Patent Citation Inflation: The Phenomena, Measurement Methods, and Proposed Solutions Based on Relative Citation Indicators

Lin, Z.; Ying, H.; Lixin, C.

STI 2018 Leiden

23rd International Conference on Science and Technology Indicators
"Science, Technology and Innovation Indicators in Transition"

STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Chair of the Conference

Paul Wouters

Scientific Editors

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

Layout

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

The articles of this collection can be accessed at https://hdl.handle.net/1887/64521

ISBN: 978-90-9031204-0

23rd International Conference on Science and Technology Indicators (STI 2018)
"Science, Technology and Innovation indicators in transition"
12 - 14 September 2018 | Leiden, The Netherlands
#STI18LDN

Patent Citation Inflation: The Phenomena, Measurement Methods, and Proposed Solutions based on Relative Citation Indicators

Zhang Lin[*], Huang Ying[**] and Chen Lixin[***]

[*] *zhanglin_1117@126.com*
School of Information Management, Wuhan University, Wuhan 430072, (China)

[**] *huangying_work@126.com;*
Department of Public Administration, Hunan University, Changsha 410082, (China)

[***] *lynnchenlixin@163.com*
School of Information Technology, Shangqiu Normal University, Shangqiu 476000, (China)

Introduction

Like scientific articles, patents contain references. Applicants are required to appropriately describe and cite prior knowledge about the technological background of their inventions (Criscuolo & Verspagen, 2008; Jaffe, Trajtenberg & Fogarty, 2000). Furthermore, during a patent's prosecution, the examiners are required to search for any related prior patents and publications to judge the application's patentability, called examiner citations (Cotropia, Lemley & Sampat, 2013). Cotropia (2009) and Sampat (2010) pointed out that providing an abundance of references could be beneficial for an applicant. Recently, the United States Patent and Trademark Office (USPTO) has been granting more and more patents with more and more references. In 1976, each patent contained an average of six references, whereas by 2015 the average number of references had increased more than eight times to an average of 50. This may hint that, the more patents that have been granted, the greater the number of references that could be provided.

Obviously, the continuing increase in the number of patents and a predilection by applicants to provide a greater number of references within their applications, would certainly cause an explosion in patent references and would inevitably result in patent citation inflation. Hall, Jaffe & Trajtenberg (2001) pointed out that: "*The combination of more patents making more citations suggests a kind of citation 'inflation' that may mean that later citations are less significant than earlier ones*". Besides, Persson, Glänzel and Danell (2004) studied inflationary bibliometric values of the basic indicators including citations counts in the view of scientific collaboration, and suggested the need for relative indicators used for evaluation studies in order to guarantee the validity of conclusions drawn from bibliometric results.

Although patent citations have been frequently analyzed in various aspects, the phenomenon of explosion or inflation in patent citations has not been widely noted and concerned. In this article, we study the phenomena of patent citation inflation, present several methods to detect and measure patent citation inflation, and propose some solutions based on relative citation indicators aiming to improve citation-based patent evaluation approaches.

The Measurement Methods

Patent citation inflation, which is caused by an enormous supply of references, can lead to a general rise in citation counts and results in a loss of impact or value. Some basic concepts related to citation indicators and their computational formulas are listed in Appendix Table 2. In all the formulas below, m represents the year under analysis, n represents the year of comparison, and p and q represent the granted year of the patent[1].

Firstly, we consider inflation of patent references. In a period from year n to year m, patent reference inflation can be measured by the rate of change in the number of references (SR) supplied in year m versus year n. This ratio is calculated by:

$$RI_{m,n,q} = \frac{SR_{m,q} - SR_{n,q}}{SR_{n,q}} = \frac{SR_{m,q}}{SR_{n,q}} - 1 \qquad (1)$$

where RI denotes patent reference inflation rate. The higher the rate of change is, the more serious the patent reference inflation occurs.

Serious patent reference inflation inevitably causes patent citation inflation. There are two kinds of patent citation inflation. Patent citation inflation received in a particular period means that significantly more citations can be received by patents in the particular period than in another period. For patents granted in year q, citation inflation within a period from year n to year m (using a diachronous method), can be measured by the rate of change in the number of citations (c) received in year m versus year n, using the following calculation:

$$ci_{m,n,q} = \frac{\frac{c_{m,q}}{1+O_{n-q,m-q}} - c_{n,q}}{c_{n,q}} = \frac{c_{m,q}}{(1+O_{n-q,m-q})c_{n,q}} - 1. \qquad (2)$$

in which ci denotes patent citation inflation rate, and $O_{n-q,m-q}$ denotes obsolescence rate of patents granted in year q, i.e., the change rate of share (or probability) of being cited from age $n-q$ to $m-q$ (see Appendix B). Here we take the effect of obsolescence into consideration, since patents are expected to receive fewer and fewer citations when aging.

We take the patents granted in 1976, and their citations received in 2013 vs. 2014 as an example:

$$ci_{2014,2013,1976} = \frac{c_{2014,1976}}{(1+O_{37,38})c_{2013,1976}} - 1 = \frac{28566}{(1-0.058)\cdot 27962} - 1 = 0.085.$$

The number of citations received in 2014 is 28,566, whereas that in 2013 is 27,962, i.e., an annual growth of 2.2%. After taking the obsolescence effect into account[2], citations inflated 8.5% in 2014 versus 2013. In some cases, patent obsolescence rate (O) may be omitted or be regarded as zero sometimes to simplify the algorithm.

