11e Séminaire SciScI - OST
Mardi 8 avril 2014
Bronwyn H. Hall (University of California, Berkeley), USA)

"Using Patent Data as Indicators"

RESUME

Prof Hall opened her presentation by noting that she would be providing a social science rather than an economic perspective on the use of patent data as indicators. One of the advantages of using patent data is that this information is, by its very definition, public in nature. The data is easily accessible, usually available for free, and seemingly objective.

I. Patent Data and Knowledge Measurement

But what exactly is it that patents indicate? First, patents indicate a property right in a piece of knowledge and can therefore be used to measure innovative output. Second, at least in the US and Europe, patents cite the older pieces of knowledge on which they are based. They therefore provide a map – albeit an imperfect one – of the links between different outputs or knowledge that can be seen to spill over from one person to another.

However, if patents are used as indicators, it is necessary to understand exactly what they mean. How and why are they taken out by firms? How are they administered, and enforced, in different jurisdictions? How does all this change over time? It is therefore unwise to assume that patents are a stable measure of innovative output denominated in the same units over time. Keith Pavitt, for example, has shown that there are 3 sources of bias in patent counts: differences across countries of the economic costs and benefits of patents; differences among technologies and sectors as to the importance of patents as a protection; and differences among firms in their propensity to patent, especially when it comes to unimportant innovations.

For example, the volume of US patent almost halved during 1980s. However, this does not mean that the US economy was less inventive during that time. Rather it reflects a number of variables: a change in the law to bring the US into line with the TRIPS Convention; a sharp increase in patents by the telecoms industry in the subsequent period; the fact that developing new drugs became more expensive over time. If these variables are not well understood, the figures for patent applications alone could lead to misconceptions.

When considering patents as indicators it is also necessary to define the unit of analysis, and understand the distinction between a patent from a single jurisdiction and a patent family. A patent family is a collection of documents from different patent offices that covers the same invention. As the precise definition and scope of a patent varies across regions, there are multiple definitions of patent families based on the concept of the priority patent.

A priority patent is the patent application that establishes the date before which the examiner searches for prior art. For most inventors in developed countries it refers to the application in the home country patent office.

- When analysing applications, grants, opposition or litigation behaviour, the appropriate unit of measurement is the individual patent.
- When analysing invention, the appropriate unit of observation is the patent family or, in fact, the priority patent because that is the actual invention.

II. Patent Data and Knowledge Value

Jacob Schmookler's 1960s work was the first to analyse economic behaviour on the basis of patent statistics. However, Zvi Griliches' 1980 project was the first to use large sample, computerised data from the US Patent Office. That showed that patents were strongly correlated to R&D spending across firms, with an elasticity close to one. However, patents did not help explain sales, profits, or market value of firms due a significant skew in the distribution of patent value: most patents were worthless but a few were worth a great deal. Patent citations were also skewed. Most patents received zero, one or two citations; a few receive hundreds.

A 2005 paper by Hall, Jaffe, and Trajtenberg explored the correlation between patent citations and the value of a firm. If a firm's inventions are highly cited, did that make the firm more valuable given its R&D spend? They found that cites per patent were more important than patent yield itself. Below the median, cites per patent had no effect, and it was only when firms averaged above the median that they saw a corresponding increase in value. That was the case for both small firms and large firms. Of interest was the fact that self-cites were worth twice as much as third party cites indicating that firms that practised self-citation owned their technology trajectory and generated higher profits from it.

In terms of timing, citations received before value was measured mattered less than those received after, although they were useful for forecasting future cites. Surprisingly, predictable and unpredictable citations carried approximately equal weight.

Since that study was carried out, other correlates of patent value have emerged: opposition and litigation, family size, backward citations, claims, and type of citation. A new indicator is a network measure that looks at business strategy complexity and technology complexity. This shows a high level of networking in the telecom sector, for example, indicating that the technologies overlap enormously resulting in mutually blocking relations. Hence, the high level of litigation in this field.

III. Patent Data and Knowledge Flows

Citations can also be used as indicators of knowledge flow. A 2002 paper by Jaffe, Trajtenberg and Fogarty surveyed 1,300 inventors and found that about 50% of patents corresponded to some kind of knowledge flow. A further 25% corresponded to very substantial flows.

Citations have also been used to study geographic localisation. Does knowledge move more easily in localised areas than across countries or across districts? Research has been carried out on the spillover effect from alliances. If Firm A forms a technology alliance with Firm B, does it learn more from that firm than if it had not formed that alliance? Finally, work has been done on citations as measures of knowledge types: generality, originality, general purpose technologies (GPTs), etc. Jones

and Uzzi suggest that a radical scientific paper is one that combines citations that are rarely seen together; Fleming and coauthors are beginning to apply this idea to patents.

IV. Conclusion

Prof Hall concluded by stating, first, that patents were useful as indicators, especially when citation-weighted. However, it was necessary to understand the political, legal and business policy environments that surrounded those patents. Second, citations were a defensible partial measure of knowledge transfer. They suggested a certain level of spillover localisation. Finally, the development of richer citation measures was a work in progress. Going forward, Prof Hall considered that there were 2 major areas that merited further investigation: achieving greater consistency in assignee names, and improving our ability to classify industry in addition to technology classes.

Prof Hall closed with a list of various sources of patent data from around the world, including NBER, Patstat, the OECD, EPO, the IIP (Japan), and SIPO (China).

V. Q&A Session

Frédérique Sachwald asked about citation to **non-patent literature**. Prof Hall noted that patents can indeed cite non-patent literature such as serious journal articles or the trade press. This was an under-exploited issue today but was clearly worth exploring.

January Sadlak asked if there was a positive correlation between the number of patents and the number of **published scientific papers**. Prof Hall noted that, generally speaking, productive researchers published more, as did productive institutions and productive countries. At the same time, some countries published a great amount relative to how much they patented, and vice versa. Jacques Mairesse added that work was underway on this topic at Leuven University, under the leadership of Dirk Czarnitzki.