Detecting Technological Originality through Cross-Domain Knowledge in Company

Chun-Chieh Wang[1], Mu-Hsuan Huang[2], Huei-Ru Dong[2]
[1]Graduate Institute of Patent, National Taiwan University of Science and Technology, Taiwan
[2]Department of Library and Information Science, National Taiwan University, Taiwan

Abstract--**Interdisciplinary knowledge flow results the appearance of research front and initiate a serious study on front evolution. Such situation is also evident as novel technologies emerge. Many scholars consider the breakthrough technologies as results of cross-domain knowledge obtained. The technological knowledge diffusion is commonly measured by patentometrics, and cross-domain knowledge obtained is analysed by the spectrum of patent classifications cited. However, the reference cited by patent not only shows patent-type document, but also Non-Patent Reference (NPR) that also presents the knowledge obtained from prior art. The purpose of this study is to advance our understanding of the advantage of employing cross-domain knowledge. U.S. patents acquired by IBM, Philips, and Samsung from 2004 to 2013 are used to identify the relationship between inventive output and cross-domain knowledge obtained. The technological originality of each company is measured by the spectrums of NBER technology field distribution in patent-type reference and ESI journal fields distribution in journal paper cited. Finally, we found Philips performs different patent originality to IBM and Samsung in Computer Hardware & Software, Semiconductor Devices, and Miscellaneous-Elec. fields.**

I. INTRODUCTION

Technological knowledge involves various degrees of specificity, tacitness, complexity, and independence and may differ greatly across technologies. Among these, previous literature mainly focuses on specificity. Generic knowledge refers to the knowledge of a very broad nature, while specific knowledge refers to knowledge specialized and targeted to specific applications. Generic or focused knowledge is also related to different types of science—that is, basic sciences generate generic knowledge, while applied sciences generate focused knowledge [1].

Concerning the nature of technological knowledge, it would be more difficult for latecomer companies to acquire generic knowledge that relates knowledge from very broad sources. On the contrary, it is rather easy to gain specific knowledge that is specialized and targeted to specific applications. However, in some technological sectors, these factors do not always function against technological catch-up. Aggregation from diverse sources is often not a task that the catching-up has to achieve directly, but in many cases, the catching-up economies simply acquire and emulate the result of innovation based on aggregation from diverse sources [12].

Patent citation counts can be used to create measures of the scope, depth, or applicability of an invention. Hall et al. [4] use "backward" citations to assess the degree to which an invention integrates broadly diffuse information (what they term ''originality''). Originality represents a measure of the technological diversity of citations made by the patent, defined by the variety of cited technology classes (''backward'' citations). Thus if a patent cites previous patents that belong to a wide range of fields, its ''originality'' will be high. Originality may be better understood as measures of an invention's interdisciplinary nature [15].

Originality, the diversity of cited patents (backward citations), is a measure of the innovation value of a patent. The measure is based on the distribution (ratio) of cited patents over classes [11]. Originality refers to the breadth of the technology fields on which a patent relies. The Originality measure, first proposed by Trajtenberg et al. [18], operationalises this concept of knowledge diversification and its importance for innovation: inventions relying on a large number of diverse knowledge sources are supposed to lead to original results (i.e. on patents belonging to a wide array of technology fields). Lerner et al. [9] mentions that the originality and generality have been interpreted as measures of the fundamental importance of the research being patented. Patent originality has been used in a wide range of studies, e.g. on the creation of venture-backed start-ups [2]; the duration and outcome of the patent examination procedure at the European Patent Office [5]; and the value of post-merger patents vis-à-vis pre-merger ones [16]. Gress [3] discusses the properties of the USPTO patent citation network and conclusions about the originality in backwards- and forwards-, and intra- and inter-citation networks. Hwang and Kim [6] also adopt the originality indicator to analysis on the multi-technology capabilities of Korea and Taiwan. Layne-Farrar and Lerner [8] measure average originality as patent value to examine patent pool participation and rent sharing rules. Moser and Nicholas [10], Shih et al. [13] [14], and Xie and Giles [19] all are survey competitive advantage of company by patent originality. However, they were not measuring technological originality by both patent reference and journal paper reference cited in patent in those previous studies.

