

**Author : Chung-Huei Kuan, Mu-Hsuan Huang, Dar-Zen Chen**

**Title : The Overlooked Citations: Investigating the Impact of Ignoring Citations to Published Patent Applications**

**Journal : Journal of Informetrics**

**Year of Publication : 2020**

### **The citations that should not be overlooked**

Patent citations are generally considered as more objective than paper citations as they do not directly contribute to whether a patent is granted. Other than that, patent citations are also distinct from paper citations in that a patent involves two documents, the early published patent application and the issued patents, and both would continue to receive citations. Even though a paper may be on-lined first and then officially published, the time gap in between is usually limited, and the on-lined version receives few citations. In contrast, there may be a lengthy period between a published application and its corresponding patent and the published application, as it is publicly available for more time, would gather significantly more citations. Even after the application is granted and a patent issued, the published application would still be cited independently, parallel to its issued version.

This separate set of citations, directed to a document other than the issued patent, would impact a true evaluation to the patent. These citations, however, are overlooked in most, if not all, patent analyses. This study is, therefore, aimed to investigate how patents and their published applications accumulate citations, and the degrees of impact if citations to the published applications are not considered.

This study collects about 280,000 pairs of patents and their published applications. These pairs are divided into five groups of roughly identical sizes based on the time difference between their early publication dates and issue dates. The average citations to these five groups' published applications and issued patents are then summarized in the table below.

	All	≤1yr.	, 2 yrs.	, 3 yrs	, 4 yrs.	>4yrs.
<b>Patents</b>	276,940	54,865	84,044	65,209	33,545	39,277
<b>Pat. citations</b>	2.34	2.58	2.25	2.24	2.22	2.47

	All	≤1yr.	, 2 yrs.	, 3 yrs	, 4 yrs.	>4yrs.
<b>Pub. app. citations</b>	6.60	2.90	4.36	6.17	8.66	15.54
<b>Citation to both</b>	0.49	0.51	0.46	0.46	0.51	0.56
<b>Combined citations</b>	8.45	4.97	6.14	7.95	10.37	17.46
<b>VCR</b>	44.76%	63.29%	49.73%	40.76%	33.65%	24.35%

From the table above, one can see that the citations to the published applications are 2, 3, or even more times than those to the issued patents if their time gap is 2, 3, or more years apart. There is little intersection between the two sets of citations, indicating that citations to published applications and issued patents are separate and independent.

This study further calculates an indicator measuring how much value is captured by considering only the citations to the issued patents, called *Value Capture Rate* (VCR), as follows.

$$VCR = \frac{|Citations\ to\ published\ application|}{|Union\ of\ the\ citations\ to\ published\ applications\ and\ patents|}$$

One can also see that, based on the VCR data above, citations to the issued patents only cover about 45% of all citations. In other words, about 55% of a patent's impact or value is ignored. This study further applies the Main Path Analysis, a type of patent citation analysis, to a real-life case and finds that the results, with and without ignoring citations to the published applications, are almost entirely different, again suggesting the importance of the citations to the published applications.

The contribution of this work lies in the demonstration of the importance of citations to the published applications using empirical data, and of the impact of ignoring them. According the findings of this work, the conduction of patent citation analyses is only reasonable when both citations to the patents and to the published applications are considered together, so that a comprehensive reality may be discovered.