The second kind of patent citation inflation refers to that, significantly more citations received by patents granted in some certain years. Using a diachronous method, as of the end of year m, patent citation inflation of a group of patents granted in year p versus another group of patents granted in year q can be measured by using the following calculation:

$$CI_{m,p,q} = \frac{ACR_{m,p}}{ACR_{m,q}} - 1. \qquad (3)$$

[1] In all the examples below, 2015≥ m≥ n≥ p≥ q≥ 1976, and occasionally q≥ 1790 (The first US patent was granted by President Washington in 1790). The examples supporting each concept are based on the data in Appendix Table 1.

[2] According to the data listed in Appendix A, for the patents granted in 1976, the obsolescence rate from age 37 (in 2013) to 38 (in 2014) should have reached $O_{n-q,m-q} = O_{37,38} = -0.058$ (it is an expected value calculated by empirical data). It means that the citation counts received by the patents should have expectedly decreased by 5.8% in 2014 versus 2013 according to obsolescence theory; but the citation counts factually increased by 2.2% in 2014 due to serious patent citation inflation in the year.

in which, *CI* denotes patent citation inflation rate. For instance, $CI_{2015,1977,1976}$ = $(SC_{2015,1977}/N_{1977})/(SC_{2015,1976}/N_{1976}) - 1 = 0.043$ means that, as of 2015 the later patents issued in 1977 compared to the earlier patents issued in 1976 averagely received 4.3% more citations. However, Equation (3) can be bias toward earlier patents, which had more time than later patents to receive citations. In order to avoid bias toward earlier patents, we normalize the citation window and present another measurement:

$$CI(m-p)_{p,q} = \frac{ACR_{m,p}}{ACR_{m-p+q,q}} - 1 = \frac{ACR_{(m-p)+p,p}}{ACR_{(m-p)+q,q}} - 1. \qquad (4)$$

For instance, $CI(2015-1977)_{1977,1976} = CI38_{1977,1976} = (SC_{2015,1977}/N_{1977})/(SC_{2014,1976}/N_{1976}) - 1 = 0.077$ means that, as of the end of 2015, the average citation rate of the later patents issued in 1977 was 12.01, whereas as of the end of 2014 that of the earlier patents issued in 1976 is 11.15, inflating by 7.7%. For the later patents, their citation inflated 7.7% instead of 4.3%. It indicates that citation inflation rate became greater when bias toward earlier patents was avoided by normalizing the time window.

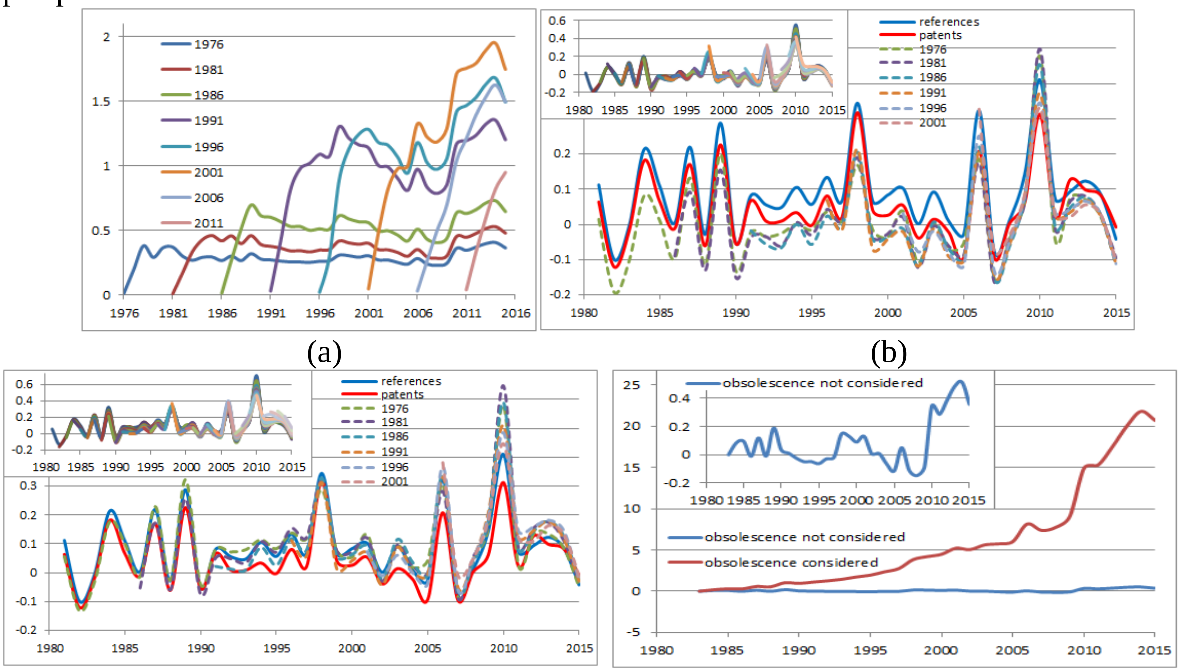
Demonstration Analysis

Data

We downloaded full-text data of USPTO patents granted between 1976 and 2015, including citation information. For simplicity, all references to USPTO utility patents that were granted during 1976–2015 were selected as our dataset (see Appendix Table 1). Ultimately, our dataset contained a total of 5,270,483 utility patents and 59,820,251 citations received by these utility patents during the period.

