Koichiro KIMURA, ed. (2021) *Impacts of Innovation on Firm Performance and Industrial Development in East Asia* (*BRC Research Report*), Bangkok: Bangkok Research Center, JETRO Bangkok.

Chapter 3

Technology Development and Similarities*

Koichiro KIMURA†, Hiroshi MATSUI‡, Kazuyuki MOTOHASHI§,
Shun KAIDA**, and Janthorn SINTHUPUNDAJA††

Abstract: Firms must choose which technologies or technological fields to focus on relative to their competitors when building their competitive technology positions. To examine the similarities of firms in terms of their technological fields, we conduct a case study on the similarities between a fast-growing Chinese robotics firm and a Japanese first-mover firm within the same industry. We show that as the number of patent applications by the Chinese firm increases, the technologies of the two firms become more similar in terms of the level of technological positions but more diverse in terms of the level of patent documents. In other words, as the Chinese firm develops technologies in major technological fields of robotics, it works on a variety of technologies within the major fields.
Key Words: patent; technology position; technological similarity; robotics industry; China

1. Introduction

An increasing number of firms in emerging countries are focusing on technology development for competitive survival and further growth. As a result, innovation centers have dispersed geographically from traditional developed countries to emerging countries. Moreover, recent technological changes arising from the Fourth Industrial Revolution have led to a rapid increase in business opportunities for firms to develop new technologies. This has allowed firms in emerging countries to increase their competitive advantages through aggressive technology development and innovation activities. [1]

Patent documents play a key role in researching and understanding the technology of firms. Patents have often been used as an indicator of innovation because they contain useful technological information and expand the availability of databases in various countries (Nagaoka et al., 2010). Much research has been conducted on patents, such as studies on the factors of patent production and how the number and quality of patents affect business performance. Consequently, patent research has revealed a great deal about the background and impact of technology development and innovation activities.

Moreover, patents can indicate technological similarities or differences among firms through International Patent Classification (IPC) codes and other technology classifications assigned to patent documents. The IPC is a hierarchical classification of patents according to technological fields. Using the IPC, the technology position of each firm can be identified according to a vector defined as a firm's number of patents in a technological field as a fraction of its total number of patents. Firms face choices in the technologies or technological fields they focus on and acquire to build their competitive technology positions relative to industry competitors. In other words, the choice of technological distance from rivals is critical for firms under fierce competition.

Therefore, we will analyze how the number of patent applications made by a firm in an emerging country changes its technology position compared with that of a first-mover firm in the same industry. Much research has already accumulated on technological

---

* The authors would like to thank Professor Gary Jefferson (Brandeis University) and Mr. Zhenyu Gong (Brandeis University) for the useful discussion at an online meeting held by the Institute of Developing Economies on December 18, 2020.
† Associate Senior Research Fellow, Institute of Developing Economies, JETRO, Japan and IDE Bangkok Research Center, JETRO Bangkok
‡ Expert, ABeam Consulting (Thailand) Ltd.
§ Professor, The University of Tokyo
** Senior Consultant, ABeam Consulting Ltd., Japan
†† Consultant, ABeam Consulting (Thailand) Ltd.
[1] Competitiveness depends on not only technology but also business models, knowledge in a broader sense, and other factors. Moreover, innovation is a broad concept not limited to technology, as it refers to bringing new value to society through channels such as new product development, reorganization, and others.

similarity; we will utilize it here as well. However, the concept of similarity has often been used to verify whether technological spillover exists among industries or firms with advanced levels of technological similarity (Bloom et al., 2013; Jaffe, 1986). In contrast, we will focus on the formation pattern of the technology position itself—that is, how technology position changes with an increasing number of patent applications. [2]

Additionally, we will use natural language processing (NLP) to calculate technological similarity. [3] Technology positions based on the IPC are determined from all or several of a firm's patent documents, not from individual patent documents. Therefore, if we can obtain a technology position from each patent document, we can also calculate the technological similarities at the level of individual patent documents. Consequently, it will even be possible to show differences between technologies classified within the same technological field. To accomplish that, we will vectorize each abstract of patent documents based on NLP in addition to vectorizing each technology position at the firm level based on the IPC. We will then compare the vectors generated by NLP with those based on the IPC and examine the potential use of NLP-generated vectors for future research on technological similarities or differences and changes in those characteristics.

