ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Bachelor Thesis International Bachelor Economics and Business Economics

Patent and Performance of Global Technology Firm

Author: Jyushin Park, 509800

Supervisor: Dr. Vadym Volosovych

Second Assessor:

Date:

ABSTRACT

One of the fundamental questions of economics is the role of knowledge capital in determining the productivity of a nation or a firm. One way to compensate the traditional R&D stock is to use the patent that a firm has applied for or registered as a proxy variable of knowledge capital. By analyzing global tech firms that has patent activity during 1980 and 2019, this research showed that patent stock can explain part of real sales that is not explained by tangible assets or labour inputs, with statistical significance. Patent stock that considers patents' citation information better explained a firm's sales productivity and showed higher correlation with sales productivity than simply calculated patent stocks.

Table of content

1. INTRODUCTION

One of the fundamental questions of economics is the role of knowledge capital in determining the productivity of a nation or a firm. A traditional way of measuring knowledge capital, under the assumption that research and development cost is accumulated as a firm's knowledge regardless of success, is to proxy knowledge capital with the R&D stock, which is a variable that accumulated the R&D cost with a constant rate of depreciation. However, this method has limitations in that it cannot reflect the fact that the R&D may not always result in innovation.

One way to improve the problem R&D stock has as a proxy variable of knowledge capital is to use the patent that a firm has applied for or registered. The patent does not reflect all the innovative activities that a firm has done. However, works of Pakes and Griliches (1980) and Griliches (1981) and many follow up studies have shown that patents can be used as a proxy variable in measuring a firm's knowledge capital. With the recent trend of matching extensive patent data to the firm's financial data, such findings are contributing to the facilitation patent data are bringing to research on areas such as economics and business administration.

The aim of this research is to see whether it is possible to measure the knowledge capital of global tech companies through their patent data. Using the method of Bloom and Van Reenen (2002), I define the knowledge capital as intangible assets that create revenue and study whether the revenue that is not explained by labour and capital inputs can be explained by patent stocks with statistical significance. Specifically, not only the number of registered patents but the number of citations were also considered in the analyzation, so that the patent stock that reflects the quality of patents are taken into account. The results showed that the patent variable could partly explain the revenue that are not explained by tangible assets or labour inputs, with statistical significance.

Many research were done on ways to measure a firm's knowledge capital by using patent data. Pakes and Griliches (1980) and Griliches (1981) each analyzed patents that were applied on USPTO from 1968 to 1975, and 121 US firms that had R&D expenditures above a certain level during 1963 to 1975. They showed that the market value of a firm measured by Tobin's q was significantly correlated to R&D stock and the number of registered patents. Cockburn and Griliches (1988) and Griliches (1990) established the notion of patent stock, which is defined by the capitalization of patent stock and proposed a production function that reflects patent stock and showed that patent stock and a firm's Tobin's q are positively correlated using the sample of Pakes and Griliches (1980).

Apart from the effort to measure knowledge capital using just the number of registered patents or simply calculated patent stocks, there also has been a continuous effort in trying to account for the quality of patents. Schankerman and Pakes (1986) took the quality of patents into consideration by using the patent renewal fee the patentee has paid. Trajtenberg (1990) used citation information of patents to measure the importance of each patent. Lanjouw and Schankerman (2004) also proposed an index for each patent quality by comprehensively taking the number of patent claims, the scale of patent family, number of citations, and impact factor into account.

Research on measuring knowledge capital by using patent information has accelerated after Hall, Jaffe and Trajtenberg (2001) established the NBER Patent Data Project (PDP) database. NBER PDP database is a comprehensive database that matched every firm's financial information reported on Compustat and patents registered on USPTO. It provides a wider range of firms and provides more comprehensive patent information than databases used for previous research, which enables wider, in-depth research than before.

Using this (NBER PDP) database, Hall, Jaffe, and Trajtenberg (2005) showed that apart from R&D stock and patent stock, citation stock also has a significant relationship with a firm's market value (Tobin's q). Such result is significant in that it proved, through a large sample, that one can measure a firm's knowledge capital more accurately by using qualitative patent information (citation count) together with quantitative information of patents (number of applied patents).

