Does Tokyo Stock Exchange Appreciate Corporate Innovations? Role of Patents' Quality and Research Productivity

Kostiantyn Ovsianikov, Ph.D.

Graduate School of Business Sciences, Humanities, and Social Sciences,
University of Tsukuba
Tsukuba, Japan

ovsianikov.kosti.gf@un.tsukuba.ac.jp

ORCID: 0000-0002-5255-3913

Abstract

This paper examines the market perception of corporate innovations in Japan. It follows the research question formulated by Hall, Jaffe, and Trajtenberg (2005): “how does innovative activity translate into market value, and what aspects of the underlying process are captured by the empirical measures available?”. The novelty of my study is twofold. First, it embraces the longitudinal innovation- and finance-related corporate records to come up with the largest ever combined data-set for Japan that encompasses 632 companies listed at the Tokyo Stock Exchange over the period of 19 years. Second, in addition to linear regressions, it applies the generalized additive models (GAMs). The latter technique allows for realistically capturing nonlinear patterns present in the data while at the same time retaining predictive features of a model. The main finding of the article is following. Amid the dominant role of research and development (R&D), especially for the Pharmaceutical and Chemical industries, market consistently rewards influential patents in the manufacturing sector.

Keywords: Patents; R&D; Tobin's q ; Generalized additive models.

Declarations

- **Funding.** This article is a part of the research project funded by the JSPS Grant-in-Aid for Young Scientists (B).
- **Conflicts of interest/Competing interests.** None.
- **Code availability.** The R code is available upon request.
- **How to cite.** This article is available as a working paper at [SocArXiv](https://arxiv.org/abs/2008.00000). Ovsianikov, Kostiantyn. 2020. “Does Tokyo Stock Exchange Appreciate Corporate Innovations? Role of Patents' Quality and Research Productivity.” SocArXiv. [doi:10.31235/osf.io/k4ys6](https://doi.org/10.31235/osf.io/k4ys6).

1 Introduction

Corporate assets can be divided into *tangible* and *intangible*. The former are directly linked to profitability and can therefore be measured with the reference to corporate financial data. The quantitative display of the latter, on the other hand, is not straightforward. Part of the intangible assets is captured by the research and development (R&D) expenses. While R&D is a strategic driver of innovations, one also needs to account for the effectiveness of the R&D spending. The output measurements of innovations commonly used in the existing literature are patents and patent citations.

It may seem that by knowing the “input” and the “output” estimations pertaining to technological innovations we can at least intuitively evaluate firm’s intangible assets. However, there is much more to add to this story. For example, Griliches (1998) (p. 299) presents patents as merely quantitative metrics for inventions that do not play an explicit economic role *per se*.¹ They rather contribute to an unobservable stock of knowledge which, in turn, might be subject to specific institutional environments – both at the national and at the industrial levels.

Patents’ count mainly helps us to comprehend the *quantitative* dimension of corporate innovations. In order to evaluate the *quality* of innovations, however, one must impose measurement restrictions such as size and industrial profiles of the companies analyzed. Furthermore, as older firms would have a higher likelihood of accumulating more patent counts till the present day, and, as older patents are more likely to have more citations, one has to address the “truncation problem”.

Finally, there exist various ways to measure the perceived impact of innovative activities. Although it may seem justified to relate the innovative output to corporate financial profitability, in practice, it takes an indefinite number of years till a patenting activity can yield profits (Czarnitzki, Hall, and Oriani 2006). Therefore, following the established line of research originated in Griliches (1981), the current paper chose to concentrate on the issue of market evaluation operationalized via the *Tobin’s q* (market to book ratio) metric. On one hand, the main limitation of using Tobin’s q relates to market volatility. In other words, investors’ concerns are not necessarily related to firm’s physical assets. The main advantage of using this parameter, on the other hand, is the fact that market does not take a long time to evaluate technological capabilities of companies. Hence, this strategy helps a researcher to circumvent the data loss due to the otherwise lagged dependent variable. This and the related methodological issues are covered in a detail in the Section 3.

