Research on the impact of cross domain fusion of patent technology on patent value--Based on propensity score matching method

Ziyan Meng[1]

[1]School of Business, Shandong Normal University, Jinan, China

E-mail address: m806459168@163.com

Abstract. Technological innovation has become the main driving force for economic and social development, and also the focus of competition between countries or enterprises. This paper uses the propensity score matching method (PSM) in causal reasoning to study and analyze the impact of cross domain fusion of patent knowledge and technology on changes in patent value. Specifically, this paper selects patent data from the "Chemistry and Metallurgy" field of the China Intellectual Property Office from 2011 to 2020, and analyzes the frequency of patent citation as a substitute variable for the value of patent technology. The experimental results indicate that cross domain fusion of patent technology has a significant positive effect on improving patent value and is a related factor in increasing the frequency of patent citations; And the processing effect of multi domain fusion is significantly better than that of dual domain fusion. Therefore, cross domain integration of technology can increase the frequency of patent citations and ultimately enhance patent value. Our research can capture the causal relationship between the cross domain phenomenon of patent technology and the frequency of patent citations, and provide corresponding suggestions for promoting knowledge innovation.

Keywords: Patented technology; Cross disciplinary; Propensity score matching; Knowledge fusion

1 Introduction

1.1 Background and purpose

In the context of knowledge economy and economic globalization, intellectual property has become a strategic resource to enhance the economy and the core element to enhance the international competitiveness [1]. Whoever owns more patents, especially more high-quality patents, can gain more share in the market and gain more advantages and initiative in the competition. Under this background, it is particularly important to study and analyze the influencing factors of patent value, which directly affects the improvement of the quality of knowledge results.

G. Guan et al. (eds.), *Proceedings of the 2023 3rd International Conference on Education, Information Management and Service Science (EIMSS 2023)*, Atlantis Highlights in Computer Sciences 16,
https://doi.org/10.2991/978-94-6463-264-4_4

In the research and analysis of patents, Narin and Harhoff believe that a patent with a high degree of technological innovation tends to be frequently cited by subsequent technologies and its citation frequency is higher [2]-[3]. Y Yoshikane et al. believe that the patent citation frequency can be used to reflect the importance of its own technology[4] ; Xiao Guohua and others pointed out that institutions with high cited patents are superior to their competitors or peers[5]; In addition, some studies also agree that the patent citation frequency as an index to evaluate the patent quality[6]. Therefore, this paper uses the citation frequency of the patent as the surrogate variable of the patent value. The more frequently a patent is cited, the higher its technical value and the more it can promote the transformation of patented knowledge achievements.

Subject integration has gradually become the mainstream of technological innovation. As the core of intellectual property, patent is difficult to play a higher technical value if it is supported by knowledge in a single field. On the contrary, when Montresor and other scholars studied green technology, they found that the "hybrid" and correlation reorganization of green and non-green knowledge technology seemed to "obtain new ecological solutions for the region better than the" pure " green technology[7]. Therefore, this paper on cross-field patent integration can effectively realize the intersection of subject knowledge, make the more influential patents tend to be frequently cited by subsequent technologies.

1.2 Literature review

In the study of the cross-domain phenomena of patents or papers, Chen Shiji and Qiu Junping Using scientific measurement measures, typical integration indicators (Rao-Stirling Index (RS) and Leinster-Cobbold Diversity Index (LCDiv)) to examine the differences between subjects of all papers published in Web of Science, the results show that, Highly cited papers always show high subject diversity and great disparity [8]; Steele and Stier found through least squares regression analysis that papers with higher interdisciplinary relevance have higher impact[9]; The study by F.Yoshikane et al is based on the Japanese patent literature published in 1998, The results show that the diversity of Japanese patent citation classification is significantly positively correlated with its citation frequency[4]. It can be seen that most scholars study the impact of interdisciplinary on the value of journal articles, only a few scholars are involved in the field of patents. Also, their empirical studies are less and their conclusions are inconsistent. Thus, more empirical researches are needed.