Therefore, the purposes of this case study are, first, to show the technological similarities between a latecomer (a rapidly growing Chinese robotics firm) and first-mover (a Japanese firm within the same industry) and, second, to examine the availability of vectors generated by NLP. Specifically, we will first show that as the number of patent applications increases, the technologies of the two firms become more similar in their level of technological positioning, but the technologies for which they submit patent documents become more diverse. In other words, the Chinese firm, while developing technologies in major robotic technological fields, works on a variety of technologies in those major fields. As for the second purpose, this study will be fundamental research to develop a means to show technological similarities using NLP.

The structure of this article is as follows. Section 2 introduces the industry covered here and the approach used in our analysis. Section 3 reports the results of our analysis. Finally, we summarize and conclude the analysis in Section 4.

---

[2] Kimura (2020) also showed some characteristics of the technology positions of Chinese firms and their formation processes.

[3] Term frequency–inverse document frequency (TF-IDF) is another way to vectorize documents. It is a method of evaluating the importance of a word based on the frequency of its occurrence.

2. Method

2.1. The Case

In this study, we focus on a Chinese industrial robotics firm as a case study. In the global industrial robotics market, a few global firms hold a large market share, but indigenous firms are expanding their production volumes and sales rapidly in China, where the market is growing rapidly (Cheng et al., 2019). In 2019, approximately 144,000 units of industrial robots were sold in China, with sales consisting of approximately 99,000 units by foreign firms and 44,600 units by indigenous firms (China Robot Industry Alliance, 2020). With market expansion, Chinese firms have increased the number of patent applications they file and are actively engaged in mergers and acquisitions to enhance their competitive advantages. In addition, robots are increasingly used in various applications beyond traditional automotive and electronics production lines at factories, expanding the opportunities to develop new technologies.

Specifically, we compare a fast-growing Chinese firm, SIASUN Robot & Automation Co., Ltd. (hereafter, Siasun) in Shenyang with a Japanese first-mover, YASKAWA Electric Corporation (hereafter, Yaskawa) in Kitakyusyu. Siasun was established in 2000 and belongs to the Chinese Academy of Sciences. [4] Yaskawa, founded in 1915, started its robotics business in the 1970s based on its motor and servo motor business. [5]

Before showing the number of patent applications filed by each firm, we first introduce the Bureau van Dijk (hereafter, BvD) Orbis Intellectual Property (hereafter, Orbis IP) database used in this study. BvD, a Moody's Analytics firm, is a provider of firm information worldwide. Orbis IP is a combination of accounting information, Orbis, and intellectual property information linked by BvD's unique firm codes. Using Orbis IP, we construct an analytical data set as follows. We download patent documents filed by the firms at the patent offices in their home countries. The patent documents are valid as of their download times in August and September 2020.

---

[4] Source: Siasun's official website: http://www.siasun.com/ (accessed on January 25, 2021).

[5] Source: Yaskawa's official website: https://www.yaskawa.co.jp/ (accessed on January 25, 2021).

The patent applications filed by Siasun and Yaskawa based on this data set are summarized in Table 1. The cumulative number of patent applications by Siasun has rapidly increased since the 2010s, and the technological fields covered by their filings have become increasingly diversified over time, as shown in Table 1(a). Moreover, their business has expanded in line with the number of patent applications, as shown in Table 2.[6] Sales have grown at a rapid compound annual growth rate of approximately 19%, and the gross profit rate has also steadily improved. In addition, we find that they have increased their R&D spending at a faster pace than their growth in sales. In contrast, Yaskawa already has many patents and has established the technological fields on which they focus, as shown in Table 1(b). In other words, their technology position has not changed significantly in recent years. Yaskawa's sales in 2019 were 3.8 billion USD, or approximately 10 times more than Siasun's sales, according to Orbis IP. Therefore, by extension, Chinese firms are growing rapidly within the market, but major existing firms are still a significant presence.

[6] However, there is a two-way causal relationship between business expansion and increased R&D expenditures, and there is a time lag for patents to materialize into business advantages. Therefore, care should be taken when discussing the relationship between business performance and the formation of technology positions.