2 Literature

The research on the market evaluation of patents dates back to 1960-70s (Scherer 1965; Schmookler 2013), while the first studies using the firm-level data appeared in 1980s (Griliches 1984). These early empirical works solely relied on the patent counts, which represent the quality of innovations only fragmentarily (Hall, Jaffe, and Trajtenberg 2005). As the nature of patented inventions can substantially differ, the count of patent citations is believed to provide a more nuanced picture of corporate innovations (Acharya, Baghai, and Subramanian 2013). In other words, the distribution of citations per patent is very uneven, meaning that there are relatively few highly influential (i.e. highly cited) patents, whereas the rest have mediocre to minor impact (Kleinknecht, van Montfort, and Brouwer 2002; Trajtenberg 1990).

While citations allow to control for patent quality, their usage presents certain methodological challenges known as truncation problems. According to Hall, Jaffe, and Trajtenberg (2005) (p. 17) “citations to a given patent typically keep coming over long periods of time, but we only observe them until the last date of the available data”. I adopted the established methodology of using citation *stocks*, also accounting for their depreciation rate over time. Furthermore I included another control variable used in the previous literature

¹While Griliches (1998) specifically refers to the limits of patent *applications* data, current research takes into account only *granted patents* to partially mitigate this issue. However, the downside result of this strategy is a reduced data set. I touch upon this issue in a more detail in the Section 3.

called *citations per patent*. This parameter introduces the qualitative dimension to otherwise quantitative metrics of “patents’ stock weighted by assets”.

Another influential contributor to firm’s intangible assets is its R&D spending. In fact, R&D expenditures play even more important role for the market evaluation than patents’ stock (Hall, Jaffe, and Trajtenberg 2005, 20). As such, higher R&D expenditure indicates that a company makes strategic investments, which signals about its long-term credibility. At the same time, investors must be sure that R&D spending, which represents the input-aspect of innovations, bears fruit. In order to operationalize the R&D efficiency, scholars suggest using the “patent-yield” variable obtained via dividing patents’ stock by R&D stock (Hall, Jaffe, and Trajtenberg 2005, 27). This variable can be described as “patents per one yen spent on R&D”. While R&D normally accounts for a more pronounced positive market reaction, this may differ depending on an industry analyzed.

Based on the sample of Italian, German and Dutch companies, Brouwer and Kleinknecht (1997) argue that R&D expenditure is of a higher importance for manufacturing sectors, while investment in fixed assets (which is a marker of capital intensity) has a higher impact for the service-related companies. I compare my findings to those of Brouwer and Kleinknecht (1997) in the Section 4.1.

2.1 Japan’s context

As most of the hitherto research on the market perception of corporate innovations has concentrated on the U.S. and Western Europe, not much attention has been paid to Japan. It matters for the following reasons. Firstly, Japan has a third-largest market capitalization for listed companies. Secondly, alongside with the growing market for corporate control, Japan has distinct institutional arrangements encapsulated by the pertaining legacy of relational corporate finance and governance. Hence, it presents a rich material for scholars focusing on interaction between markets and institutions.² From a comparative perspective, Japanese companies stand out for placing comparatively high value into patenting their inventions. In fact, Cohen et al. (2002) argue that the capacity to patent innovations is akin the transmission of a message about the efficient usage of R&D funds among companies occupying a similar niche.

The pioneering research on the impact of technological assets on a firm’s perception by market in Japan was performed by Haneda and Odagiri (1998). Their analysis relies on the sample of pharmaceutical, chemical, and electrical equipment firms. There exist three major problems with this study. Firstly (1), it operates with the small sample size of only 90 companies. Since the authors include three industries (“Drugs”, “Chemicals” and “Electrical”) into their analysis, this arouses doubts about the validity of industrial representation. The second (2) issue follows from the one above. Although the most influential relevant articles use random-effects models with an optional addition of industry-level fixed effects, Haneda & Odagiri prefer the firm-wise fixed effects. They justify it by the fact that the Hausmann Test rejected the null hypothesis about the nonexistent individual effects. This, however, does not automatically show the legitimacy of including firm-level fixed effects. The decision has to be ultimately justified by the economic theory, which, in this case, presents persuasive counter-arguments, such as that firms tend to *collectively* (at least on an industrial level) change their strategies for meeting market expectations (Hall, Jaffe, and Trajtenberg 2005, 26). The last (3) problem is that the authors do not control for the quality of patents’ stock by accounting for citations.³