The analysis of the Patent Citation Inflation during 1976–2015

Patents incline to receive more citations along with the explosion of patent references during 1976-2015. Figure 1-3 present the phenomena of patent citation inflation from different perspectives.

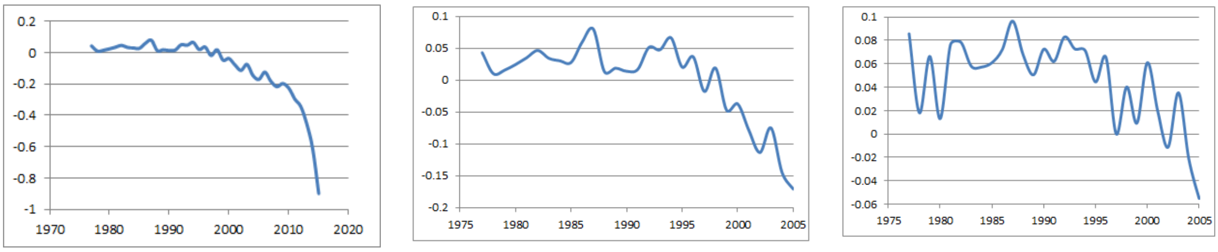


(a) (b)

(c) (d)

Figure 1. The first kind of patent citation inflation (a group of patents in a given period)

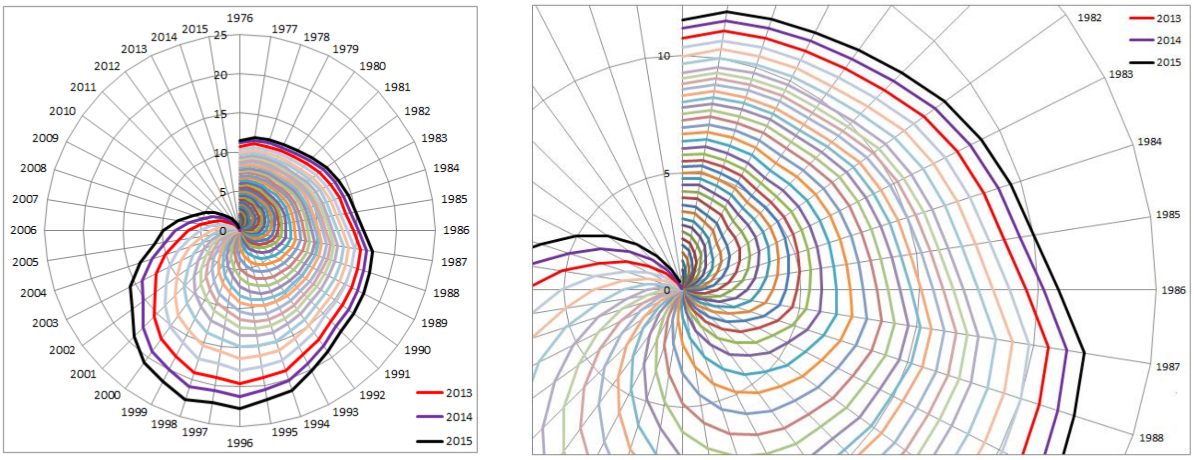
Note: The (***a***) presents the annual average numbers of the citations received by the patents granted in the selected years; (***b***) indicates the annual citation inflation rates of the patents granted in the six years ($ci_{m,m-1,q}$; $q+5 \leq m \leq 2015$; q = 1976, 1981, 1986, 1991, 1996, 2001; "references" denotes the annual patent reference inflation rate, and "patents" denotes the annual growth rate of patents; the inset shows the annual citation inflation rates of the patents granted during 1976−2010); the (***c***) shows the modified annual citation inflation rates in the six years; and (***d***) for the patents granted in 1976, their citation inflation rates during 1983−2015 versus 1983 ($ci_{m,1983,1976}$; 1983≤ m ≤ 2015). In Figure 1, X-axis denotes the year when the patent was cited.



(*a*) (*b*) (*c*)

Figure 2. The second kind of patent citation inflation (a group of patents versus to another)

Note: The distributions of (***a***) as of 2015, the citation inflation rate for two groups of patents granted in two adjacent years ($CI_{2015,q+1,q}$; 1976< q< 2015); (***b***) the highlight of the rate ($CI_{2015,q+1,q}$; 1976< q< 2005); and (***c***) the inflation rate with a 10-year time window ($CI10_{q,q-1}$; 1976< q< 2005). In the figures, X-axis denotes the granted year of the cited patent, i.e., q.



(*a*) (*b*)

Figure 3. The annual growth rings of the average citation rates

Note: The annual growth rings of (***a***) the average citation rates ($ACR_{m,q}$; 1976≤ q≤ m≤ 2015) of the patents granted during 1976−2015. For instance, the curve in 2015 $ACR_{2015,q}$ means as of 2015 the average number of the citations had been received by the patents granted in the year q. $ACR_{2015,1998}$ = 22.7 in 2015, meaning that as of 2015 the patents granted in 1998 had been cited, on average, 22.7 times. The (***b***) shows the zoom-in of the rates of some old patents; the angular coordinate denotes the year of the cited patent was granted, i.e., q.