Table 1: Summary of the Data Set

(a) Siasun

		2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Application Number		11	2	13	7	71	31	14	44	41	20	187
Cumulative Number		20	22	35	42	113	144	158	202	243	263	450
A	Human Necessities	0	0	0	0	0	4	5	8	11	12	27
B	Performing Operations; Transporting	7	7	13	18	52	66	70	86	114	127	212
C	Chemistry; Metallurgy	2	2	3	3	7	7	7	7	7	7	9
D	Textiles; Paper	0	0	0	0	0	0	0	0	0	1	1
E	Fixed Constructions	6	6	6	6	6	7	7	7	9	9	11
F	Mechanical Engineering; Lighting; Heating; Weapons; Blasting	0	0	0	0	6	6	6	7	8	8	37
G	Physics	5	7	10	11	23	30	37	50	54	58	97
H	Electricity	0	0	3	4	19	24	26	37	40	41	56

Source: The authors' creation based on the Orbis IP.

(b) Yaskawa

		2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Application Number		498	349	484	546	442	518	265	235	96	118	69
Cumulative Number		5,185	5,534	6,018	6,564	7,006	7,524	7,789	8,024	8,120	8,238	8,307
A	Human Necessities	87	89	89	89	90	107	130	156	159	162	166
B	Performing Operations; Transporting	924	1,027	1,184	1,394	1,550	1,772	1,844	1,906	1,933	1,985	2,013
C	Chemistry; Metallurgy	116	122	124	128	128	128	132	140	141	141	141
D	Textiles; Paper	9	9	9	9	9	9	9	9	9	9	9
E	Fixed Constructions	16	19	23	23	23	23	23	23	23	23	24
F	Mechanical Engineering; Lighting; Heating; Weapons; Blasting	135	143	153	165	175	187	189	194	194	197	198
G	Physics	1,387	1,448	1,511	1,556	1,599	1,674	1,722	1,757	1,766	1,779	1,793
H	Electricity	2,511	2,677	2,925	3,200	3,432	3,624	3,740	3,839	3,895	3,942	3,963

Source: Same as that of Table 1(a).

Table 2: Siasun's Business Performance, 2010–2019

Item (Unit) \ Year		2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Sales	(1,000 USD)	81,392	121,783	162,673	213,015	245,868	256,554	290,020	371,941	446,265	389,064
Gross Profit	(1,000 USD)	25,376	38,926	48,779	75,019	91,887	91,735	96,016	150,857	165,134	144,467
Gross Profit Rate	(%)	31.2	32.0	30.0	35.2	37.4	35.8	33.1	40.6	37.0	37.1
R&D Expenses	(1,000 USD)	761	5,390	5,904	6,171	11,127	9,461	10,159	17,219	21,297	22,214
The Ratio of R&D Expenses to Sales	(%)	0.9	4.4	3.6	2.9	4.5	3.7	3.5	4.6	4.8	5.7
Number of employees	(Persons)	815	1,191	1,479	2,117	2,480	3,097	3,611	4,150	4,513	4,559

Source: Same as that of Table 1(a).

2.2. Preprocessing and Vectorization

Next, we take the following steps for data cleaning and preprocessing of the abstracts in the downloaded patent documents. First, we eliminate signs in the abstracts—e.g., ",", %, and ?—and we lowercase all of the alphabetical characters. In addition, the headings in the abstracts of the patent documents filed in Japan, such as "Problem to be solved" and "Solution," were eliminated because although they are useful for classifying the issues and trends in technology development emphasized by each firm, they do not directly indicate the technological field itself. Second, we use only nouns, verbs, and adverbs and convert them into stems with the natural language toolkit. Then, we vectorize each abstract and make it a 100-dimensional vector with Doc2Vec.

Finally, we calculate the technological similarity between Siasun and Yaskawa. Specifically, we calculate the cosine similarity, s_t, between the cumulative patent applications of Siasun until the year, $patent_t^S$, and those of Yaskawa until the same year, $patent_t^Y$:

$$s_t = similarity\ \ (patent_t^S, patent_t^Y),$$

where t is time. The cumulative patent applications at each year are the accumulation of each component of the document vectors. The similarity here is obtained by the following calculation:

$$s_t = \frac{\sum_{i=1}^{n} S_i Y_i}{\sqrt{\sum_{i=1}^{n} S_i^2}\sqrt{\sum_{i=1}^{n} Y_i^2}},$$

where S_i and Y_i are components of each vector of the patent documents of Siasun and Yaskawa, respectively. It should be noted that the similarity is based on the fraction of each component, so it does not reflect the size of each component. In other words, the number of patent applications does not affect technological similarity. Additionally, the degree of similarity is based on technological fields and features in terms of words. Therefore, the similarity does not consider the quality of patent applications.