Besides the research mentioned above, Kogan, Papanikolaou, Seru, and Stoffman (2017) suggested a way to measure the economic value of each patents and showed that patent values that were measured by this method were consistent with Schumpeterian growth model's main prediction. Although they did not use the NBER PDP database, the research of Bloom and Van Reenen (2002) shows that citation stock, which takes the number of citations per patent into account, explains revenue productivity better than simply calculated patent stock.

Firms' patent activity based on one, particular country and a country's patent office have been continuously researched upon, but not much research have been done on global tech firms, using combined data base of different patent office. This research is meaningful in that unlike previous research, it combines international patent information constructed from AcclaimIP and patent information registered on WIPO and USPTO to analyse global tech firms.

This paper is organized as follows: Section 2 discusses model that is used in regression analysis. Then, Section 3 gives a detailed explanation of the data used in the analysis and methods to establish various patent stock. Section 4 discusses main results of the regression analysis. Lastly, Section 5 presents conclusions.

2. MODEL

To study the effect of patent activities on a firm's productivity, Bloom and Van Reenen (2002) assumes a Cobb-Douglas production function as follows:

(2.1) $$logQ_{it} = logA_{it} + \alpha logG + \beta logN_{it} + \gamma logK_{it}$$

where Q represents output, A represents overall efficiency parameter, G represents the stock of accumulated knowledge capital, N represents labour input and K represents capital input. In other words, it uses a typical Cobb-Douglas production function but separates the total factor productivity part into overall efficiency parameter (A) and knowledge capital (G^a). The purpose of this separation is to explicitly investigate the role the knowledge capital plays in the total factor productivity. The subscript i represents a firm and t represents period.

The linear regression model that Bloom and Van Reenen (2002) set up to estimate above production function is as follows:

(2.2) $$logQ_{it} = \alpha logPAT_{it} + \beta logN_{it} + \gamma logK_{it} + \eta_i + \tau_i + \nu_{it}$$

Efficiency parameter A is composed of firm specific fixed effect (η_i), time effect(τ_i), and error term (ν_{it}) and is expressed as $A_{it} = \exp\,(\eta_i + \tau_i + \nu_{it})$. To quantify knowledge capital (G), which is an intangible asset that one cannot directly observe, firms' patent activities (PAT) is used as proxy variable for the empirical analysis.

When using patent information as a proxy for the knowledge capital a firm has accumulated, it is better to use cumulated stock rather than flow (the number of applied patents). This is because patents that were applied in the past can be technologically valid until now and have influence on the firm's present productivity.

On the other hand, it is also not realistic to say that technologies which were developed a long time ago has equal effect on the productivity as a recently developed technology. Considering these facts, this research follows the methods of Bloom and Van Reenen (2002), Hall, Jaffe, and Trajtenberg (2005) and use perpetual inventory method to construct the variable patent stock as a proxy for knowledge capital. If we note patent stock in period t as X_t, the number of patents applied on period t as X_t, and depreciation rate as $\delta \in (0,1)$, then patent stock X_t can be constructed recursively as follows:

(2.3) $$X_t = (1 - \delta)X_{t-1} + X_t$$

In this research, depreciation δ is set as 30%, as that of Cockburn and Griliches (1988) and Bloom and Van Reenen (2002). Year of patent application, and not registration, is used for patent stock because of the time lag between the application and registration. Generally, it takes two to three years for a technology to be filed as a patent, be evaluated, and be finally published. However, because the development of the technology is already finished at the time of filing, the technology can influence a firm's productivity even though it is not finally published as a patent. Therefore, it would make sense to use the year of filing to construct patent stock.

On the other hand, various patents inherently have heterogeneous technological value. In this sense, the weakness of patent data would be that a firm's fundamental core technology and subsidiary inventions are all counted as one patent. Therefore, to improve this weakness and construct more accurate proxy variable, one needs to measure the quality of patents and calculate patent stock that reflects the quality.