As a result, patent citations have been cumulated greatly over the past four decades (see Figure 3*a*). Figure 3*a* indicates that for different patents, grouped by granted year, the growths of their average citation rates (annual growth rings of citation) are different – later patents' average citation rates usually increased faster than earlier patents' (later patents received much more citations than earlier ones). In addition, the figure indicates that generally the growths of the average citation rates increased faster in later years than in earlier years – most patents received more citations in later years and even old patents also did well in recent years (see Figure 3*b*).

In conclusion of this section, we measured two different kinds of patent citation inflation rates (*ci* and *CI*): patent citation inflation received in a particular period, and patent citation inflation received by a certain group of patents, both based on the diachronous measures. In terms of the first kind, patent obsolescence should be considered and can be used to modify patent citation inflation. As for the second kind, time window should be considered and can be used to modify patent citation inflation. Patent citation inflation rate with a certain time window, such as *CI*10, is a recommendable method to detect and measure patent citation inflation, especially appropriate for recent patents with short time window.

Solutions to Eliminate Effects of Patent Citation Inflation: Relative Indicators

Patent citation inflation has occurred frequently and seriously in the recent four decades; so it would not be judicious that a patent's impact is measured by the citation count it has received, without considering citation inflation. In these regards, Jaffe and his collaborators (Hall, Jaffe & Trajtenberg, 2001a; Henderson, Jaffe & Trajtenberg, 1998; Jaffe & Trajtenberg, 1998) proposed two methods to eliminate some disadvantageous effects of patent citation inflation. One is the fixed-effects approach, scaling citation counts by dividing them by the average citation count for a group of patents; the other is quasi-structural approach, which attempts to distinguish the multiple effects on citation rates via econometric estimation (Hall, Jaffe & Trajtenberg, 2001b; Jaffe & Lerner, 2001).

Inspired by the above two approaches, we present some new integration approaches, seen in Equations (6)–(9) in Table 1. Compared to the solutions proposed in Hall, Jaffe & Trajtenberg (2001b) and Jaffe & Lerner (2001), our approaches take more parameters and indicators into consideration, such as mean value μ (average citation rate), standard deviation σ and percentile rank *p*, in order to more accurately and reasonably calculate the relative value of patent citation count when patent citation inflation occurs. To better interpret and compare these measures, we take two patents as examples in Table 1: patent *a* granted in 1990 (No. 4,979,787) and patent *b* granted in 1985 (No. 4,529,663), and both of them were cited 65 times as of the end of 2015.

Table 1. Relative Citation Indicators to Eliminate Effects of Patent Citation Inflation

Eq.	Description	Formula	Parameter	Example (patents *a* and *b*)
(5)	The average citation rate (also named the fixed-effects approach [3]): the ratio of the citation count to the mean (the average citation rate).	$RC_{\mu} = C/\mu$	mean: $\mu = ACR$	$RC_{\mu}(a)$ =65/17.96 = 3.62 $RC_{\mu}(b)$ = 65/14.99 = 4.34 $RC_{\mu}(b)/RC_{\mu}(a)$ =1.20
(6)	The annual average citation rate: the average of the ratio of the annual citation count to the annual mean (annual average citation rate).	$RC_{\mu i} = \frac{\sum_n^m (c_i/u_i)}{n-m+1} = \frac{\sum_n^m (c_i/acr_i)}{n-m+1}$	annual mean: $\mu_i = acr_i$	$RC_{\mu i}(a)$ =96.91/26 =3.73 $RC_{\mu i}(b)$ =136.15/31 =4.39 $RC_{\mu i}(b)/RC_{\mu i}(a)$ = 1.18
(7)	The mean and standard deviation approach: the ratio of the deviation of the citation count from the mean to the standard deviation.	$RC_Z = (C - \mu)/\sigma$	mean: μ standard deviation: σ	$RC_Z(a)$ =(65−17.96)/33.24 =1.42 (P_a =92.15%) $RC_Z(b)$ =(65−14.99)/26.64 =1.88 (P_b =96.97%) $RC_Z(b)/RC_Z(a)$ =1.33 (P_b/P_a =1.05)
(8)	The annual mean and standard deviation approach: the ratio of the deviation of the annual citation count from the annual mean to the annual standard	$RC_z = \frac{\sum_n^m (c_i - \mu_i)/\sigma_i}{n-m+1}$	annual mean: μ_i annual standard deviation: σ_i	$RC_z(a)$ =28.59/26 =1.10 $RC_z(b)$ =43.29/31 =1.40 $RC_z(b)/RC_z(a)$ = 1.27

[3] The fixed-effects approach was presented by Hall, Jaffe and Trajtenberg (2001a).

Eq.	Description	Formula	Parameter	**Example** (patents *a* and *b*)
	deviation.			
(9)	The percentile rank approach: the percentage of counts in its frequency distribution that are lower than it.	RC_p $=\mathrm{P}\{X < C\}$	percentile rank: *p*	$RC_p(a)$ =95.62 $RC_p(b)$ =97.27 $RC_p(b)/RC_p(a)$ = 1.02

Note: For a patent granted in year *n*, as of the end of year *m*, it received citation count *C*; only in year *i* ($m \geq i \geq n$), it received citation count c_i. For the patents granted in year *n*, as of the end of year *m*, each patent on average received citation count μ (mean value, which is equal to average citation rate *ACR*), and standard deviation is σ; only in year *i*, each patent on average received citation count acr_i (annual average citation rate which is equal to annual mean value μ_i), and standard deviation is σ_i. As of the end of year *m*, for all citation counts received by the patents granted in year *n*, percentile rank *p* of a citation count is given.