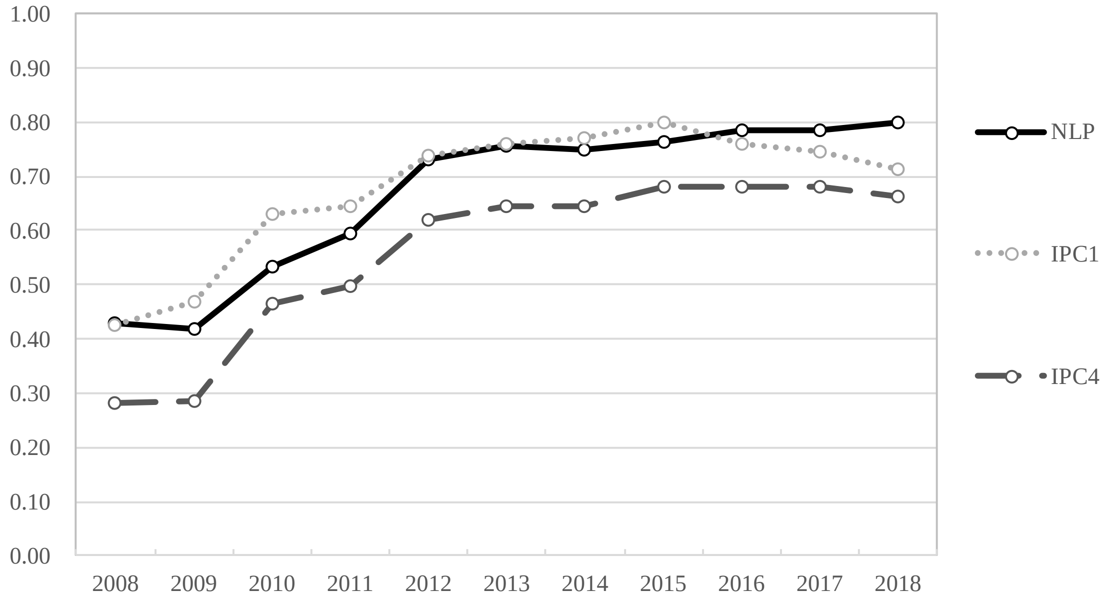
3. Results

3.1. Similarity

Figure 1 shows the technological similarities between Siasun and Yaskawa. It includes not only the similarity based on the vectors obtained by NLP but also the similarities based on the single- and four-digit levels of the IPC (hereafter, IPC1 and IPC4, respectively) to compare the NLP-based similarity. Because the number of technological fields is 8 in IPC1 and 203 in IPC4 in our data set, the number of components in the vectors for IPC1 and IPC4 are also 8 and 203, respectively. [7]

[7] The single-digit level of the IPC originally has a total of eight categories.

Figure 1: Similarities between Siasun and Yaskawa, 2008–2018



Source: The authors' calculation.

The NLP-based similarity shows the technological similarity of Siasun's technology position based on that of Yaskawa. The similarity increases from approximately 0.40 in 2008 to approximately 0.80 in 2018. After a rapid increase in similarity as the number of patent applications begins to increase, the similarity flattens out around the mid-2010s, and there is not much change even when the number of patent applications in each year is relatively high. Even if a primary business of each firm is the industrial robotics business, Siasun is also developing factory automation-related technologies, and Yaskawa has a large business in motors which are a core component of robots. Therefore, as the number of patent applications increases to a certain extent, even if the two firms are in the same industry, the differences in their technological positions become clear.