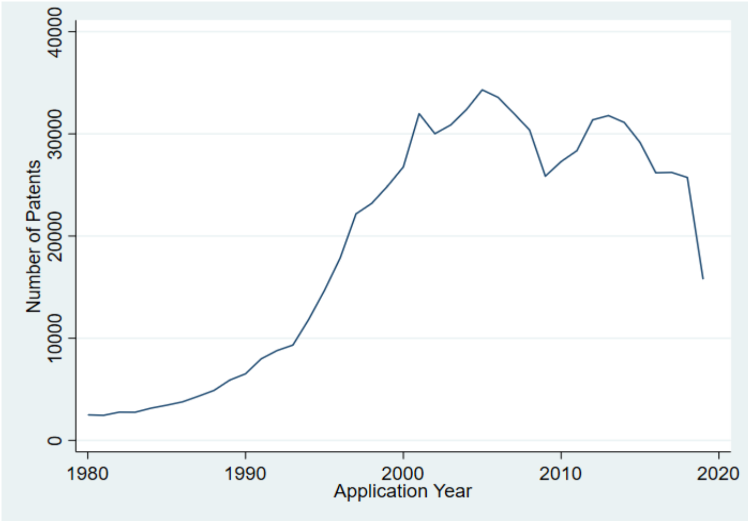
Although there are many ways for measuring the quality of patents, I am going to use number of patent citation , which Trajtenberg (1990), Bloom and Van Reenen (2002), Hall, Jaffe and Trajtenberg (2005 used in their research. However, the number of patent citation has the risk of downward bias due to data disconnection issue, and ways to make up for this problem will be discussed in Section 3.

3. DATA

3.1 WIPO-USTO panel data

The research subject firms selected are reported as top 100 global tech companies Thomson Reuters' (2018) report on 2018. They restricted set of tech companies to those that have at least $1 billion in annual revenue. They implemented the eight pillars of performance and related parameters to construct statistical model employing Bayesian inference. The content of the eight pillars are as follows: financial performance, management & investor confidence, innovation, legal compliance, environmental impact, people & social responsibility, reputation, and risk & resilience. The 100 companies with the highest score were selected as global tech leaders. WIPO-USPTO patent data of all firms provided by AcclaimIP were used and matched with financial data to construct the panel data. Patents that were filed and published during 1947 and 2019 were extracted from the database and with this data perpetual inventory method (2.3) was carried out to construct annual patent stocks. Further explanation on this will be made below. Figure 1 shows the graph that number of patents against their application year. The graph shows truncation bias since the data was collected based on published year.

Figure 1: Number of patents



Using the financial analysis that Bloomberg database provides, information on real sales, tangible asset and total number of employees were acquired from each firm's balance sheet and income statement. With this information, Q, K, N in (2.2) is proxied each. To convert nominal values to real values, World Bank's GDP deflator data with 2015 as the base year was used. Since the subjects of research are global firms and they have various nationality, GDP deflator of the country they are each located in was applied to. This would be reasonable since the information that is reported through a firm's audit is made in the country's own currency. The nationalities of the firms that are analysed in this research are as follows: Australia, Canada, Switzerland, China Germany, Finland, France, United Kingdom, Hong Kong, India, Ireland, Japan, Netherlands, South Korea, Sweden, United States and Taiwan(as China for political reason)

Regression analysis is carried out using the 1980-2019 panel data of firms for which WIPO-USPTO were all matched with its financial data. Data with missing observations such as sales, asset, total number of employees were excluded from the analysis. As a result of this process, 98 firms and observations 2,161 and 1,500,000 patent information are relevant sample for the analysis.