This study tries to advance our understanding of the advantage of employing cross-domain knowledge. U.S. patents acquired by IBM, Philips, and Samsung from 2004 to 2013 are used to identify the relationship between inventive output and cross-domain knowledge obtained. The technological originality of each company is measured by the spectrums of NBER technology field distribution in patent-type reference and ESI journal fields distribution in journal paper cited.

II. METHODOLOGY

A. Patent data collection and technology fields

The empirical data utilized by this study were collected from the United States Patent and Trademark Office (USPTO)

Granted Patent Database. The sample was restricted to utility patents filed by IBM, Philips and Samsung, (top assignee in America, Europe, and Asia) and granted in 2004-2013. The USPTO Patent databases do not have any authority control for assignees' names. Therefore, this study uses authority control to establish unified assignee names, which are collocated with all versions of an assignee patent, even if they were issued under alternative names.

The technology fields used in this study were categorized using the Jaffe et al. [7] classification, which divided the US Patent Classification (USPC) into six technology categories with subcategories. The patent filed by IBM, Philips and Samsung are counted based on technology fields, shown as Table I. The top 3 fieldes of patent count in the three companies are *Communications*, *Computer Hardware & Software*, *Information Storage*, *Semiconductor Devices*, and *Miscellaneous-Elec*. The patent filed by IBM, Philips, and Samsung in these five technology fields are analysed to detect difference of their technological originality.

B. ESI journal fields for journal paper reference

Prior art cited in U.S. utility patent includes patent reference and non-patent reference. Non-patent references covers a diverse set of journal paper, conference paper, book, patent search report, company catalogue, product manual, newspaper, magazine, marketing material, and newsletter etc. We focus on journal paper in NPR and classify them into Essential Science Indicators (ESI) journal fields [17]. Essential Science Indicators categorizes journals into 22 broad disciplines. Each journal is assigned to one of the 22 disciplines. Similarly, Essential Science Indicators then assigns each paper to a discipline—and only one discipline—based on the journal in which it appears. The 22 ESI journal fields are shown as follows:

- Agricultural Sciences
- Biology & Biochemistry
- Chemistry
- Clinical Medicine
- Computer Science
- Economics & Business
- Engineering
- Environment/Ecology
- Geosciences
- Immunology
- Materials Science
- Mathematics
- Microbiology
- Molecular Biology & Genetics
- Multidisciplinary
- Neuroscience & Behavior
- Pharmacology & Toxicology
- Physics
- Plant & Animal Science
- Psychiatry/Psychology
- Social Sciences, general
- Space Science

C. Originality index

Originality index is count of cited patents (backward citations), publication date and IPC class codes for each cited patents. Building on Trajtenberg et al. [18] and on Hall et al. [4], the originality indicator was redefined in this study as:

$$\text{Originality}_p = 1 - \sum_{j}^{n_p} s_{pj}^2$$

where s_{pj} is the percentage of citations made by patent p to technology field (or journal field) j out of the n_p NBER technology fields (or ESI journal fields) contained in the patents cited by patent p. Citation measures are built on USPTO patents. This definition represents the rationale that the broader the technological root of the underlying

