ISSN:2790-1688 DOI: 10.56028/aetr.4.1.148.2023

The Influence of Prince on Sleeping Time: An Empirical Study Based on Propensity Score Matching

Xiao Zheng [1], Yinghong Ma [1,*]

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

* yinghongma71@163.com

Abstract. With the rapid development of science and technology, it is of great significance to study the influence of prince on sleeping time to explore the knowledge value of sleeping beauty. In this work, we use the data of American Physical Society to identify sleeping beauties according to B-index, and use logistic model and propensity score matching model to analyze the data. Then, we explore the impact of prince on the sleeping time of sleeping beauty. The results show that the prince has a significant inhibitory effect on the sleeping time of sleeping beauty, and the sleeping time of sleeping beauty cited by prince is shorter. The citation behavior of prince has the feature of self-selection. Using propensity score matching method can effectively overcome the problem of sample selection bias and obtain more reliable estimation results.

Keywords: Sleeping Beauty; sleeping time; Prince; Propensity Score Matching.

1. Introduction

With the mature use of knowledge databases and modern retrieval tools, the increasing availability of publication information, and the increasing standardization of scientific research, the research based on publication citations is also increasing. The researches find that most of papers are cited quickly after publication, but there are also some important scientific studies that are re-recognized many years after publication. Barber proposed the term “resisted discovery” to explain the above phenomenon[1]. Later, Garfield. E put forward the term “delayed recognition” and suggested using citations to identify papers with delayed recognition[2]-[5]. Gl änzel estimated the occurrence of delayed recognition and emphasized some common features in the papers on delayed recognition[6]. The proposal of “Sleeping Beauty”(SB) was put forward in 2004. Van Raan replaced “delayed recognition” with this vivid name, gave the definition of SBs for the first time, and defined the “Prince” as paper that first quoted SBs[7].

The existing paper pay less attention to the influence of prince’s citation on the sleeping time of SBs, and mainly focuses on the definition and identification of prince, which is conducive to study the awakening mechanism of SBs. Ohba proposed three characteristics to define the prince[8]. Zong believed that prince and SBs have quite a number of co-citations, and the SBs without prince can also wake up[9]. Fang found that a SB may be awakened by several different princes[10].

There are few studies in academic circles on the quantitative analysis of the influence of prince’s citation on the sleeping time of SBs. Therefore, this work uses the data of American Physical Society (APS) from 1893-2016 to control the influence of sample selection bias by using the propensity score matching method, and deeply examines the effect of prince on the sleeping time of SBs, which in order to give full play to the awakening effect of prince on SBs, and help to explore the knowledge value of SBs.

2. Data

2.1 Data source and sample selection

Using the data of American Physical Society (APS) from 1893 to 2016, this work calculates the B-index of all papers and arranges them in descending order. The first 1% papers are defined as SBs[11]. The B-index of a paper is calculated by

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.148.2023

$$B = \sum_{t=0}^{t_m} \frac{\frac{c_{tm} - c_0}{t_m} \cdot t + c_0 - c_t}{max\{1, c_t\}} \quad (1)$$

where t represents the length of time after paper is published, c0 is the number of citations in the year of publication, ctm is the maximum number of citations, and tm is the year when paper reaches the maximum number of citations. Since this study explores the influence of prince on the sleeping time of SBs, a total of 4710 study samples are obtained after removing the missing values of relevant variables.

2.2 The definition of variables

There are three variables in this study: result variable, treated variable and control variable. First, the result variable selected in this work is the sleeping time of SBs. Second, this work takes whether the SB is cited by prince as treated variable, and sets the SBs cited by prince as treated group, and the SBs not cited by prince as control group. Third, based on the existing research and relevant paper, this work selects academic age, the number of coauthors, the influence of authors, the frequency of citations, the quality of references, the reputation of journals and the volume of journals as control variables. See TABLE I and TABLE II for the definition of variables and descriptive statistics.

2.3 Research method

Stata is used for statistical analysis, and the relationship between prince and the sleeping time of SBs is estimated using logistic model. Because the conventional logistic estimation model cannot eliminate the endogenous nature of control variables, and it is difficult to obtain accurate results. This study further uses the propensity score matching method (PSM) to control potential confounding factors, mainly to explore the impact of prince on the sleeping time of SBs. The treated group is matched with control group to obtain the treatment effect of prince on the sleeping time of SBs.