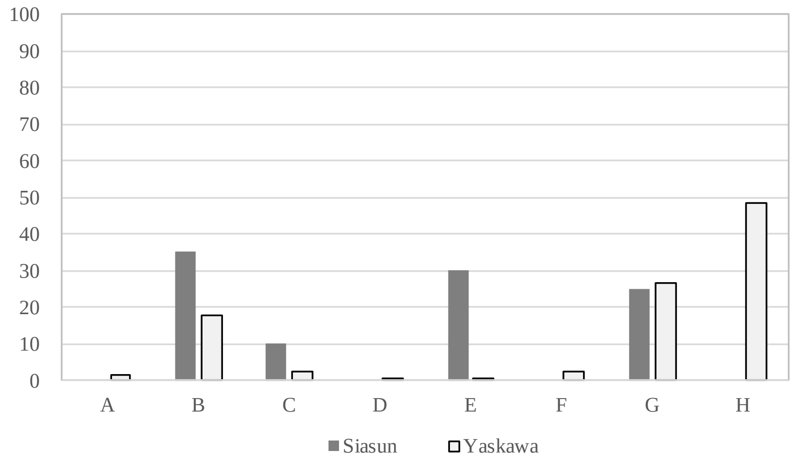
As mentioned in the previous section, the similarity here is based on technological fields, so we also compare the NLP-based similarity with that based on IPC1 and IPC4. Figure 1 shows the trends in similarity among the three different vectors as a whole. However, there are at least two characteristics when we compare the similarities in detail. First, the IPC1-based similarity is consistently greater than the IPC4-based similarity during the observation period. This is likely because the differences in the fraction of patent applications among broader classifications such as those of IPC1 generally tend to be smaller than the differences among narrower ones such as IPC4 for firms in the same industry. Second, the NLP-based similarity was located between the IPC1- and IPC4-based similarities from 2009 to 2015, after which the levels between the IPC1- and NLP-based similarities reversed because the IPC1-based similarity decreased, although that based on NLP did not change much. In other words, it is possible that the differences in technological fields in terms of IPC became greater than the differences in the technologies themselves in terms of the words in the abstracts. [8]

Thus, although there are differences between the three similarities, their trends are broadly similar as mentioned above. To understand the background, we compare the fractions of technological fields based on IPC1 in 2008 with those a decade later. Figure 2 illustrates the fractions of the number of patent applications in Table 1. Both Siasun and Yaskawa had higher fractions in Section B related to manipulators and in Section G related to controls in 2008 and 2018, both of which are major technologies for robots. In contrast, the main difference in 2008 was that Siasun and Yaskawa had higher fractions in Section E related to drilling operations, which were not for robots, and in Section H related to servo motors for robots, respectively. However, over the next 10 years, Siasun did not increase the number of its patent applications in Section E as much as it had previously, but it increased its patent applications in Section H related to servo motors and power supply. As a result, the technological similarity of Siasun to Yaskawa was closer than it had been in 2008.

[8] In addition, changes in the writing style of the abstracts should be accounted for in the NLP analyses.

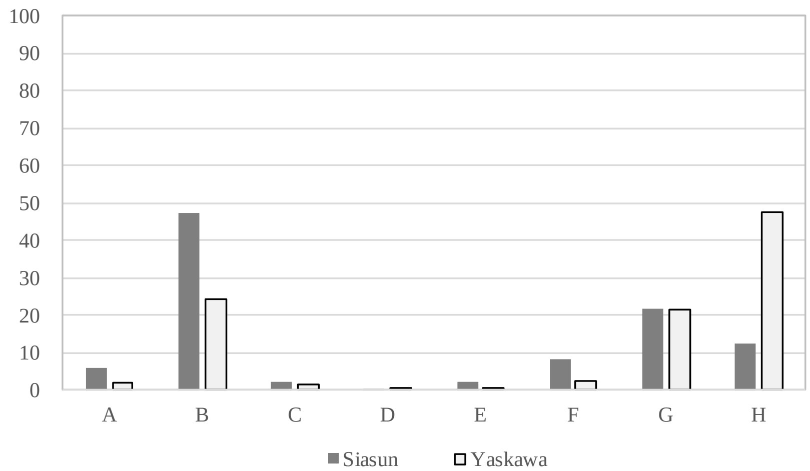
Figure 2: Fractions Based on IPC1 as of (a) 2008 and (b) 2018

(a) 2008



Source: Same as that of Table 1(a).

(b) 2018



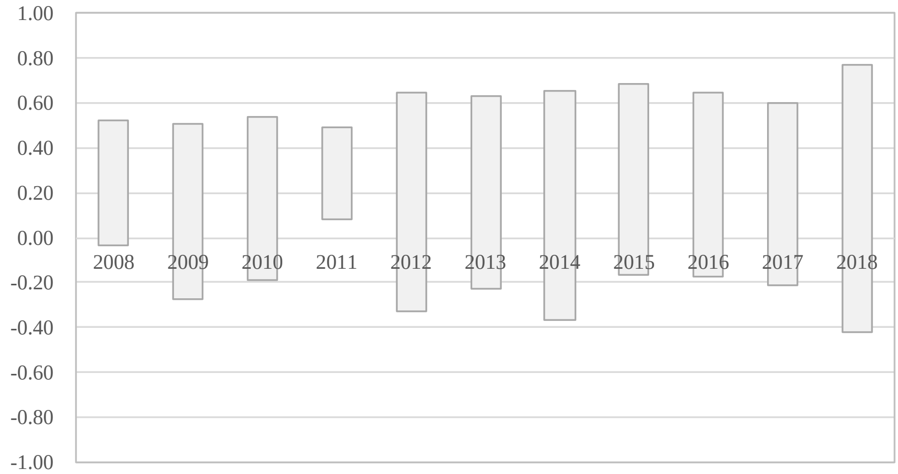
Source: Same as that of Table 1(a).