The third aspect was taken care of by Nakanishi and Yamada (2007). Another advantage of their article is that it only focused on the successful patent applications. At the same time, the size of the data-set hardly increased (only 101 companies distributed across drugs, chemicals, electrical machinery, precision instruments and transportation equipment industries), which again raises the concerns about representativeness. Touching upon details, authors do not explicitly discuss the values of obsolescence rates pertaining to knowledge stock variables, which would have been a useful reference for those willing to reproduce their study.

²Back in the day, when most of the research on the market valuation of intangible assets was based on the U.S. data, Hall and Oriani (2006) made a claim of conducting the first study on the continental countries of the European Union.

³On p. 314 the authors admit that “the *quality* of patents more than the *number* of patents may well be a better measure of the firm’s technological assets”. However, they do not mention the obvious possible solution, which is to control for patent citations.

The latter problem was addressed in the article of Ishii (2010). He explicitly mentions the adoption of the depreciation rate of 15% per year for the patents and citations stocks. At the same time, Ishii’s method of calculating R&D stock is unclear. The author also stressed the rationale for not including firm-level fixed effects, which, as mentioned above, is justified by the influential related studies such as the one by Hall, Jaffe, and Trajtenberg (2005). The main confusion that the article by Ishii (2010) causes is related to the choice of the dependent variable. Instead of using the most widespread option, which is the *Tobin’s q*, author opted for the so-called “value added” without a robust reference to previous studies. As mentioned in the Section 1, this choice would cause a need for introducing lags for the dependent variables – the fact that the author did not consider.

The last currently available study on this topic with the focus on Japan belongs to Mahlich (2010). He concentrated on the impact of technological innovation on corporate performance within the Japanese pharmaceutical industry. His approach seems problematic for the same reasons discussed in regard to the article of Ishii (2010). The author claims that “the use of market value receives a lot of criticism” without providing any supporting references. As a result, the choice of the “profit on total assets” as a dependent variable seems very arbitrary. In addition to this, the analyzed sample consists of 34 Japanese pharmaceutical companies, which is certainly not enough to come up with industry-specific generalizations. Lastly, the author justifies the choice of the depreciation rate for patents’ stock in a following way: “We assume δ [obsolescence rate] to be 20% which is between the 13.5% used by Haneda and Odagiri (1998) and the 25% estimated by Pakes and Shankerman (1984)”. Needless to say that an approach based on averaging the rate from the two studies separated by 14 years and concerned with different countries lacks a proper justification. In a nutshell, the author’s conclusion about the absent effect of patenting on economic performance cannot be compared to other related studies due to methodological issues.

As we can see from the summary of literature pertaining to the case of Japan, there has not been a single comprehensive study answering the question of market evaluation of corporate technological assets. Furthermore, due to the previously existing data limitations, the number of observations was much lower compared to the researches centered on the U.S. or the European cases.

3 Methodology

The whole working process, including manuscript preparation, statistical analysis and data visualization was conducted in the R-environment (R Core Team 2020). The article was put together with the help of *knitr* (Xie 2020) and *bookdown* (Xie 2016) packages. I did most of the data wrangling and visualization with the *tidyverse* toolkit (Wickham et al. 2019). Packages such as *plm* (Croissant and Millo 2018) and *mgcv* (Pedersen et al. 2019) enabled me to run regressions. Finally, *kableExtra* (Zhu 2019), *pander* (Daróczi 2014), *stargazer* (Hlavac 2018), *ggcorrplot* (Kassambara 2019) and *reporttools* (Rufibach 2009) packages helped me to accurately present the results of my statistical analysis.

