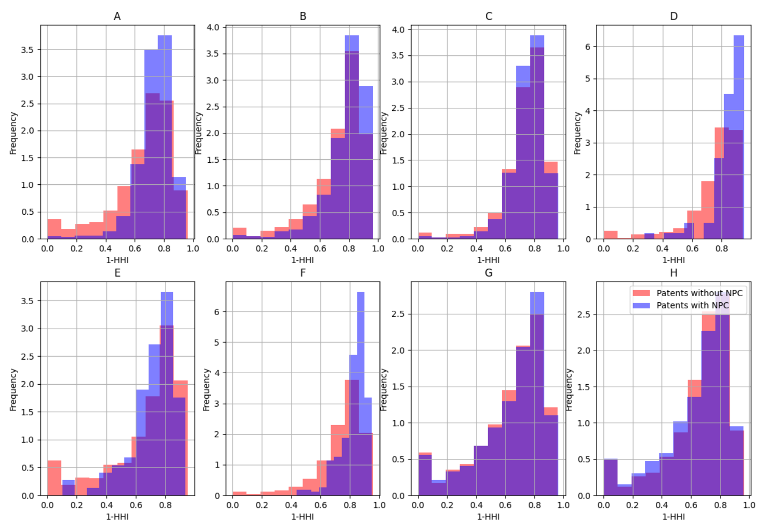

Figure 8: Patent distribution of HHI in technological fields

4.4 Discussion

The initial analysis of our data reveals a convincing relationship between the profitability of firms and citation of academic papers. This trend suggests a deep interaction between a firm's absorptive capacity for academic knowledge and its overall profitability. More profitable firms appear to be more likely to actively engage in and use academic research than less profitable firms. This situation implies a possible synergy between academic knowledge and commercial success.

Further analyses reinforce this idea. Citing academic papers positively influences the value and technological breadth of patents. These findings emphasize the usefulness of academic knowledge in developing patents that are not only valuable but also have various technological effects. This situation is important in a rapidly evolving technological landscape where the integration of diverse knowledge can be an important driver of innovation.

The positive correlation between paper citations and firm profitability has several possible explanations. The first consideration is the role of advanced R&D capabilities as a latent variable. Firms or individuals with higher R&D skills, often sharpened through rigorous academic training such as graduate studies, are probably more skilled at applying academic knowledge than those without. Given that such firms can better understand and utilize state-of-the-art research than other firms, this skill level may lead to patents of increased value. As a result, article citations and patent value are positively correlated. This correlation suggests that academic engagement is not simply a function of the availability of specific resources but is also closely tied to the skill level of individuals within a firm.

Another plausible explanation relates to the nature of academic research itself. Academic research often includes novel ideas that cross multiple disciplines, thus introducing fresh perspectives and innovative approaches. Such cross-disciplinary knowledge can help increase the value of patents. In our analysis, patents that cite academic papers tend to cover a broader range of technical fields than those that do not. This observation is consistent with the notion that exposure to various academic studies can lead to a broadened technical scope in patent applications.

5 Conclusion and Future Work

We attempted to clarify the relationship between academic knowledge and industrial activities. We constructed a database that combines academic papers, patents, and firms to analyze knowledge transfer from academic research to industry through patent citations.

We obtained three main findings through our analysis. First, profitable firms tend to cite academic papers. Second, patents that cite academic papers have more forward citations than those that do not. Third, patents that cite academic papers are cited in a wider range of technical fields than those that do not.

Our findings illustrate a beneficial relationship between the academic field and the patent domain. Firms that effectively absorb and apply academic knowledge appears to gain a competitive advantage, as indicated by their patents with increased value and technological diversity. This relationship emphasizes the importance of developing strong connections between industry and academia, wherein the exchange of knowledge and ideas can have good results in terms of innovation and profitability.

We now discuss our future work. First, an analysis from a social welfare perspective is required. A macroeconomic study is also necessary. Our results imply that basic research can be applied in a wide range of fields. Spillover and long-term growth will become questions within the framework of macroeconomic growth theory. Finally, our study takes the position of a linear model, wherein the results of basic research lead linearly to applied research, development, and product commercialization. We assume a relatively simple picture wherein publicly available academic knowledge leaks out to the private sector and has a positive influence. Our study assumes only a supply-side effect from academic research activities that have nothing to do with the market. In reality, however, some innovations may be realized by firms investing R&D resources in areas where demand is growing. These phenomena are referred to as "technology push" and "demand pull" in the context of innovation studies. Demand–pull innovation was not covered in this study and will be a subject of future research.