This may be due to different methods and data used, but the biggest reason should be the presence of external factors that affect the experiment. To reduce the error caused by external factors, this paper intends to take patent data from Chinese invention applications as an example, and use the propensity score matching method (PSM) in causal inference to give the probability of the patent receiving cross-domain treatment under the condition that the control confounding factors are similar, that is, the propensity score, and then the treatment group and the control group are matched to achieve the purpose of controlling confounding bias. At this point, if there is a difference in outcomes between the treatment and control groups, the difference can be attributed entirely to the impact of cross-domain integration.

2 Data and model

2.1 Sample data

We used simple random sampling to select sample data from the population data with a final sample size of 223,522. Based to the convergence of the patented technology, we designed two sets of PSM experiments (**Table 1**).

Table 1. Sample overview of five PSM trials

Trails	Group	Classification	Number of patents
PSM 1	Treated	Patents for cross domain technology integration	85878
	Untreated	Patents for non cross domain integration of technology	137644
PSM 2	Treated	Patents for technology integration across multiple fields	7431
	Untreated	Patents for cross domain integration of technology	78447

The variables used in this study include processing variables, outcome variables, and covariates (**Table 2**). For covariates, we selected quantitative factors that affect the frequency of patent citations from published research literature as the majority of covariates, while adding two latent variables obtained from the dimensionality reduction of patent applicant feature variables. Suitable covariates can greatly improve the fit and interpretation rate of the model [10].

Table 2. Covariate set

Variable	Meaning	Max	Min	Mean	Sd
Number of claims(X1)	Number of requests for protection of patent rights	1	84	8.24	33.725
Patent title number(X2)	Number of patent title words	0	134	18.23	47.302
Document pages(X3)	Number of pages of application documents	1	300	11.99	211.175
Length of patent publication(X4)	Time from publication of patents to present	2	11	6.12	6.811
Number of inventors(X5)	Number of patented inventions	1	42	3.93	7.307
Number of citations(X6)	Number of cited applied patents	0	25	3.71	7.876
Number of patent families(X7)	Number of patents in the same family	1	225	2.67	19.135
Latent Variable (FAC1 and FAC2)	Common factor 1	-0.35674	10.97358	0.0009065	1.005
	Common factor 2	-0.56542	2.53456	-0.0037135	0.993

The selected set of covariates was then assessed for multicollinearity. The results are shown in **Table 3**. Therefore, it can be considered that the collinearity of variables can be ignored.

Table 3. Multicollinearity diagnosis of covariates

Dimension	Eigenvalue	Condition indicators	Variance ratio								
			X1	**X2**	**X3**	**X4**	**X5**	**X6**	**X7**	**FAC1**	**FAC2**
1	5.844	1.000	0.01	0.01	0.00	0.00	0.01	0.01	0.01	0.00	0.00
2	1.031	2.381	0.00	0.01	0.00	0.00	0.01	0.00	0.03	0.80	0.02
3	1.014	2.401	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.88
4	0.842	2.635	0.01	0.09	0.01	0.01	0.02	0.03	0.21	0.14	0.01
5	0.347	4.107	0.00	0.14	0.01	0.01	0.06	0.61	0.18	0.01	0.01
6	0.332	4.195	0.02	0.37	0.09	0.00	0.00	0.17	0.36	0.00	0.00
7	0.245	4.880	0.04	0.01	0.08	0.01	0.79	0.00	0.11	0.02	0.05
8	0.184	5.633	0.80	0.38	0.06	0.00	0.01	0.05	0.01	0.00	0.00
9	0.126	6.819	0.06	0.00	0.28	0.60	0.08	0.03	0.06	0.00	0.00
		VIF	1.559	1.620	1.091	1.069	1.122	1.050	1.542	1.042	1.063
		1/VIF	0.641	0.617	0.916	0.936	0.892	0.952	0.648	0.960	0.941

2.2 PSM

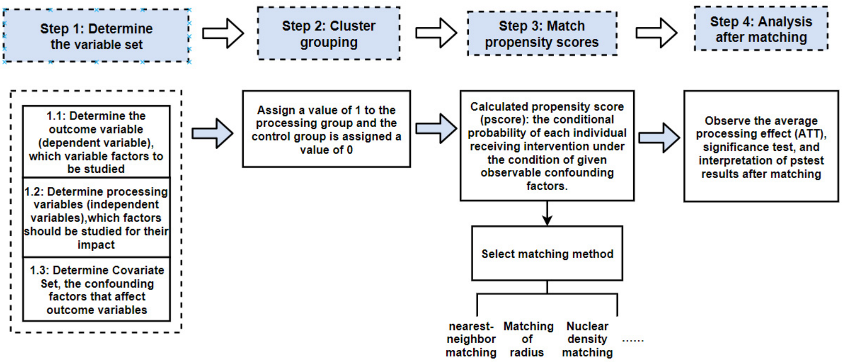

Fig. 1. Steps of propensity score matching method

The algorithm steps of the propensity score matching method (PSM) used in this parper are shown in **Fig. 1**, which can be roughly divided into four steps.