TABLE I PATENT COUNT OF COMPANIES IN EACH FIELD

Fields	IBM	Philips	Samsung
1. Chemical	886(1.9%)	224(3.1%)	2814(7.9%)
2. Computers & Communications			
2.1 Communications	3322(7.3%)	**1032(14.2%)**	**6843(19.3%)**
2.2 Computer Hardware & Software	**24015(52.8%)**	**1293(17.7%)**	3906(11.0%)
2.3 Computer Peripherals	536(1.2%)	334(4.6%)	1864(5.3%)
2.4 Information Storage	**4098(9.0%)**	435(6.0%)	**4408(12.4%)**
3. Drugs & Medical	33(0.1%)	446(6.1%)	186(0.5%)
4. Electrical & Electronic			
4.1 Electrical Devices	710(1.6%)	174(2.4%)	1056(3.0%)
4.2 Electrical Connectors	221(0.5%)	26(0.4%)	95(0.3%)
4.3 Electrical Lighting	37(0.1%)	871(11.9%)	917(2.6%)
4.4 Power Systems	726(1.6%)	170(2.3%)	957(2.7%)
4.5 Semiconductor Devices	**5865(12.9%)**	350(4.8%)	**6314(17.8%)**
4.6 Miscellaneous-Elec.	931(2.0%)	**1289(17.7%)**	2336(6.6%)
5. Mechanical			
5.1 Materials Processing & Handling, Metal Working	470(1.0%)	53(0.7%)	503(1.4%)
5.2 Motors, Engines & Parts, Transportation	22(0.1%)	4(0.1%)	85(0.2%)
5.3 Optics	121(0.3%)	170(2.3%)	1912(5.4%)
5.4 Miscellaneous Mechanical	298(0.7%)	38(0.5%)	214(0.6%)
6. Others	3217(7.1%)	380(5.2%)	1035(2.9%)
Total	**45508(100%)**	**7289(100%)**	**35445(100%)**

Border: The top 3 fieldes of patent count in companies

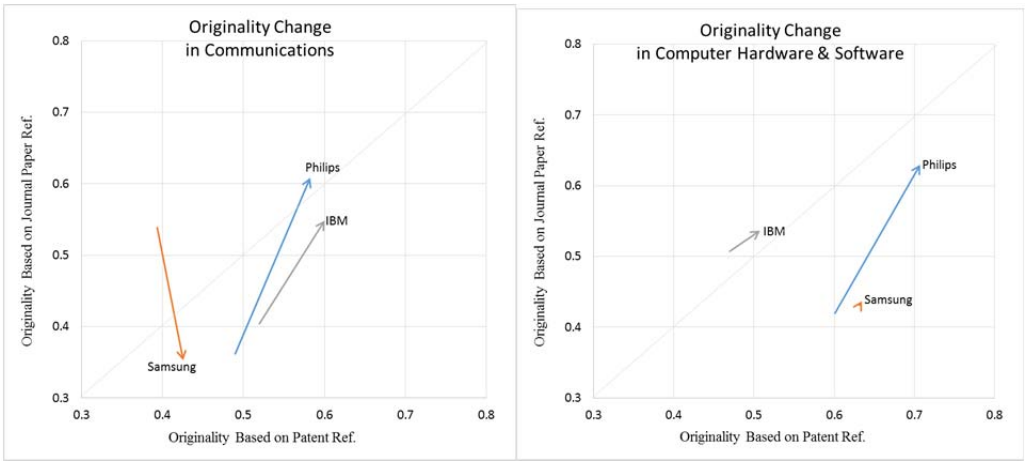
knowledge or research related to the patents, the higher is the originality of a patent. It is probable that the synthesis of divergent ideas is characteristic of research that has high originality and is hence basic in that sense [18]. If a patent cites previous patents that belong to a narrow set of technologies, the originality score will be low, whereas citing patents from a wide range of fields would render a high score. As a proxy for the nature of technological knowledge, we calculate the mean originality of each technology field (or journal field).