In terms of the second kind of patent citation inflation, all the above five approaches can be used in patent citation evaluation, just as the example listed in Table 1. But in condition of the first kind of patent citation inflation, only the Equation (6) and (8), can be used, because they consider annual patent citation count. For instance, patent *c* (No. 4,950,222) and patent *d* (No. 4,959,774), which were both granted in 1990 and were both cited 171 times as of the end of 2015. Obviously, measured by Equations (5), (7) and (9), the two patents are equal in citation impact, with the same relative citation count. Whereas when they are measured by Equations (6) and (8), the result suggests that patent *d* had greater impact or higher quality than patent *c*[4]. The reason is that patent *c* received most citations (109 citations) during the period 2009−2011 when patent citation inflation occurred seriously (see Figures 1*b*, 1*c* and Figure 4).

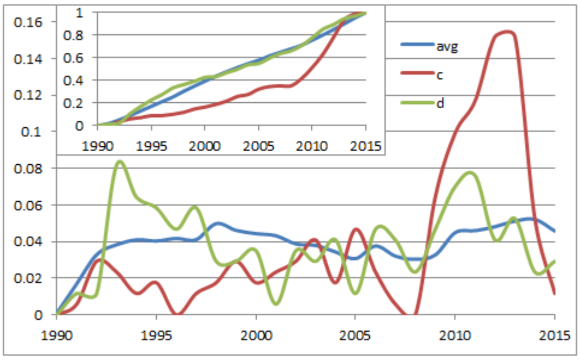


Figure 4. The distributions of share (probability densities) of citation counts received by patents *c* and *d*, and that of average citation count received by the patents granted in 1990 (the inset shows their cumulative distributions)

The above example shows that even though two patents received the same citation count, it does necessarily not mean that they have the same citation impact. That is because patent citation inflation can cause devaluation in citation count, a decline in the overall value or impact of citations. In addition, patent citation inflation had occurred frequently and seriously, meaning that relative indicators (such as five approaches in Table 1) instead of absolute indicators (such as citation count) should be used in patent evaluation[5]. Thereinto, the

[4] The relative values of citation counts received by patents *c* and *d* are respectively: $RC_{\mu i}(c)$ =224.22/26 =8.62, $RC_{\mu i}(d)$ =238.32/26 =9.17, $RC_{\mu i}(d)/RC_{\mu i}(c)$ = 1.06, calculated by Equation (6); $RC_{\sigma i}(c)$ =67.57/26 =2.60, $RC_{\sigma i}(d)$ =87.07/26 =3.35, $RC_{\sigma i}(d)/RC_{\sigma i}(c)$ = 1.29, calculated by Equation (8).

[5] We note that using relative indicators or absolute indicators is mainly determined by whether patent citation inflation occurred and its extent, which can be detected and measured by the methods as presented above.

Equation 8 is more recommendable than the others, since it involves two informative parameters and it can cope with both two kinds of patent citation inflation.

A Case Study of Patent Evaluations for Eight different Universities

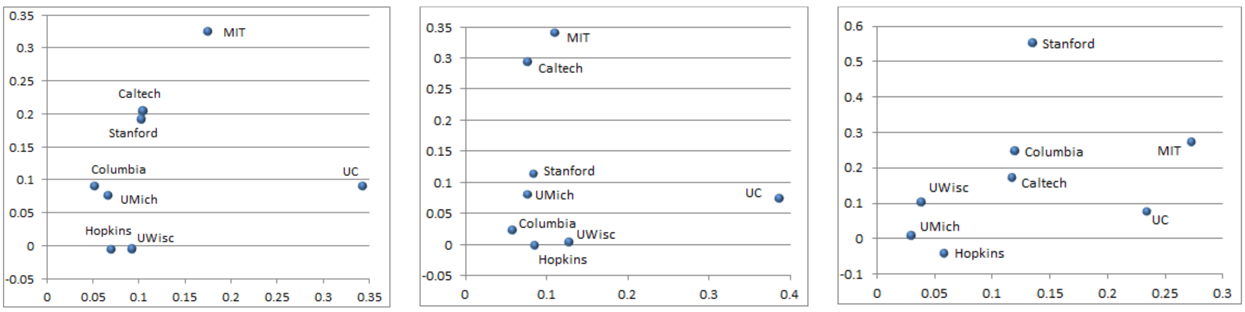
In order to examine and verify these relative indicators proposed in Table 1, we present a case study of patent evaluations for the eight American universities: UC (University of California), MIT (Massachusetts Institute of Technology), Stanford (Stanford University), Caltech (California Institute of Technology), UWisc (University of Wisconsin), Hopkins (Johns Hopkins University), UMich (University of Michigan) and Columbia (Columbia University in the City of New York). The relative indicators of the citation counts received by patents in the eight universities are presented in Table 2.