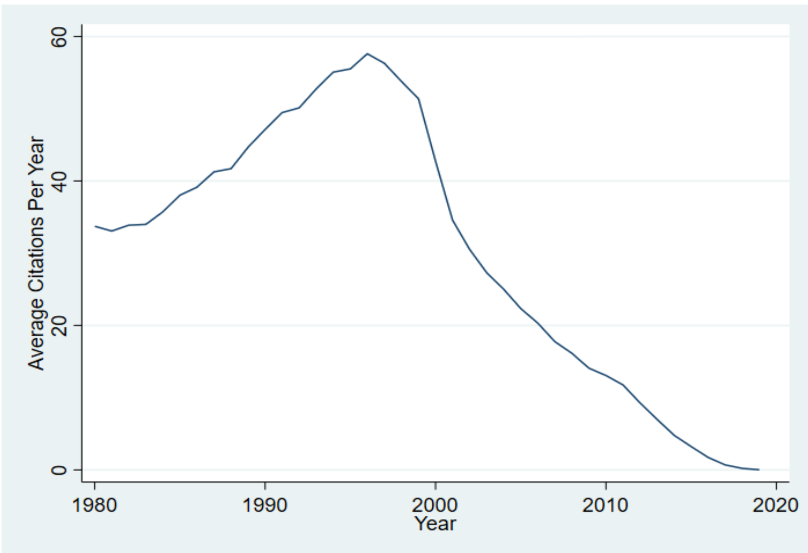
3.2 Patent stock

Among several ways of combining WIPO and UPSTO patents to calculate patent stocks, a naïve sum of the patents filed to the two patent offices was used for this research. Although this method may be convenient, it may cause two problems. The first problem is that patents published in WIPO and patents published in UPSTO may be in the same family so that multiple patents may actually be describing one technology. Just summing the patents of two patent offices without taking this into account may cause patent stocks to be overcounted. The second problem is that patents published in WIPO and UPSTO may differ in their citation pattern. USPTO is actively used by engineers from all around the world in US, and it is more likely that a patent published in USPTO will be browsed and cited by more people than a patent with similar level of technology that is published elsewhere. Therefore, simply adding the patents of the two offices without any modification may cause issues. To solve these problems, accurate information regarding patent family is needed and more research on ways to adjust for size effect should be made. The solution for the first problem can be solved through using the filter that is provided by AcclaimIP. This filter removes all documents from the search results except for the representative document of each application family(AFAM)[1]. Over-representing problem due to multiple patent documents published from a single filing can be resolved by turning on this filter. However, this research does not discuss the ways to adjust for size effect. Areas for further research remain in this sense.

As mentioned in Section 2, there is a need to consider the quality, and not only quantity, of patents when using patents as proxy variable for knowledge variable. In this line, two ways of calculating patent stock – simple summation of patents (patent stock) and summation that considers the weight for the citation counts 'citation-adjusted patent stock'– are used and compared.

One of the problems that occur when using patent information is the data disconnection issue that occurs from citation time lag. Citation may occur throughout decades, so despite having same scientific value, patent that was published 30 years ago will have been cited more than the one published cited 3 years ago.

Figure 2: Average citations per year



[1] An application family is the set of publications derived from one filing in one country or jurisdiction, which commonly includes an application and a grant, but can include supplementary documents including search reports and bibliographic reprints.

There are many ways to improve the data disconnection issue that arises from citation time lag. This research will use one of the two methods that Bloom and Van Reenen (2002) suggested. That is, assuming that the number of citations during a patent's lifetime will be equal regardless of year of publishment, yearly weight is given to normalize the number of citations so that it solves the issue of downward bias. In case of WIPO-USPTO data that was used in this research, even without using Fourier expansion on yearly citation counts, the shape of curve is smooth enough as depicted in figure 2 unlike those observed in Bloom and Van Reenen (2002). In this case, executing a Fourier expansion will distort the curve even more than intended. Therefore, number of citations was normalized using log of yearly average citation count without using Fourier series.

3.3 Descriptive statistics

Table 1: Descriptive statistics

Variables	Median	Mean	Std.Dev	Min	Max
Real Capital($m)	2,421	7,706	16,852	1.4	199,884
Employment	20,105	45,871	66,591	44	798,000
Real Sales($m)	5,907	15,902	27,805	0.7	261,660
Patents	3,230	14,994	28,320	0	173,731
Total Citations	48,278	192,933	346,499	0	2,275,654
Observations per firm	23	22.3	7.28	1	34

The main characteristic of descriptive statistics is that the standard deviation of every variable is very high. Also, the difference between mean and median implies that every variable is asymmetrically distributed. This positive skewness may be incurred by some giant tech companies among the 100 samples. For example, the number of patents of Samsung Electronics was 12,730 in 2015 while the average number of patent activity was 837 in the same year.

4. RESULTS

4.1 Regression result

Result based on the panel data constructed by observations of 98 firms in Section 3 is depicted in Table 2. Numbers inside the parenthesis are standard error for parameter estimator. Column 1 and column 2 shows OLS estimator of production function of 98 tech firms. Column 1 is the estimation result that does not account for firm specific fixed effect, and column 2 does account for it. Both columns' beta and gamma estimators show significance at 0.1% level. Also, both columns show constant return to scale where the sum of β and γ estimators are approximately 1. The adjusted R-squared value for regression analysis is 0.838 and 0.851 each.