III. FINDINGS

A. Comparing Originality changes among IBM, Philips, and Samsung

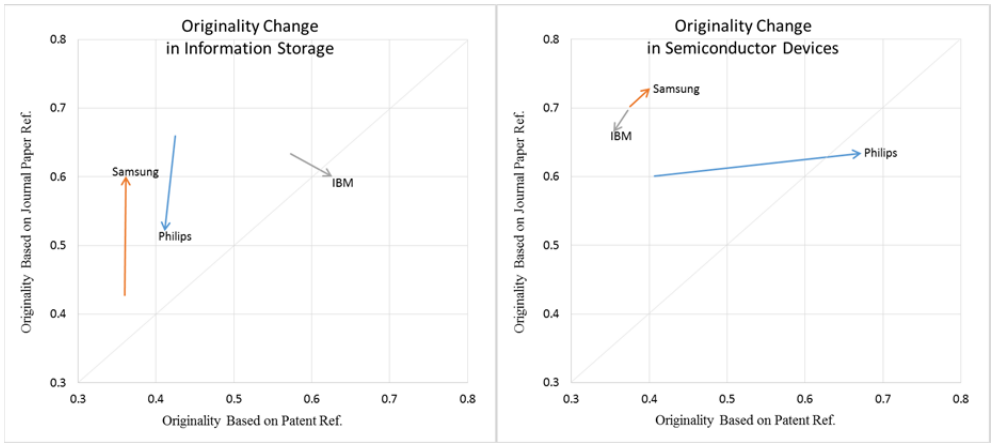
Figures below are the plot of originality based on patent reference versus based on journal paper reference that depicts technological originality in terms of diversity of prior art referred. The NBER technology field and ESI journal fields considered for the plot are from all segments of technology and science. In Fig. 1 (a), it is observed that IBM and Philips have higher Originality values than Samsung in Communications field. From prior (04-08) to posterior (09-13) periods, IBM and Philips both have upgraded their Originality, but Samsung has downgraded his Originality. In Fig. 1 (b), it is observed that Philips has higher Originality value than IBM and Samsung in Computer Hardware & Software field. From prior to posterior periods, Philips has increased his Originality more, but IBM and Samsung have increased their Originality a little.

In Fig. 2 (a), it is observed that Samsung and Philips both have higher Originality based on journal paper reference than based on patent reference in Information Storage field. From prior to posterior periods, Samsung has increased his Originality based on journal paper reference but Philips has decreased his Originality. IBM has similar Originality values based on patent/journal paper reference. In Fig. 2 (b), it is observed that Philips has higher Originality value based on patent reference than IBM and Samsung in Semiconductor Devices field. From prior to posterior periods, Philips has increased his Originality more based on patent reference, but IBM and Samsung have changed their Originality a little.



(a) Communications (b) Computer Hardware & Software

Fig. 1. Originality change in Communications/Computer Hardware & Software



(a) Information Storage (b) Semiconductor Devices

Fig. 2. Originality change in Information Storage/Semiconductor Devices

In Fig. 3, it is observed that IBM, Philips, and Samsung get significant difference in Originality in Miscellaneous-Elec. field. IBM gets higher Originality based on patent reference; Philips gets higher Originality based on journal paper reference; Samsung changes his Originality from higher value based on journal paper reference to higher value based on patent reference. From prior to posterior periods, Samsung has decreased his Originality based on journal paper reference. IBM and Philips have similar Originality values based on patent/journal paper reference.

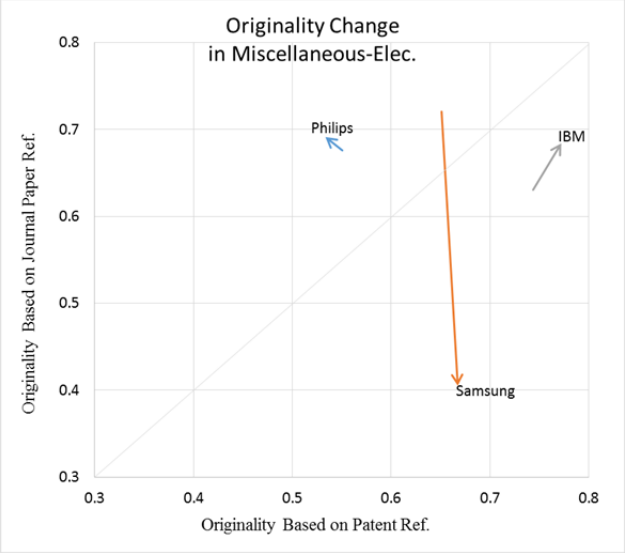


Fig. 3. Originality change in Miscellaneous-Elec.

B. Distributions of patent/journal paper reference

The proportions of patent reference in NBER technology field are measured in Table II, and the proportions of journal paper reference in ESI journal field are measured in Table III. Comparing to Figs. 1-3, we can observe the difference of Originality among IBM, Philips, and Samsung.