3.2. Diversity

The technology position of Siasun approaches that of Yaskawa as described in 3.1, but there is also a great deal of diversity among the patent applications. If we calculate and compare the similarities between Siasun's new patent applications in a given year and Yaskawa's technology position in that year, we find a wide range in the similarities. Figure 3 shows the range of maximum and minimum similarities between Siasun's and

Yaskawa's individual patent applications from 2008 to 2018. During the period when the annual number of patent applications was low (until 2011), the difference between the maximum and minimum similarities ranged from 0.46 in 2011 to 0.78 in 2009. In contrast, as the number of patent applications in each year increased, the differences widened from 0.81 in 2017 to 1.12 in 2018. A large number of patent applications tend to develop into a wide variety of technologies.

Figure 3: Range of Maximum and Minimum Similarities, 2008–2018



Source: The authors' calculations.

As we mentioned, however, because firms generally develop and sell a variety of products, we calculate the width only in "B25J9: Programme-controlled manipulators," which has the highest number of patent applications in terms of the IPC. Similar to the trend of all the patent applications taken as a whole, even within the same code of B25J9, the greater the number of patent applications is, the greater the range of similarity. Specifically, it was just 0.19 in 2010 with two applications but 0.69 in 2018 with 17 applications.[9] As technological development in a certain technological field increases, a variety of technologies are developed even within the same field.

[9] The range cannot be calculated for 2011 because there was only one patent application in B25J9.

4. Discussion

We conducted a case study on the technological similarity between Siasun and Yaskawa and showed that as the number of patent applications increases, the technologies of the two firms become more similar in the level of their technological positions but at the same time more diverse in terms of the technologies reflected in their patent documents. Consequently, Siasun can be seen as differentiating and improving the quality of its products by developing a variety of technologies within the main technological fields of robotics. Thus, through this case study on similarity, we were able to understand how a firm in an emerging country develops its technologies in comparison with its competitors.

Therefore, we need to accumulate more cases on the technological similarity among firms in other industries. By doing so, we can extract more general characteristics about the patterns of technological similarities or differences among firms in the process of industrial development. Further research is also needed on the relationship between changes in technological similarity and business performance. This would allow for further discussion of the effective technological development of firms in terms of technology position. Through the study of technological similarities, we will be able to gain a better understanding of the relative market advantages of firms.

References

Bloom, Nicholas, Mark Schankerman, and John van Reenen (2013) “Identifying Technology Spillovers and Product Market Rivalry,” *Econometrica* 81(4): pp. 1347–1393.

Cheng, Hong, Ruixue Jia, Dandan Li, and Hongbin Li (2019) “The Rise of Robots in China,” *Journal of Economic Perspective* 33(2): pp. 71–88.

China Robot Industry Alliance (2020) *The 2020 Market Report of China’s Industrial Robotics Industry* (*2020 Zhongguo Gongye Jiqiren Chanye Shichang Baogao*), Beijing: China Robot Industry Alliance.

Jaffe, Adam (1986) “Technological Opportunity and Spillovers of R & D: Evidence from Firms’ Patents, Profits, and Market Value,” *The American Economic Review* 76(5): pp. 984–1001.

Kimura, Koichiro (2019) "Overseas Expansion and Technological Capabilities: The Case of Chinese Electronics Firms," *Global Journal of Emerging Market Economies* 11(1-2): pp. 119–131.

Kimura, Koichiro (2020) "Technological Specialization: The Case of China's Construction Machinery Industry," in Koichiro Kimura (ed.) *Innovation in East Asia* (*BRC Research Report*), Bangkok: IDE Bangkok Research Center, JETRO Bangkok.

Nagaoka, Sadao, Kazuyuki Motohashi, and Akira Goto (2010) "Patent Statistics as an Innovation Indicator," in Bronwyn Hall and Nathan Rosenberg, eds. *Handbook of the Economics of Innovation*, vol. 2, Amsterdam: North-Holland.