3 Empirical analysis

By performing one-on-one K-nearest neighbor propensity score matching on the samples of the clustered processing group and the control group. The results are shown in **Table 4**. This indicates that cross domain integration of technology can increase the average citation frequency of patents by 0.2642 times; The integration of technology across multiple fields can increase the average citation frequency of patents by 0.5019 times, and the results are all significant at the 1% level.

Table 4. Estimated results

Trails	Matching method	ATT				
		Treated	Un-treated	Difference	S.E.	T-stat
PSM 1	K-nearest neighbor matching (K=1)	4.7294	4.4652	0.2642***	0.0312	8.48
PSM 2	K-nearest neighbor matching (K=1)	4.9857	4.4839	0.5019***	0.0944	5.32

Results of the balance tests are shown in **Table 5** and **Fig. 2**. Matching can effectively balance the two sets of sample data, significantly reduce the standardized deviation of variables, and make the distribution as close as possible.

Table 5. Balance test results

variable	Before matching				After matching				%bias (reduct)
	Treated (mean)	Un-treated (mean)	% Bias	t	Treated (mean)	Un-treated (mean)	% Bias	t	
X1	9.1581	7.667	25.1	59.51	9.1573	9.2076	-0.8	-1.52	96.6
X2	19.257	17.584	24.1	56.34	19.257	19.627	-5.3	-10.32	77.9
X3	15.036	10.082	32.3	79.49	15.027	14.299	4.8	7.94	85.3
X4	5.9868	6.2067	-8.4	-19.39	5.9688	5.8775	4.2	8.61	50.4
X5	4.0516	3.8552	7.3	16.72	4.0516	4.0431	0.3	0.65	95.7
X6	3.916	3.5776	12.0	27.78	3.9161	4.0016	-3.0	-6.03	74.7
X7	3.3169	2.2739	22.8	55.21	3.3127	3.2316	1.8	3.14	92.2
FAC1	0.03327	-0.01929	5.2	12.06	0.03329	0.05715	-2.4	-4.63	54.6
FAC2	-0.04434	-0.03369	7.8	18.02	0.04432	0.05379	-0.9	-1.91	87.9

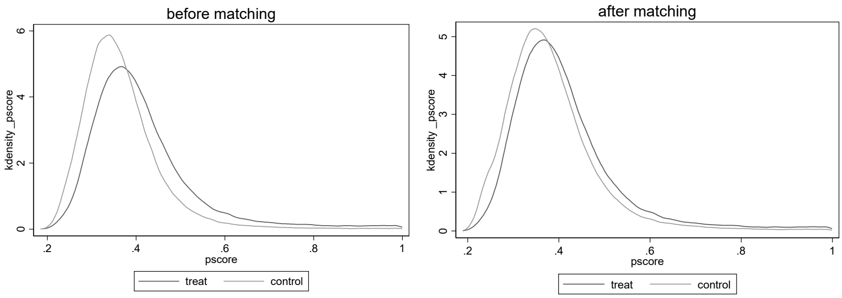

Fig. 2. Kernel density function before and after PSM 1 matching

The other five propensity score matching methods were used to test the robustness of the estimated results of the two groups (**Table 6**). The results were similar to those of one-to-one K-nearest neighbor matching, and all passed the t-test. The estimated results were significant.