Table 2. Relative indicators of the citation counts received by patents of eight universities

Univ.	All fields				Chemistry				Communication			
	No.	**ACR**	**μ App**	**μ-σ App**	**No.**	**ACR**	**μ App**	**μ-σ App**	**No.**	**ACR**	**μ App**	**μ-σ App**
UC	9440	16.29	1.259	0.092	3495	10.95	1.272	0.076	272	19.52	1.017	0.078
MIT	4809	25.70	2.003	0.326	993	17.47	2.347	0.342	317	27.36	1.773	0.274
Caltech	2857	17.90	1.622	0.206	692	16.01	2.032	0.295	136	28.92	1.341	0.174
Stanford	2822	18.51	1.532	0.193	757	9.76	1.360	0.115	157	51.23	2.625	0.555
UWisc	2535	10.96	0.909	−0.003	1150	8.05	0.972	0.005	44	23.52	1.172	0.104
Hopkins	1907	15.15	0.971	−0.005	767	9.15	1.015	−0.0004	67	10.48	0.820	−0.040
UMich	1826	13.81	1.202	0.078	690	10.27	1.290	0.082	34	13.18	0.711	0.011
Columbia	1411	15.85	1.276	0.092	517	9.65	1.115	0.025	139	15.38	1.813	0.249

Note: The field of chemistry refers to the International Patent Classification (IPC) C01−14, C40 or C99; and the field of communication is IPC H04 referring to "electric communication technique". In the table, "No." is the number of patents; "ACR" is the average citation rate, i.e., the average number of citations received by per patent; "μ App" represents that the mean of values calculated by Equation (5) listed in Table 1; "μ-σ App" represents that the mean of values calculated by Equation (8) listed in Table 1. During 1976−2015, UC had obtained 9440 patents and those patents had received 153,813 citations as of the end of 2015; MIT had obtained 4809 patents with 123,602 citations; Stanford had obtained 2822 patents with 52,228 citations; Caltech had obtained 2857 patents with 51,145 citations; Wisconsin had obtained 2535 patents with 27,785 citations; Hopkins had obtained 1907 patents with 28,894 citations; UMich had obtained 1826 patents with 25,310 citations; Columbia had obtained 1411 patents with 22,389 citations.

Figure 5 shows that UC obtained the most patents, but MIT, Caltech and Stanford's patents are of higher relative citation impact – on average their patents received much more citations than the others did. In the field of chemistry, MIT and Caltech's patents are of higher relative citation impact; whereas in the field of communication, MIT obtained the most patents, and Stanford's patents are of higher relative citation impact – the patents averagely received much more citations than the others did.



(*a*) (*b*) (*c*)

Figure 5. Patent evaluations in quantity and citation impact for the eight universities

Note: The scatter diagrams of the eight universities: (***a***) in all fields, (***b***) chemistry and (***c***) communication; X-axis represents quantity − the share of patents of each university; and Y-axis indicates citation impact − the relative value calculated by Equation (8) listed in Table 1.

Figure 6 further shows that MIT, UC, Caltech and Stanford are the top four in the ranking based on quantity-impact integration approach ("*q-μ-σ* App") [6]: they not only obtained the most patents, but also their patents received higher citation impact ("ACR") as well as higher relative citation impact ("*μ-σ* App"). However, the ranking positions of the other four universities are interesting. Hopkins obtained more patents and had higher citation impact ("ACR") than UMich, but Hopkins is the last (the eighth) and UMich is the fifth when measured by "*q-μ-σ* App". The ranking based on quantity-impact integration approach ("*q-μ-σ* App"), which considers the patent citation inflation and patent age, might be more reasonable and informative.

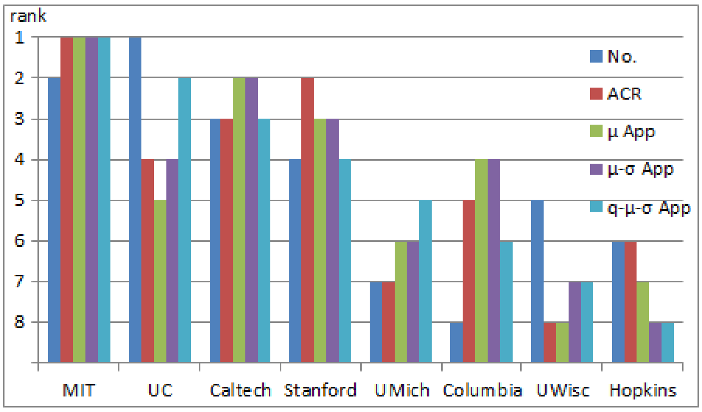


Figure 6. The ranks of eight universities according to different patent indicators[7]

Conclusions

Over the last forty years, the USPTO annually has granted an increasing number of patents with a great large number of references to other patents. Such an increase in the number of references has led to serious patent citation inflation in the period. In this study, we explore the phenomena of patent citation inflation, present some measures used to calculate citation inflation rates, and propose some solutions using relative citation indicators, aiming to improve patent evaluation approaches based on citation impact analysis. Taken as a whole, the results of this study verify that patent citation inflation has occurred seriously and frequently over the past four decades, and confirm the ability of some proposed indicators to detect and measure patent citation inflation.