3.1 Data

The data-set used in this article is the result of merging two files: one, obtained from the Nikkei NEEDS Financial Quest, includes corporate financial information, another, extracted from the Institute of Intellectual Property Patent Database (IIP-DB) – captures companies’ technological activity registered with the Japan Patent Office (JPO). The latter source was originally compiled by Goto and Motohashi (2007). Following the example of Hall, Jaffe, and Trajtenberg (2005), I included the companies from pharmaceutical, chemical, machinery and electrical equipment industries as specified in the NEEDS Database.

The total count of companies is 632, with each unit extending over 19 years (from 1998 till 2016), making the obtained time-series considerably larger than any of those used in existing studies on Japan. My sample is also comparable to the majority of the international studies cited here. The resulting data-set analyzed in this article contains 7,813 observations. For the comparison, the data-set employed by Sandner (2010)

encompasses 6,757 observations. My data includes fewer companies but stretches across a longer time-frame, which, at the end, brings the number of total cases to the common benchmark.

The problem of classifying patents was addressed by Griliches (1998). On one hand, there is a “target” approach based on the classification of the World Intellectual Property Organization (WIPO) or the National Bureau of Economic Research (NBER). The IIP-DB, on which the current study relies, uses this classification as well (Goto and Motohashi 2007, 1433). It answers the question which field can the respective patent be most fitting at. This, however, entails serious problems for the firm-level study, as the same company can file patents belonging to multiple classes.

“The Patent Office classifies patents into many classes (300+ in the mid-1950s) and even many more subclasses (over 50,000), based on its need to ease the search for prior art. The resulting classification system is based primarily on technological and functional principles and is only rarely related to economists’ notions of products or well-defined industries (which may be a mirage anyway). A subclass dealing with the dispensing of liquids contains both a patent for a water pistol and one for a holy water dispenser.”

— (Griliches 1998, 292)

Hence, the alternative “origin” approach adopted in this article “starts from patent totals for particular firms and then groups them into industries according to the firm’s primary activity” (Griliches 1998, 294). Following the methodological suggestion by Haneda and Odagiri (1998) (p. 309) that was implemented in all the subsequent works related to Japan, I used the industrial classification present in the NEEDS Database.

The chronological frames are motivated by the following circumstances. On one hand, the lower limit was dictated by the earliest availability of the market-related corporate information. On the other, the latest accessible year in the IIP-DB for the patent applications approved ex post is 2019. However, it takes on average several years for an application to be either rejected or granted. That is why the years 2017 to 2019 so far contain very few information, which will however become more substantial as time passes. As a result, 2016 is the last significant year, for which the patent-related data is available. The length of the patent-registration process is one of the reasons why the curves on the Figure 1 are downward-sloping from the mid-2000s. Application-filing process is especially time-consuming for the Pharmaceutical industry, where the average period for granting a patent is 7.1 years – well above the cumulative mean of 6.45 years (see also the Table 1).

Table 1 Average Timing (in Years) of Patent Grants per Industry

	1971-1980	1981-1990	1991-2000	2001-2014	Combined
Chemical	6.87	8.09	7.64	4.81	6.66
Electronics	7.28	8.66	7.29	4.38	6.67
Machinery	6.64	7.83	6.73	4.51	6.25
Motor	6.56	8.29	7.17	4.37	6.39
Pharm	7.59	8.34	8.24	5.04	7.10
Precision	7.62	8.94	7.93	4.65	7.05
Cumulative	6.90	8.14	6.95	4.55	6.45

Although an average time for receiving a patent within the Chemical industry throughout 1971-2014 was roughly the same as within the Electronics, the Table 1 shows that starting from 1990s Chemical-related applications have required more approval time than the Electric-related ones. Hence, especially for the period analyzed in this article, most challenging conditions for licensing appear to be in Chemical and the related Pharmaceutical industry. This goes hand in hand with the importance of patenting activities for these sectors both in Japan and internationally (Cohen et al. 2002; Hall, Jaffe, and Trajtenberg 2005). In general, it is certainly too early to provide the full information on the number of granted patents filed during

the last decade. However, this problem is mitigated by accounting for patents' *stock* rather than flow, which I discuss in the Section 3.2.