Table 6. Results of the other five matching methods

Trails	Matching method	ATT				
		Treated	Un-treated	Differ-ence	S.E.	T-stat
PSM 1	K-nearest neighbor matching (K=4)	4.7294	4.4781	0.2513***	0.0267	9.43
	Radius matching (cal = 0.01)	4.7294	4.5078	0.2217***	0.0245	9.03
	Local linear regression matching	4.7294	4.4669	0.2625***	0.0312	8.42
	Markov matching (K=4)	4.7299	4.5422	0.1878***	0.0247	7.59
	Nuclear density matching	4.7294	4.5185	0.2109***	0.0244	8.66
PSM 2	K-nearest neighbor matching (K=4)	4.9857	4.5032	0.4825***	0.0794	6.08
	Radius matching (cal = 0.01)	4.9857	4.5148	0.4709***	0.0722	6.52
	Local linear regression matching	4.9857	4.5513	0.4344***	0.0944	4.60
	Markov matching (K=4)	4.9845	4.5507	0.4338***	0.0735	5.90
	Nuclear density matching	4.9857	4.6203	0.3655***	0.0720	5.07

4 Conclusions

This paper studies the impact of patent technology's cross domain situation on patent value from the perspective of the classification number and causal relationship of patent technology itself. Specifically, we have found that: (1) patents with technology cross domain integration are more likely to receive more citations than patents with non technology cross domain integration, which can more promote the increase in patent value. This is mainly because interdisciplinary integration can better promote patent innovation and empower technological innovation. (2) Among patents with the phenomenon of technology cross domain integration, patents with technology cross domain integration are cited more frequently than patents with technology cross single domain integration.

From the above empirical results, it can be seen that the cross disciplinary situation and quantity of patented technology can serve as factors affecting the positive increase in patent citation frequency, and can predict the technical value and importance of knowledge achievements of a certain patent when combined with other factors.

Due to the limitations of the experimental samples in terms of country, field, and time, the research conclusions drawn in this paper do not yet have universality, and further research is needed to determine whether it can be extended to other fields.

5 References

1. Wan Xiaoli. (2009) Research on Patent Quality Indicators and Their Functions in the Performance Evaluation of Intellectual Property Strategy Implementation. Science and Technology Management, 30(11):69-74. https://kns.cnki.net/kns8/defaultresult/index.
2. Narin, F. (1995) Patent as Indicators for the Evaluation of Industrial Research Output. Scientometrics, 34(3) :489-496. DOI:10.1007/BF02018015.
3. Harhoff D, Narin F, Scherer FM, et al. (1999) Citation Frequency and the Value of Patented inventions. The Review of Econmics and Statistics, 81(3):511-515. DOI:10.2307/2646773.
4. Yoshikane, Fuyuki, Suzuki, Yutaka, Tsuji, Keita, et al. (2011) Analysis of the relationship between citation frequency of patents and diversity of their backward citations for Japanese patents. Scientometrics, 92(3):721-733. DOI:10.1007/s11192-011-0612-6.
5. Xiao Guohua, Wang Chun, Jiang He. (2008) Design and construction of patent analysis and evaluation index system. Library and Information Work,52(2):96-99. https://kns.cnki.net/kns8/defaultresult/index.
6. Ma Tingcan, Li Guiju, Jiang Shan. (2012) Patent quality evaluation indicators and their application in patent measurement. Library and Information Work, 56(24):89-95,59. https://kns.cnki.net/kns8/defaultresult/index.
7. Montresor Sandro, Quatraro Francesco. (2020) Green technologies and Smart Specialisation Strategies: a European patent-based analysis of the intertwining of technological relatedness and key enabling technologies. Regional Studies, 54(10): 1354-1365. DOI: 10.1080/00343404.2019.1648784.
8. Chen Shiji, Qiu Junping, Arsenault Clément, Larivière Vincent. (2021) Exploring the interdisciplinarity patterns of highly cited papers. Journal of Informetrics, 15(1). DOI: 10.1016/j.joi.2020.101124.
9. Steele TW, Stier J C. (2000) TheImpact of Interdisciplinary Researchin the Environmental Sciences: a Forestry Case Study. Journal ofthe American Society for Informa- tion Science, 51(5):476-484. DOI:10.1002/(SICI)1097-4571(2000)51: 53.0.CO;2-G.
10. STEINER P M, COOK T D, SHADISH W R. (2011) On the Importance of Reliable Covariate Measurement in Selection Bias Adjustments Using Propensity Scores. Journal of Educational & Behavioral Statistics, 36(2): 213-36. DOI:10.3102/1076998610375835.