References

Akcigit, Ufuk, Douglas Hanley, and Nicolas Serrano-Velarde, "Back to Basics: Basic Research Spillovers, Innovation Policy, and Growth," *Review of Economic Studies*, 2021, *88* (1), 1–43.

Bloom, Nicholas, Charles I. Jones, John van Reenen, and Michael Webb, "Are ideas getting harder to find? † ," *American Economic Review*, 2020, *110* (4), 1104–1144.

Cohen, Wesley M. and Daniel A. Levinthal, "Innovation and Learning: The Two Faces of R & D," *The Economic Journal*, 1989, *99* (397), 569.

Cohen, Wesley M, Daniel A Levinthal, J. R. COBB, Wesley M Cohen, and Daniel A Levinthal, "Absorptive capacity: a new perspective on learning and innovation," *Administrative science quarterly*, 1990, *40 A* (3), 128–152.

Eugene, Garfield, "Citation Indexes for Science," *Science*, 1955, *122* (3159), 108–111.

Griliches, Zvi, "Issues in assessing the contribution of research and development to productivity growth," *The Bell Journal of Economics*, 1979, *10* (1), 92–116.

Hall, Bronwyn H, Adam Jaffe, and Manuel Trajtenberg, "Market value and patent citations," *RAND Journal of economics*, 2005, pp. 16–38.

Harhoff, Dietmar, Frederic M Scherer, and Katrin Vopel, "Citations, family size, opposition and the value of patent rights," *Research Policy*, 2003, *32* (8), 1343–1363.

Jaffe, Adam B, "Real effects of academic research," *The American Economic Review*, 1989, pp. 957–970.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, "Technological Innovation, Resource Allocation, and Growth*," *The Quarterly Journal of Economics*, 2017, *132* (2), 665–712.

Kortum, Samuel S, "Research, patenting, and technological change," *Econometrica: Journal of the Econometric Society*, 1997, pp. 1389–1419.

Lemley, Mark A. and Bhaven Sampat, "Examiner characteristics and patent office outcomes," *Review of Economics and Statistics*, 2012, *94* (3), 817–827.

Narin, Francis, Elliot Noma, and Ross Perry, "Patents as indicators of corporate technological strength," *Research Policy*, 1987, *16* (2), 143–155.

_ , **Kimberly S. Hamilton, and Dominic Olivastro**, "The increasing linkage between US technology and public science," *Research Policy*, 1997, *26* (3), 317–330.

Olsson, Ola, "Knowledge as a Set in Idea Space: An Epistemological View on Growth," *Journal of Economic Growth*, 2000, *5* (3), 253–275.

_ , "Technological opportunity and growth," *Journal of Economic Growth*, 2005, *10* (1), 35–57.

A Data Columns Details

Details of each columns used in this analysis are as follows.

Paper ID We use Web of Science[6] accession number (UT) as a unique identifier for scientific papers. IDs are for example "WOS:000075806100061" or "WOS:A1977DL33400271"

Patent ID We use USPTO[7] publication number as a unique identifier for patents in US. ID is for example "US760499B2". The first two letters "US" are the region symbol. Kind Code "B2" gives status of patents[8].

Firm ID We use Bureau van Dijk[9] ID number as a unique identifier for firms. ID is for example "US140689340".

Academic Category We use Essential Science Indicators(ESI) from Web of Science as scientific categorization. In our analysis, bigger classification based on ESI is used. That classification which is aggregated by NISTEP [10] consists of eight categories in total.

Technological Category We use International Patent Classification(IPC) classified by World Intellectual Property Organization(WIPO)[11] as technological categorization. IPC is for example "H01B" which represents "CABLES; CONDUCTORS; INSULATORS". The first letter, which is from A ("Human Necessities") to H ("Electricity"), is the section symbol. Two digit number following gives a class symbol. For example "H01" represents "BASIC ELECTRIC ELEMENTS". The final letter makes up the subclass symbol.

[6] https://www.webofscience.com

[7] https://www.uspto.gov/

[8] https://www.uspto.gov/learning-and-resources/support-centers/electronic-business-center/kind-codes-included-uspto-patent

[9] https://www.bvdinfo.com/ja-jp/

[10] http://www.nistep.go.jp/research/science-and-technology-indicators-and-scientometrics/benchmark

[11] http://www.wipo.int/classifications/ipc/en/