In Communications field, IBM and Philips get higher Originality than Samsung, we can observe that patent references of IBM, Philips, and Samsung most are distributed to Communications and Computer Hardware & Software fields. And journal paper references of IBM, Philips, and Samsung most are distributed to Computer Science and Engineering fields. It signifies that the three companies all refer patent references and journal paper references from similar fields, therefore, their communications patents may be in the same technology topics.

In Computer Hardware & Software field, Philips get higher Originality than IBM and Samsung, we can observe that patent references of Philips and Samsung most are distributed to Communications and Computer Hardware & Software fields, but patent references of IBM most are distributed to Computer Hardware & Software and Others fields. IBM referred different patent references to Philips and Samsung. Journal paper references of IBM and Samsung most are distributed to Computer Science and Engineering fields, but patent references of Philips most are distributed to Engineering and Clinical Medicine. Philips referred different journal paper references to IBM and Samsung. It signifies that IBM refers patent references from different field, and Philips refers journal paper references from different field. Communications patent of IBM and Philips may be in different technology topics from Samsung.

In Information Storage field, IBM get higher Originality than Philips and Samsung, we can observe that patent references of IBM, Philips, and Samsung most are distributed to Computer Hardware & Software and Information Storage fields. Journal paper references of Philips and Samsung most are distributed to Engineering and Physics fields, but patent references of IBM most are distributed to Engineering and Computer Science. IBM referred different journal paper references to Philips and Samsung. It signifies that IBM refers journal paper references from different field. Information Storage patent of IBM may be in different technology topics from Philips and Samsung.

In Miscellaneous-Elec. field, IBM, Philips, and Samsung get significant difference in Originality, we can observe that patent references of IBM, Philips, and Samsung most are distributed to Miscellaneous-Elec. and Computer Hardware & Software fields. Journal paper references of IBM, Philips, and Samsung most are distributed to Engineering. Besides, IBM also refers more Computer Science journal papers, Philips also refers more Clinical Medicine journal papers, and Samsung also refers more Physics journal papers. It signifies that Miscellaneous-Elec. patent of IBM, Philips, and Samsung may be in different technology topics.

IV. CONCLUSION

To conclude, this study is preliminary research on Originality based on patent reference and journal paper reference. The major findings are that there is significant different from Philips to IBM and Samsung in Originality performance, especially their patent granted in Computer Hardware & Software, Semiconductor Devices, and Miscellaneous-Elec. fields. These results indicate that the Originality could not be measured only based on patent reference, but the non-patent reference, especially journal paper reference. It can be reasoned that the technology overlapping with alliance partner is needed by company. But it remains unclear the causal relation between technology dependence and alliance relationship.

Although the present study has yielded that findings have practical implications, its design is not without flaws. The technological originality based on journal paper reference is arguably a good proxy for diversity of technological knowledge, but does not capture knowledge diversity with other types of non-patent reference. The other non-patent reference, excluding journal paper, not concerned in this study limits methodology of patent and journal paper reference in the observed knowledge diversity.