Table 1. The Definition Of Variables

Variables		Symbol	Definition
Result variable	sleeping time	Tsl	The length of time from the publication of SBs to awakening
Treated variable	prince	prince	Whether SB is cited by prince: Yes = 1, No = 0
Control variable	academic age	age	The length of time the first author spent publishing the first to last paper
	the number of coauthors	coauthor	The number of cooperators with the first author
	the influence of authors	influence	The average citation of papers published by first author
	the frequency of citations	citation	The number of citations from paper
	the quality of references	reference	The average citation of references
	the reputation of journals	reputation	The normalization of journal rankings
	the volume of journals	volume	The cumulative amount of papers published in journal

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.148.2023

Table 2. Descriptive Statistics

Variables	Mean	Std	Max	Min
Tsl	30.007	14.176	116.000	1.000
prince	0.486	0.500	1.000	0.000
age	26.641	18.191	110.000	1.000
coauthor	7.998	10.939	103.000	0.000
influence	10.190	21.269	785.500	0.000
citation	28.076	76.181	3087.000	0.000
reference	85.549	175.785	5883.000	0.000
reputation	0.506	0.313	0.909	0.091
volume	56983.270	36553.240	171143.000	124.000

3. Empirical Analysis

3.1 Propensity score estimation

In order to match the SB cited by prince or not, this work uses the logistic model to estimate the impact of prince on sleeping time, and the estimated results are shown in TABLE III. The research results show that the logistic model has a good fitting effect. The number of coauthors, the frequency of citations, the quality of references, the reputation of journals and the volume of journals have a significant positive impact on the possibility of prince citing SB.

Table 3. Logistic Model Estimation Results Of The Influence Of Prince On Sleeping Time

Variables	Coefficients	Std	z	P
age	-0.0005	0.0011	-0.42	0.671
coauthor	0.0039	0.0019	2.07	0.038
influence	0.0017	0.0012	1.39	0.166
citation	0.0036	0.0005	7.30	0.000
reference	0.0005	0.0001	3.31	0.001
reputation	0.1865	0.0610	3.06	0.002
volume	3.83e-6	5.31e-7	7.21	0.000
Constant	-0.5143	0.0522	-9.86	0.000
Pseudo R2	0.033			

3.2 Examination of matching results

The matching quality directly affects final estimation results. An ideal match can satisfy both the common support condition and the balancing hypothesis. The distribution of propensity scores of the SBs in treated group and control group after meeting the common support hypothesis need matching is as close as possible, while the characteristics of treated group and control group after meeting the balance hypothesis need matching are balanced and there is no significant difference.

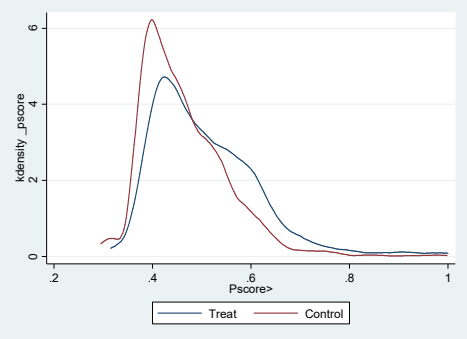
In this work, the radius matching method is selected for matching. The kernel density functions before and after matching are shown in Fig. 1. Fig. 1 shows that there is a significant difference in the probability distribution of two groups' propensity scores before matching, and the distribution of two groups' propensity scores after matching is very close, indicating that the gap between the two groups of SBs have been significantly narrowed, and the matching effect is better. Fig. 2 intuitively shows the common value range of propensity score. It is found that most of observed values are within the common value range. Only a small number of samples will be lost in the propensity score matching, and the common support condition can be satisfied.

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.148.2023

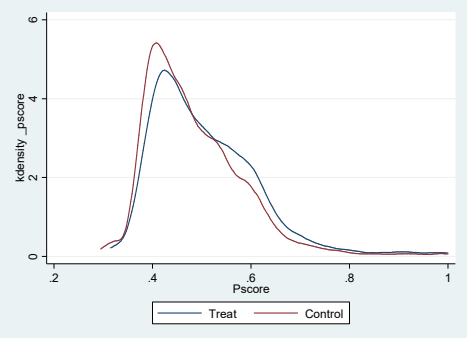
Next, we check the balancing hypothesis[12]. It can be seen from TABLE IV that the reduct bias of each matching variable in treated group and control group is less than 20% after matching, and it can be considered that the standardization deviation after matching is small enough[13]. To sum up, matching significantly reduces the difference of matching variables between treated group and control group, minimizes the error of sample self-selection, satisfies the balance hypothesis, and the sample matching is relatively successful.

3.3 Matching effect analysis

TABLE V shows the estimated results of the average treated effect (ATT) of prince on sleeping time. It can be seen that the estimation result after using the propensity score matching method to alleviate the potential sample selection bias shows that ATT is negative and passes the significance test at the level of 1%. The prince's citation significantly reduces the sleeping time of SBs. If the SB is not cited by prince, its sleeping time is 30.221. However, due to the introduction of prince, the sleeping time of SBs is reduced to 29.205, which decreased by 1.015. Therefore, the prince has a significant negative impact on the sleeping time of SBs. The treated effect estimated by propensity score matching method is lower than that estimated before matching, which indicates that there is a selective bias in the citation behavior of prince. It is necessary to analyze the influence of prince on sleeping time by the propensity score matching method.

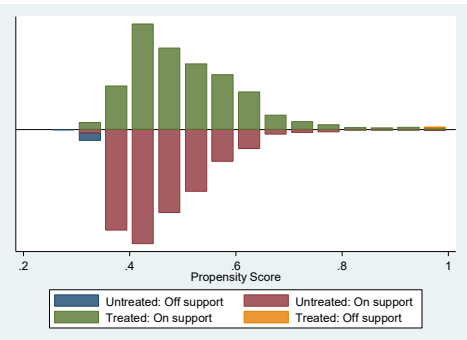


(a) before matching



(b) after matching

Figure 1. Kernel density function.