Table 2: Regression result

Log Real Sales	(1)	(2)	(3)	(4)	(5)
Log Capital	0.347***	0.337***	0.286***	0.278***	0.276***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Log Employment	0.694***	0.706***	0.62***	0.614***	0.61***
	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)
Log Patent stock			0.299***		0.062
			(0.019)		(0.032)
Log Cite_adj Stock				0.306***	0.26***
				(0.017)	(0.029)
Firm dummies	no	yes	yes	yes	yes
Time dummies	yes	yes	yes	yes	yes
Adj. R-Squared	0.838	0.851	0.867	0.872	0.872
No. firms	98	98	98	98	98
No. observations	2,187	2,187	2,187	2,187	2,187

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

From column 3, firm specific fixed effect is continuously adopted. It differs from Bloom and Van Reenen (2002) in that parameter estimator for capital of tech firms are not as capital intensive as those firms suggested by Bloom and Van Reenen (2002). Rather, one can see that the parameter estimator for labour input is 0.6, which is close to the overall average that Bloom and Van Reenen (2002) suggested.

Column 4 presents parameter estimator where patent stock without weight for citation count was used as proxy variable for knowledge capital. Column 5 presents the result of regression analysis where normalized citation count was weighted for calculation of patent stock and the patent stock was used as a proxy variable for knowledge capital. The parameter estimator for labour and capital still shows significant result at 0.1% level, and the parameter estimator for patent stock is significant at 0.1% level each. In column 5, the parameter estimator of citation count-weighted-patent stock being 0.26 means that when the patent activity doubles, total factor productivity can increase by 26%. Parameter estimator for patent stock is not significant even at 5% level unlike citation adjusted patent stock which is significant at 0.1% level.

These results suggest that when using patent activity as a proxy variable for knowledge capital, giving weights for citation counts to calculate patent stock is able to deliver more information than otherwise. We can see from here that it is meaningful to use citation information when measuring how much the knowledge capital contributes to productivity. The value of citation information has also been noted in research of Hall, Jaffe, and Trajtenberg (2005) where the corporate value of patent was studied.

4.2 Robustness check

Table 3: Robustness check

Log Real Sales	(1)	(2)
δ	30%	
Log Capital	0.26***	
	(0.011)	
Lagged Log Capital		0.327***
		(0.012)
Log Employment	0.605***	
	(0.014)	
Lagged Log Employment		0.461***
		(0.014)
Log Cite_adj stock	0.171***	
	(0.036)	
Lagged Log Cite_adj stock	0.141***	0.242***
	(0.03)	(0.016)
Firm dummies	yes	yes
Time dummies	yes	yes
Adj R-squared	0.872	0.857
No. firms	98	97
No. observations	2161	2058

* p < 0.05, ** p < 0.01, *** p < 0.001

Subsection 4.2 shows the results of robustness check of regression analysis that was carried out above. It may take time for the implementation of new technology to be actually used in production and to increase the total factor productivity. This is because it takes time to prepare the production process of a new product and to marketize the core technology to suit consumers. In this case, it may be more appropriate to use patent stock from previous year, and not this year, for the proxy variable of this year's productivity. Therefore, I aim to analyze the influence of patent activity on total factor productivity using lagged variables of patent stocks.

Regression analysis conducted in column 5 of Table 2 is repeated using all patent data from WIPO-USPTO, and adopts lagged citation adjusted patent stock variable. The result of this analysis is presented in column 1 and 2. In other words, patent stock that was adjusted with its was used as the proxy variable for knowledge capital in Table 3. Column 1 conducts regression analysis by adding a time lag regarding the patent stock weighted for citation count to the model of column 5 of Table2. In column 2, time lag for capital, labour, and knowledge capital is introduced for analysis. When conducting regression analysis that uses time lag, there are two points to consider. The first is about the result of the analysis. In column 1 the current year's patent stock and previous year's patent stock is added to the model for the regression analysis. The difference of significance between the two variables can be interpreted as which variable carries more information in estimating productivity. If the previous year's patent stock has more significant result, this can be interpreted as the existence of time lag between the patent activity and actual implementation in production. The second point is the number of samples. In column 2 and 4, time variable is added not only for patent stock but also capital and labour inputs. This caused the analysis period to be 1981 – 2019, which is 1 year shorter. As a result, the number of sample firms subject for the analysis is 97, and the number observations of panel analysis that is carried out by pairing firm and year is 2,058, which all has decreased.