TABLE II PROPORTION OF PATENT REFERENCE IN NBER TECHNOLOGY FIELDS

NBER Technology Fields	Comm.[1]			Comp. H&S			IS			Semi.			Misc.-Elec.		
	I[2]	P	S	I	P	S	I	P	S	I	P	S	I	P	S
Chemical	-	-	.01	-	.01	.01	.01	.01	-	**.05**[3]	.06	**.07**	.05	.02	.03
Computers & Communications															
Communications	**.61**	**.68**	**.76**	.07	**.10**	**.14**	.02	.02	.01	.01	.02	.01	.04	.05	.06
Computer Hardware & Software	**.25**	**.13**	**.11**	**.69**	**.57**	**.58**	**.26**	**.08**	**.06**	.01	.01	.01	**.14**	**.08**	**.09**
Computer Peripherals	.01	.01	.01	.01	.03	.04	.01	.01	-	-	-	-	.01	.02	.03
Information Storage	.02	.01	.01	.06	.04	.08	**.57**	**.76**	**.80**	.03	.01	.05	.02	.01	.04
Drugs & Medical	-	.02	-	-	.07	-	-	-	.01	-	.04	-	.01	.06	.01
Electrical & Electronic															
Electrical Devices	.02	.03	.02	.01	.01	.01	.01	.01	.02	.03	.02	.02	.05	.02	.02
Electrical Connectors	-	-	-	-	-	-	-	-	-	-	-	-	.01	-	-
Electrical Lighting	-	.01	-	-	.01	-	-	.01	-	-	**.09**	.01	.01	.02	.01
Power Systems	.01	.01	.01	.01	.01	.01	.01	.01	.01	.02	.03	.02	.04	.01	.02
Semiconductor Devices	.01	-	-	.01	-	.01	.05	.01	.04	**.80**	**.65**	**.78**	.08	.01	.03
Miscellaneous-Elec.	.02	.06	.04	.01	.10	.06	.01	.04	.02	.01	.03	.01	**.45**	**.66**	**.56**
Mechanical															
Materials Processing & Handling, Metal Working	-	-	-	-	-	-	.01	-	.01	.02	.01	.01	.02	-	.01
Optics	.01	-	.01	-	-	.01	-	.02	.01	-	.01	-	.02	.01	.03
Miscellaneous Mechanical	.01	.01	.01	.01	-	.01	-	-	-	-	-	-	.01	-	-
Others	.03	.03	.01	**.12**	.05	.04	.04	.02	.01	.02	.02	.01	.04	.03	.06
Total	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

[1]**Comm.:** Communications; **Comp. H&S:** Computer Hardware & Software; **IS:** Information Storage; **Semi.:** Semiconductor Devices; **Misc.-Elec.:** Miscellaneous-Elec.
[2]**I:** IBM; **P:** Philips; **S:** Samsung.
[3]**Border**: The top 2 fieldes of patent count in the company.

TABLE III PROPORTION OF JOURNAL PAPER REFERENCE IN ESI JOURNAL FIELDS

ESI Journal Fields	Comm.[1]			Comp. H&S			IS			Semi.			Misc.-Elec.		
	I[2]	P	S	I	P	S	I	P	S	I	P	S	I	P	S
Biology & Biochemistry	-	-	-	-	-	.01	-	-	-	-	-	-	-	.02	-
Chemistry	.01	.01	-	-	-	-	-	-	.02	.09	.09	.12	.01	.04	.02
Clinical Medicine	-	-	-	-	**.25**	.01	-	-	-	-	-	-	-	**.28**	-
Computer Science	**.58**	**.23**	**.22**	**.58**	.11	**.21**	**.50**	.19	.11	.05	.05	.02	**.34**	.03	.10
Economics & Business	-	-	-	.01	-	-	-	-	-	-	-	-	-	-	-
Engineering	**.37**	**.65**	**.75**	**.37**	**.57**	**.72**	**.35**	**.34**	**.58**	**.49**	**.34**	**.31**	**.39**	**.44**	**.60**
Materials Science	-	.01	-	-	-	-	.01	-	-	.09	.08	.09	.05	.01	.02
Mathematics	-	-	-	.01	-	.01		-	-	-	-	-	-	.01	-
Multidisciplinary	.01	-	-	.01	-	-	.01	-	.01	.03	-	.07	.01	.01	.05
Neuroscience & Behavior	-	-	-	-	.01	-	-	-	-	-	-	-	-	-	-
Physics	.03	.10	.03	.01	.05	.04	.13	**.47**	**.28**	**.25**	**.44**	**.39**	.20	.16	**.20**
Space Science	-	-	-	.01	.01	-	-	-	-	-	-	-	-	-	.01
Total	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

[1]**Comm.:** Communications; **Comp. H&S:** Computer Hardware & Software; **IS:** Information Storage; **Semi.:** Semiconductor Devices; **Misc.-Elec.:** Miscellaneous-Elec.
[2]**I:** IBM; **P:** Philips; **S:** Samsung.
[3]**Border**: The top 2 fieldes of patent count in the company.