Acknowledgments

We thank Professor Liming Liang (Henan Normal University) for valuable advices and discussions. This study is supported by the National Natural Science Foundation of China (Grant No. 71573085 and No. 71373252).

[7] Note: "No." means the number of patents obtained by a university during 1976−2015 (patent quantity); "ACR", "*μ* App", "*μ-σ* App" represent patent citation impact (seen in Table 2); and "*q-μ-σ* App" represents patent quantity-impact integration − its value is calculated by the values of "No." and "μ-σ App", i.e., the product of the share of patents of a university (patent quantity) and the mean of values calculated by "μ-σ App" (relative citation impact, i.e., Equation (8); seen in Tables 1−2).

References

Brookes, B.C. (1970). The growth, utility, and obsolescence of scientific periodical literature. *Journal of Documentation*, 26(4), 283–294.

Burton, R.E., & Kebler, R.W. (1960). The “half-life” of some scientific and technical literatures. *American Documentation*, 11(1), 18-22.

Cotropia, C.A. (2009). Modernizing the inequitable conduct doctrine in patent law. *Berkeley Technology Law Journal*, 24(2), 723–783.

Cotropia, C.A., Lemley, M.A., & Sampat. B.N. (2013). Do applicant patent citations matter? *Research Policy*, 42(4), 844–854.

Criscuolo, P., & Verspagen, B. (2008). Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. *Research Policy*, 37(9), 1892–1908. Hall, B., Jaffe A., & Trajtenberg, M. (2001a). The NBER patent citation data file: Lessons, insights and methodological tools. *NBER Working Paper (No. 8498)*, National Bureau of Economic Research, Inc, Oct. 2001.

Hall, B., Jaffe, A., & Trajtenberg, M. (2001b). Market value and patent citations: a first look. *Rand Journal of Economics*, 36(1), 16–38.

Henderson, R., Jaffe A., & Trajtenberg, M. (1998). Universities as a source of commercial technology: A detailed analysis of university patenting, 1965-1988. *Review of Economics and Statistics*, 80(1): 119–127.

Jaffe, A., & Lerner J. (2001). Reinventing public R&D: patent policy and the commercialization of national laboratory technologies. *Rand Journal of Economics*, 32(1), 167–198.

Jaffe, A., & Trajtenberg, M (1998). International knowledge flows: evidence from patent citations. *Economics of Innovation & New Technology*, 8(1-2), 105–136.

Persson, O., Glänzel, W. & Danell, R. (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for rela- tive indicators in evaluative studies. *Scientometrics*, 60(3), 421–432.

Price, D.J. de S. (1965). *Networks of scientific papers*. Science, 149(3683), 510–515.

Sampat, B. (2010). When do applicants search for prior art? *Journal of Law and Economics*, 53(2), 399–416.

Appendix *A*: Appendix Table 1−2

Appendix Table 1. The citation counts received by USPTO utility patents during 1976−2015, grouped by granted year

granted year	*Citation year* 1976	*1977*	*1978*	···	*2013*	*2014*	*2015*	*subtotal*
1976	869	13,705	26,756	···	27,962	28,566	25,596	808,420
1977		734	13,661	···	27,442	28,459	25,497	783,467
1978			548	···	27,990	29,102	25,751	802,079
···				···	···	···	···	···
2013					12,322	94,752	166,991	274,065
2014						12,888	101,297	114,185
2015							11,449	11,449
sum	869	14,439	40,965	···	5,184,186	5,638,617	5,406,315	59,820,251

Note：The cell “869” means that the patents granted in 1976 received 869 citations in the same year in total. The cell “13,661” (seen in row 2 and column 3) means that the patents granted in 1977 received 13,661 citations in the year of 1978 in all.

Appendix Table 2. Concepts and indicators associated with patent citation

Concept and indicator description	**Formula**	**Example**
The number of references given to patents granted in year *q* by patents granted in year *p*	$R_{p,q}$	$R_{2015,1976}$ = 25,596 The patents granted in 1976 were

	Concept and indicator description	Formula	Example
			referenced 25,596 times in 2015.
	The number of citations received by patents granted in year *q* for patents granted in year *p;* of course, $R_{p,q} = C_{p,q}$	$C_{p,q}$	$C_{2015,1976}$ = 25,596 The patents granted in 1976 received 25,596 citations in 2015.
	The number of patents granted in year *q*	N_q	N_{1976} = 701,94 USPTO granted 701,94 utility patents in 1976.
(A1)	The annual average reference rate: the average number of references given to patents granted in year *q* by patents granted in year *p*	$arr_{p,q}$ $= R_{p,q}/N_p$	$arr_{2015,1976}$ = 25,596/299,382 = 0.09 On average, for every 100 patents granted in 2015 contain 9 references to the patents granted in 1976.
(A2)	The annual average citation rate: the average number of citations received by patents granted in year *q* for patents granted in year *p*	$acr_{p,q}$ $= C_{p,q}/N_q$,	$acr_{2015,1976}$ = 25,596/ 701,94 = 0.36 On average, for every 100 patents granted in 1976, they were cited 36 times in 2015.
(A3)	The sum of references given to patents issued since year *q* by patents granted in year *p*	$SR_{p,q}$ $= \sum_q^p R_{p,i}$	$SR_{2015,1976}$ $= \sum_{1976}^{2015} R_{2015,i}$ = 5,406,315
(A4)	The sum of citations received by patents granted in year *q* as of the end of year *m*	$SC_{m,q}$ $= \sum_q^m C_{i,q}$	$SC_{2015,1976}$ $= \sum_{1976}^{2015} C_{i,1976}$ = 808,420
(A5)	The average reference rate: the average number of references given to patents granted since year *q* by patents granted in year *p*	$ARR_{p,q}$ $= SR_{p,q}/N_p$	$ARR_{2015,1976}$ $= SR_{2015,1976}/N_{2015}$ = 5,406,315/299,382 = 18.06
(A6)	The average citation rate: the average number of citations received by patents granted in year *q* as of the end of year *m*	$ACR_{m,q}$ $= SC_{m,q}/N_q$	$ACR_{2015,1976}$ $= SC_{2015,1976}/N_{1976}$ = 808,420/701,94 = 11.52
(A7)	The total number of patents granted between year *q* and year *p*	$SN_{p,q}$ $= \sum_q^p N_i$	$SN_{2015,1976}$ $= \sum_{1976}^{2015} N_i$ =5,250,263