Fig. 1 Timing for turning patent applications into grants

While corporate financial data (except the R&D spending, which I touch upon in the Section 3.2) does not entail any particular technical issues, patent-related information needs to meet some criteria before running a regression. Following the methodology adopted in the existing literature, I have used the entries based on the *date of filing the patent application*, which is a point in time, from which a corresponding invention would be legally protected. However, the copyright mechanism can only be enforced if a patent is subsequently granted. In this case, a corresponding company can receive benefits associated with license possession, apart from the technological improvements an invention may generate (Czarnitzki, Hall, and Oriani 2006, 123). This is the main reason behind choosing only the successful patent applications for this study.

In fact, as seen from the Figure 2, the median granting success rate of patents filed with the Japan Patent Office (JPO) from 1971 to 2014 was 37 percent. This number is very similar to the 35 percent in Germany and is considerably lower compared to countries such as France (90 percent) and the U.K. (80 percent) (Griliches 1998, 289). Although the granting rate has been on a sharp rise in Japan since early 2000s, it is still indicative of a stringent control mechanism that the local application process entails. Schankerman (1998) notes the following: “Patents originating in Japan are substantially more valuable than in other countries, and the difference is most striking for electronics patents” (p. 80). This trend is depicted, among others, on the Figure 2, where Electronics have the lowest granting rate of 30 percent.

At the same time, the comparatively high industrial success rates for Pharmaceuticals (50 percent) and Chemicals (48 percent) do not automatically imply the low quality of the respective patents. It is rather a cause of patents being particularly important for realizing the gains from intangible assets. Differently from Electronics, “in the pharmaceutical industry, patents are perceived as a relatively effective tool of appropriating rents from technological innovations as compared to other types of mechanism such as trade

secret, increasing complexity of product design and faster market introduction” (Motohashi 2008, 1548). This has been a common finding across numerous studies concluding that the knowledge assets have the most pronounced role for the market evaluation of companies within the Pharmaceutical sector (Czarnitzki, Hall, and Oriani 2006, 15). Finally, if we look at the yearly cohorts displayed in the Table 2, the figures for Pharmaceuticals and Chemicals within the period of 1991-2000 (part of which is included in this article’s econometric analysis) hardly surpass the cumulative median.

Table 2 Median Ratio of Grants to Patent Applications per Industry

	1971-1980	1981-1990	1991-2000	2001-2014	Combined
Chemical	0.52	0.45	0.34	0.55	0.48
Electronics	0.31	0.24	0.28	0.46	0.30
Machinery	0.45	0.38	0.38	0.52	0.41
Motor	0.44	0.38	0.43	0.50	0.44
Pharm	0.58	0.51	0.36	0.50	0.50
Precision	0.35	0.38	0.28	0.42	0.36
Cumulative	0.38	0.33	0.34	0.49	0.37

Fig. 2 Granting success rate of patents filed with the JPO

The lesson that follows is that the patents are only important for companies inasmuch as they originate from the increase of a “knowledge stock” or “intangible assets”. This is the relationship between “ ΔK ” (unobservable gain in knowledge) and “Pats” shown in the Figure 3. While “Pats” is an observable *output* variable, “R&D” represents observable *inputs* to the knowledge stock. Since “Pats” are not solely determined by intangible assets’ increase, this is reflected in the presence of unobservable impacts ν . Likewise, intangible

assets depend not only on the R&D expenditures but also on the other input factors marked as ω . “Z’s”, on the other hand, denote economic benefits stemming from “ ΔK ”. As we will see later, among others, “Z’s” personify market perception of corporate tangible and intangible assets, which constitutes the gist of this paper. “Z’s” indices are additionally influenced by observable inputs “X’s” “such as investments and labor input” (Czarnitzki, Hall, and Oriani 2006, 116), as well as unobservable effects “ ε ’s”.

Fig. 3 The measurement of knowledge capital (Pakes and Griliches 1984)

3.2 Model

Differently from most of the financial variables that represent a *flow*, i.e. a value at a certain time-point exclusively, knowledge-related items accumulate their value over time, thus corresponding to a type of a *stock*. Since all the relevant time-series data is originally available as a flow, one needs to justify the technique of converting it to a stock. Below is the example of distinguishing two measurements:

“Income is a flow. It corresponds to the quantity of goods produced and distributed in a given period (which we generally take to be a year). Capital is a stock. It corresponds to the total wealth owned at a given point in time. This stock comes from the wealth appropriated or accumulated in all prior years combined.”