From Table 4, one can see that adjusted R-squared value of regression analysis is about 0.87, which is still high. This implies that regression model that was set in this section has high explanatory power. Column 1 shows that the influence of patent stock that adopts time lag on productivity is significant at 0.01% level. The productivity contribution of patent stock that did not adopt lag was significant at 0.01% level. This result is notably different from that of Bloom and Van Reenen (2002), where analysis with time lag had more significant result than the analysis without time variable. This implies that it does not take a long time for a patent to be filed and be reflected in the production function. Lee (2013) suggested countries like Taiwan and South Korea were able to make an economical leap from developing countries to the level of developed country by focusing on industries with short product life cycle in the 1980s. He mentioned industries like electronics, electrics, computer, and telecommunication have a short product life cycle in which innovative activities have an immediate impact on factor productivity. However, the estimator of patent stock with time lag is also significant at 0.01% level in column 1. One possible explanation is that even if the product life cycle is short in tech industries, accumulation of core technology that is necessary for production also play a huge role in global tech companies.

5. CONCLUSION

By analyzing global tech firms that has patent activity during 1980 and 2019, this research showed that patent stock can explain part of real sales that is not explained by tangible assets or labour inputs, with statistical significance. Patent stock that considers patents' citation information better explained a firm's sales productivity and showed higher correlation with sales productivity than simply calculated patent stocks. These findings suggest that patent information can be used for measuring knowledge capital of global tech firms, and especially that patent stock that accounts for patents' quality (citation count) can be used as a proxy variable that represents a firm's knowledge capital. Moreover, by comparing the regression result with precedent research, this research suggests a specific trait of global technology companies. However, there are two main limitations to this research. The first downside is that the construction of patent stock could be done more precisely. The different size effect between WIPO and USPTO is ignored in the patent stock construction. The second downside is that research fails to provide a clear-cut explanation of the significance of lagged citation-adjusted patent stock in robustness check. Rooms for further research remains.

REFERENCE

Bloom, N., & Van Reenen, J. (2002). Patents, real options and firm performance. *The Economic Journal*, *112*(478), C97-C116.

Cockburn, I., & Griliches, Z. (1988). The estimation and measurement of spillover effects of R&D investment-industry effects and appropriability measures in the stock market's valuation of R&D and patents. *The American Economic Review*, *78*(2), 419-423.

Griliches, Z. (1981). Market value, R&D, and patents. *Economics letters*, *7*(2), 183-187.

Griliches, Z. (1990). *Patent Statistics as Economic Indicators: A Survey. part 1-2* (No. 3301). National Bureau of Economic Research.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). *The NBER patent citation data file: Lessons, insights and methodological tools* (No. w8498). National Bureau of Economic Research.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. RAND Journal of economics, 16-38.

Kogan, L., Papanikolaou, D., Seru, A., & Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics*, *132*(2), 665-712.

Lanjouw, J. O., & Schankerman, M. (2004). Patent quality and research productivity: Measuring innovation with multiple indicators. The Economic Journal, 114(495), 441-465.

Lee, K. (2013). *Schumpeterian analysis of economic catch-up: Knowledge, path-creation, and the middle-income trap*. Cambridge University Press.

Pakes, A., & Griliches, Z. (1980). Patents and R&D at the firm level: A first report. *Economics letters*, *5*(4), 377-381.

Pakes, A. (1985). On patents, R & D, and the stock market rate of return. *Journal of political economy*, *93*(2), 390-409.

Schankerman, M., & Pakes, A. (1986). Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period, 96 ECON. *J*, *1052*, 1052.

Thomson Reuters (2018). Top 100 Global Tech Leaders. Retrieved from https://www.thomsonreuters.com/en/products-services/technology/top-100.html

Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. *The Rand journal of economics*, 172-187.