REFERENCES

[1] Breschi, S., F. Malerba and L. Orsenigo; "Technological regimes and Schumpeterian patterns of innovation," *Economic Journal,* vol. 110, no. 463, pp. 388–410, 2000.

[2] Gompers, P., J. Lerner and D. Scharfstein; "Entrepreneurial Spawning: Public Corporations and the Genesis of New Ventures, 1986 to 1999," *The Journal of Finance,* vol. 60, no. 2, pp.577-614, 2005.

[3] Gress, B.; "Properties of the USPTO patent citation network: 1963–2002," *World Patent Information,* vol. 32, no. 1, pp.3-21, 2010.

[4] Hall, B.H., A. Jaffe and M. Trajtenberg; "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools", *NBER Working Papers No. 8498*, National Bureau of Economic Research, Inc., 2001.

[5] Harhoff, D. and S. Wagner; "The Duration of Patent Examination at the European Patent Office," *Management Science,* vol. 55, no. 12, pp.1969-1984, 2009.

[6] Hwang, J.T. and B.K. Kim; "Analysis on the multi - technology capabilities of Korea and Taiwan using patent bibliometrics," *Asian Journal of Technology Innovation,* vol. 14, no. 2, pp.183-199, 2006.

[7] Jaffe, A. B., M. Trajtenberg and P. M. Romer; *Patents, citations, and innovations: A window on the knowledge economy*, The MIT Press, 2005.

[8] Layne-Farrar, A. and J. Lerner; "To join or not to join: Examining patent pool participation and rent sharing rules," *International Journal of Industrial Organization,* vol. 29, no. 2, pp.294-303, 2011.

[9] Lerner, J., M. Sorensen and P. Strömberg; "Private Equity and Long-Run Investment: The Case of Innovation," *The Journal of Finance,* vol. 66, no. 2, pp.445-477, 2011.

[10] Moser, P. and T. Nicholas; "Was electricity a general purpose technology? Evidence from historical patent citations," *American Economic Review*, pp.388-394, 2004.

[11] Narin, F.; *Tech-Line Background Paper*. CHI Research, Inc., Haddon Heights, NJ, 2000.

[12] Park, K.H. and K. Lee; "Linking the technological regime to the technological catch-up: analyzing Korea and Taiwan using the US

patent data," *Industrial and Corporate Change,* vol. 15, no. 4, pp.715-753, 2006.

[13] Shih, M.-J., D.-R. Liu and M.-L. Hsu; "Discovering competitive intelligence by mining changes in patent trends," *Expert Systems with Applications,* vol. 37, no. 4, pp.2882-2890, 2010.

[14] Shih, M.-J., D.-R. Liu and M.-L. Hsu; "Mining changes in patent trends for competitive intelligence," *Advances in Knowledge Discovery and Data Mining* ,pp.999-1005: Springer, 2008.

[15] Squicciarini, M., H. Dernis and C. Criscuolo; *Measuring Patent Quality: Indicators of Technological and Economic Value*, OECD Science, Technology and Industry Working Papers, 2013/03, OECD Publishing, 2013.

[16] Stahl, J. C.; "Mergers and Sequential Innovation: Evidence from Patent Citations," *Finance and Economics Discussion Series*, No. 2010-12. Division of Research and Statistics and Monetary Affairs Federal Reserve Board, Washington, DC, 2010.

[17] Thomson Reuters; Journal List For Essential Science Indicators, SCIENCEWATCH, 2013, Retrieved January 20, 2014, from http://sciencewatch.com/info/journal-list

[18] Trajtenberg, M, A. Jaffe and R. Henderson; "University versus Corporate Patents: A Window on the Basicness of Inventions," *Economics of Innovation and New Technology,* vol. 5, no.1, pp.19-50, 1997.

[19] Xie, Y., and D.E. Giles; "A survival analysis of the approval of US patent applications," *Applied Economics,* vol. 43, no. 11, pp.1375-1384, 2011.