— (Piketty 2017, 63)

R&D stock RS_t is calculated based on Cockburn and Griliches (1988) using the following formula:

$$RS_t = RD_t + (1 - \delta)RS_{t-1} \quad (1)$$

where RD_t is R&D expenditures during the time t (year of the observation), δ is the R&D yearly depreciation rate, and RS_{t-1} is the R&D stock accumulated during years preceding the time t . The patent stock PS_t and citations stock CS_t are calculated using the same method.

3.2.1 Depreciation rates and licences' length

Since the current study is preoccupied with Japan, I have adopted the yearly depreciation rates of patent stock from Nakanishi and Yamada (2008). According to them (p. 8), the rate of obsolescence for Electrical industry is 24.2 percent, for Chemicals – 32.3 percent, for Pharmaceuticals – 27.4 percent, for Precision Machinery – 24.5 percent and for Transportation Machinery – 26.1 percent. The previous obsolescence rates used in the Japan-related studies were based on Haneda and Odagiri (1998), who in turn, relied on the 1985 data provided by the Science and Technology Agency (p. 319). Indeed, those rates may no longer be accurate today.

On the international level, most of the articles adopted the approach by Hall, Jaffe, and Trajtenberg (2005) who used the conventional depreciation rate of 15 percent for all industries. Although I follow Hall, Jaffe, and Trajtenberg (2005) in using the same rates of obsolescence for citations, R&D stock and patents, I rely on Nakanishi and Yamada (2008) in adopting industrially-specific figures. Hall, Jaffe, and Trajtenberg (2005) admit that “this is an issue that deserves some serious revisiting, in light of the much more detailed data that we now have at our disposal” (p. 24).

Depreciation rates are calculated based on “the rate of decay of the private revenues accruing to industrially produced knowledge” (Nadiri and Prucha 1996, 51). As Pakes and Schankerman (1984) (p. 75) stress, the depreciation of knowledge stock depends on the *market valuation* rather than on a physical decline in technological assets. Market opinion is based on the fact that fresh innovations outshine the previous ones, making it difficult to appropriate rents from inventions over long periods of time. One of the components of the knowledge stock decay is the length of patents' licenses (Pakes et al. 1989). The shorter their lifespan is, the more intangible assets would depreciate annually. The fact that a firm chooses not to pay an extension fee for its patent is the sign that the accompanying “maintenance costs” outstrip the profits accruing to a license holder. Shorter lifespans of patents therefore reflect an intensive industrial competition that prompts companies to develop new technologies rather than to hold onto diminishing rents from older knowledge assets.

Figure 4 shows the general correspondence of the obsolescence rates to the industrial patterns of Japanese companies. The current research is based on the “origin” classification of Nikkei NEEDS, as shown on the *left half* of the Figure 4. Here, companies belonging to the Chemical industry have one of the longest lifespans (7.4) for patents' licenses. On the other hand, Pharmaceuticals have the lowest figure of 6.69 years. Notably, this industrial order of patent licenses from lowest to shortest agrees with the one identified for Japan by Schankerman (1998) (p. 83).

On the other hand, Nakanishi and Yamada (2008), whose rates I adopt, apply the “target” approach demonstrating where the invention “belongs”. It is depicted on the *right half* of the Figure 4. In this case, the highest obsolescence rate for the Chemical industry indeed corresponds to the lowest lifespan of its patents (7.01 years). Conversely, the lowest rate for Electronics reverberates with the relatively long lifespan of 7.26 years (7.31 years for the ICT industry). Moreover, differently from the “origin” approach, Pharmaceuticals have one of the longest patents' validity of 7.28 years. To sum up, I group the companies by “origin”, but adopt the obsolescence rates based on the “target” of innovation. This seeming discrepancy does not constitute a methodological problem, since the rates of depreciation relate to *market perception* of typical industrial trends that are more likely to be estimated via the “target” approach. Thus, one may say that companies are subject to technological expectations pertaining to the destinations of innovations, notwithstanding their industrial profiles.