ISSN:2790-1688 DOI: 10.56028/aetr.4.1.148.2023

Figure 2. Common value range of propensity score.

Table 4. The results of balancing hypothesis

Variables	Sample	Mean		%bias	%reduct bias
		Treated	Control		
age	Unmatched	27.184	26.128	5.8	53.3
	Matched	27.171	26.677	2.7	
coauthor	Unmatched	8.582	7.447	10.4	48.7
	Matched	8.584	8.002	5.3	
influence	Unmatched	11.953	8.523	16.1	39.4
	Matched	11.303	9.223	9.8	
citation	Unmatched	36.922	19.713	22.5	37.3
	Matched	33.314	22.528	14.1	
reference	Unmatched	101.620	70.356	17.7	33.7
	Matched	98.157	77.442	11.7	
reputation	Unmatched	0.527	0.486	12.9	55.6
	Matched	0.527	0.509	5.7	
volume	Unmatched	61930	52306	26.5	37.4
	Matched	61862	55839	16.6	

Table 4. The Estimated Results Of ATT

Method	Treated	Control	ATT	t
Unmatched	29.236	30.736	-1.500	-3.63
Radius	29.205	30.221	-1.015	-2.47

4. Conclusion And Suggestion

4.1 Conclusion

This work takes 4710 SBs in the field of physics as research objects, and uses the propensity score matching method to empirically explore the influence of prince on sleeping time, and draws the following conclusions.

The prince's citation has a significant promoting effect on shortening the sleeping time of SBs. In terms of the overall effect, if the SB is not cited by prince , the sleeping time is 30.221, and the sleeping time of SB that cited by prince is 29.205, and the ATT is -1.015.

The citation behavior of prince to SBs has the feature of self-selection. Ignoring the problem of self-selection will lead to biased parameter estimation. However, using the propensity score matching method can effectively overcome the problem of sample selection bias and obtain more reliable estimation results.

Academic age can shorten the sleeping time of SBs. This study shows that the influence coefficient of academic age on sleeping time is negative, and the sleeping time of SBs published by the author with old academic age is shorter than that of the author with young academic age. Older academic authors have been active in their research field for a long time. They have a certain reputation and have a large number of collaborators. Therefore, their published papers can be noticed by others more quickly, get more citations, speed up the awakening of SBs, and shorten the sleeping time of SBs.

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.148.2023

4.2 Suggestion

In order to effectively shorten the sleeping time of SBs, we need to cooperate with prestigious and older academic researchers. The newly published paper needs to be consistent with the research theme of SB and cite it, so that the new paper can get more attention, and the frequency of citation increases rapidly, which may become the prince of SB, thus effectively reducing the sleeping time of SBs.

References

[1] Barber, Bernard. Resistance by Scientists to Scientific Discovery[J]. American Journal of Clinical Hypnosis, 1963, 5(4):326-335.

[2] Garfield, E. Premature discovery or delayed recognition-why?[J]. Current Contents, 1980, 4:488-493.

[3] Garfield E. Delayed recognition in scientific discovery: Citation frequency analysis aids the search for case histories[J]. Current Contents, 1989, 23:3-9.

[4] Garfield E. More delayed recognition. Part 1. Examples from the genetics of color blindness, the entropy of short-term memory, phosphoinositides, and polymer rheology[J]. Current Contents, 1989, 38:3-8.

[5] Garfield E. More delayed recognition. Part 2. From inhibin to scanning electron microscopy[J]. Current Contents, 1990, 9:3-9.

[6] Glänzel W, Schlemmer B, Thijs B. Better late than never? On the chance to become highly cited only beyond the standard bibliometric time horizon[J]. Scientometrics, 2003, 58(3):571-586.

[7] Van Raan A F J. Sleeping Beauties in science[J]. Scientometrics, 2004, 59:467-472.

[8] Ohba N, Nakao K. Sleeping beauties in ophthalmology[J]. Scientometrics, 2012, 93(2):253-264.

[9] Zong, Z., Liu, X. , Fang, H. Sleeping beauties with no prince based on the co-citation criterion[J]. Scientometrics, 2018, 117:1841-1852.

[10] Fang H. A transition stage co-citation criterion for identifying the awakeners of sleeping beauty publications[J]. Scientometrics, 2019, 121(1):307-322.

[11] Ke Q, Ferrara E, Radicchi F, et al. Defining and identifying sleeping beauties in science[J]. Proceedings of the National Academy of Sciences of the United States of America, 2015, 112(24):7426-7431.

[12] Rubin D B. Using propensity scores to help design observational studies: Application to the tobacco litigation[J]. Health Services and Outcomes Research Methodology, 2001, 2(3):169-188.

[13] Rosenbaum P R, Rubin D B. The bias due to incomplete matching[J]. Biometrics, 1985, 41(1):103-116.