Note: The examples supporting each concept are based on the data in Appendix Table 1.

Appendix *B*: Obsolescence Rate for Patent Citation and its Expect Value

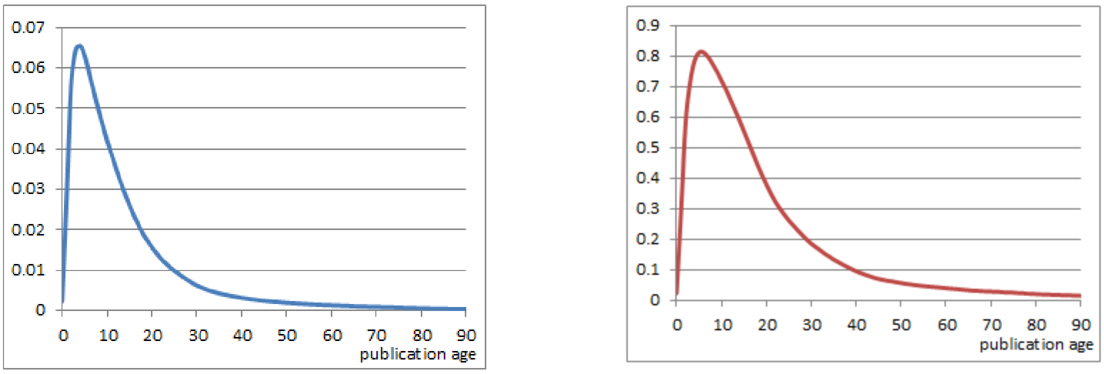
According to the theory of obsolescence in scientific literature (Burton, & Kebler, 1960; Gosnell, 1944; Marton, 1985; Price, 1965), patents should receive more and more citations in the beginning, several years later reach a peak and then gradually get fewer and fewer citations. It indicates that newer patents granted in recent several years should have higher probability of being cited than old ones. Hall, Jaffe & Trajtenberg (2001*a*) point out that for patents more than 10 years older, the older they are, the fewer citations they receive. Hence, patent obsolescence rate can be measured by the rate of change in share (or probability) of being cited for patents from publication age *a* to *b*, using the following calculation:

$$O_{a,b} = \frac{s_b - s_a}{s_a} = \frac{s_b}{s_a} - 1. \quad \text{(A8)}$$

in which, $O_{a,b}$ is obsolescence rate of patents from publication age a to b, s_a is share (or probability) of being cited for patents at the publication age of a. In a given year m, for patents granted in year n, the share (or probability) of being cited at the age of $m-n$ is the share of references given to them:

$$s_{m-n} = \frac{R_{m,n}}{SR_m} = \frac{R_{m,n}}{\sum_{1790}^{m} R_{m,i}}. \quad \text{(A9)}$$

in which, $R_{m,n}$ is reference count given to patents granted in year n and SR_m is the sum of references given to patents issued since 1790 by patents granted in year m. The share can be calculated by a synchronous method.



(*a*) (*b*)

Appendix Figure 1. The distributions of (*a*) the average probability density (or the average share) of being cited for patents at different publication ages; and (*b*) probability of being cited, i.e., the average number of citations received by patents at different ages (the annual average citation rate, *acr*)

Using the data of references provided during 1976–2015 given to USPTO utility patents granted since 1790, we calculated the average probability density (or the average share) of being cited for patents at the different publication ages, using Equation (A9) by a synchronous measure (see Appendix Figure 1*a*). Thus, we can get patent obsolescence rate in a given period using the empirical data shown in Appendix Figure 1*a*. For instance, for patents granted in 1976, their publication age was 37 in 2013 and 38 in 2014; from age 37 to 38, its expected value of patent obsolescence rate is $O_{37,38} = s_{38}/s_{37} - 1 = 0.036/0.038 - 1 = -0.058$, in which, the average share of being cited for patents at the age of 37 is equal to 0.038, and that at the age of 38 falls to 0.036. In other words, the patents should have received 5.8% less citations when ageing from 37 to 38. Alternatively using Equation (A2) by a diachronous measure, we can calculate probability of being cited, i.e., the average number of citations received by patents at different publication ages (the annual average citation rate, *acr*), shown as Appendix Figure 1*b*.