Fig. 4 Lifespan of patents' licenses by industry (NBER vs. NEEDS classification)

3.2.2 Baseline regression

To sum up the three main indicators of innovation: “R&D reveals the commitment of a firm’s resources to innovation, patents catalog the success in generating codifiable new knowledge that the firm can in principle appropriate, and citations indicate the extent to which those innovations turn out to be ‘important’ and hence presumably more valuable to the firm” (Hall, Jaffe, and Trajtenberg 2005, 24).

I operationalize the value of innovation (the “Z’s” from the Figure 3) via the *Tobin’s q*, known as a market-to-book-ratio. “Tobin’s q [...] is the ratio of the firm’s market value to tangible assets. Firm’s market value is defined as the sum of market capitalization (price multiplied by the number of outstanding shares at the end of the year) and non current liabilities plus net current liabilities less inventories” (Hall, Thoma, and Torrisi 2007, 12).

Most of the studies on this topic have followed the model developed in the seminal work by Griliches (1981) that was further elaborated in Cockburn and Griliches (1988). According to them, the market value of the firm, V , can be expressed as follows:

$$V_{it} = \alpha(A_{it} + \beta K_{it}) \quad (2)$$

where A stands for the firm’s tangible assets that include “plant, equipment, inventories, and financial assets” (Griliches 1981, 183), K represents the knowledge stock, and α – “current market valuation [...] of the firm’s assets, reflecting its differential risk and monopoly position” (Griliches 1981, 183).

As the definition of *Tobin’s q* is the ratio of the market value V of a firm to its book value A , we can rewrite the formula as:

$$q_t = \frac{V_{it}}{A_{it}} = \alpha \left(1 + \beta \frac{K_{it}}{A_{it}} \right) \quad (3)$$

Taking the logarithms, we get

$$\log q_{it} = \log \alpha_t + \log \left(1 + \beta \frac{K_{it}}{A_{it}} \right) + \varepsilon_{it} \quad (4)$$

where ε_{it} is the error term or the noise of the model.

Here, the knowledge stock K consists of three components, such as: *R&D weighted by Assets*, *Patents per yen of R&D*, and *Citations per Patent*. Incorporating these measurements, we can rewrite the last equation as follows, according to the suggestion of Hall, Jaffe, and Trajtenberg (2005) (p. 24):

$$\log q_{it} = \log \alpha_t + \log \left(1 + \beta_1 \frac{R\&D_{it}}{A_{it}} + \beta_2 \frac{Pat_{it}}{R\&D_{it}} + \beta_3 \frac{Cites_{it}}{Pat_{it}} \right) + \varepsilon_{it} \quad (5)$$

4 Data Analysis

Current analysis encompasses four industries: Electric, Machinery, Chemical and Pharmaceutical. Due to significant industrial patterns of technological innovations, authors such as Hall, Jaffe, and Trajtenberg (2005) include industrial effects into their regression analysis. I follow this methodology and present below the summary of my findings.

4.1 Overview

Table 3 General Descriptive Statistics

Statistic	N	Mean	Median	Min	Max	St. Dev.
R&D/Assets	10,463	0.363	0.221	0.0001	8.511	0.535
Patents/Assets	10,447	0.004	0.001	0.000	0.088	0.008
Citations/Assets	10,349	0.009	0.002	0.000	0.199	0.019
Patents/R&D	10,038	0.021	0.006	0.000	0.455	0.047
Citations/Patents	11,153	2.246	1.965	0.0001	12.538	1.588
Tobin's q	8,081	1.159	0.814	0.0001	7.937	1.174
R&D Intensity	10,532	0.039	0.030	0.00005	0.250	0.037
Process Innovation	11,187	1.444	0.948	0.002	12.733	1.578
Leverage	11,247	0.468	0.471	0.016	2.232	0.203
Capital Intensity	11,137	18.234	14.799	0.334	99.887	12.683
Exports	8,478	0.345	0.309	0.00004	1.000	0.221
ROA	11,269	0.043	0.042	-0.318	0.254	0.057

The summary statistics are displayed in the Table 3. Compared with the data used by Hall, Jaffe, and Trajtenberg (2005), the values of “market-to-book value” or Tobin’s q are almost identical. There are, however, notable differences among other variables. Having converted Japanese yen to U.S. dollars (the average monthly exchange rate from 1998 till 2016 was 107.625), we can see that the median value of “Patents per U.S. dollars spent on R&D” pertaining to the current data-set is considerably higher ($0.006 * 107.625 = 0.65$) than the one in the study mentioned above (0.51). This means that Japanese companies use their R&D funds more efficiently judged by the scope of the patented output.

On the other hand, the median ratio of “citations per patent” (the variable reflecting patents’ quality) for the current Japanese data-set is more than three-times lower (1.97) than the one recorded by Hall, Jaffe, and Trajtenberg (2005) (6.33). From this we can infer that despite being less numerous, the U.S.-originated patents are more influential, if judged by the number of acquired citations. The explanation of differences between Japan and the U.S. would be, however, incomplete without the reference to the institutional contexts of both countries. According to Cohen et al. (2002), “Japanese firms tend to use patents more to assure their inclusion and freedom of operation even in discrete product markets where American firms tend to use patents more for exclusion. The use of patents in Japan for maintaining market access and freedom of operation rather than for exclusion is reinforced by the comparatively low payoffs to infringement suits in Japan” (p. 1365).

Table 4 Industrial Summary

Variable	Industry	N	Mean	Median	Min	Max	St. Dev.	NA
RDA	Chemical	3191	0.336	0.264	0.002	5.544	0.332	442
	Electric	3323	0.553	0.325	0.000	8.511	0.787	597
	Machinery	3279	0.246	0.169	0.001	5.078	0.338	501
	Pharm	670	0.127	0.106	0.006	1.020	0.102	200
	all	10463	0.363	0.221	0.000	8.511	0.535	1740
PSA	Chemical	3116	0.004	0.001	0.000	0.088	0.008	517
	Electric	3333	0.004	0.002	0.000	0.071	0.006	587
	Machinery	3336	0.005	0.001	0.000	0.088	0.010	444
	Pharm	662	0.001	0.000	0.000	0.009	0.001	208
	all	10447	0.004	0.001	0.000	0.088	0.008	1756
CSA	Chemical	3114	0.009	0.001	0.000	0.199	0.022	519
	Electric	3268	0.010	0.005	0.000	0.165	0.014	652
	Machinery	3307	0.009	0.002	0.000	0.198	0.020	473
	Pharm	660	0.002	0.001	0.000	0.054	0.004	210
	all	10349	0.009	0.002	0.000	0.199	0.019	1854
PatRD	Chemical	2968	0.021	0.003	0.000	0.453	0.052	665
	Electric	3309	0.015	0.007	0.000	0.440	0.027	611
	Machinery	3131	0.031	0.009	0.000	0.455	0.058	649
	Pharm	630	0.006	0.003	0.000	0.038	0.007	240
	all	10038	0.021	0.006	0.000	0.455	0.047	2165
CitPat	Chemical	3358	2.358	2.002	0.001	12.424	1.811	275
	Electric	3532	2.295	2.066	0.046	12.538	1.419	388
	Machinery	3514	2.007	1.730	0.000	11.698	1.443	266
	Pharm	749	2.632	2.340	0.053	12.339	1.759	121
	all	11153	2.246	1.965	0.000	12.538	1.588	1050
Q	Chemical	2610	1.087	0.786	0.000	7.899	1.081	1023
	Electric	2412	1.197	0.771	0.000	7.826	1.268	1508
	Machinery	2402	1.021	0.690	0.000	7.937	1.088	1378
	Pharm	657	1.809	1.482	0.315	7.912	1.254	213
	all	8081	1.159	0.814	0.000	7.937	1.174	4122
RDInt	Chemical	3190	0.032	0.028	0.000	0.218	0.022	443
	Electric	3386	0.047	0.040	0.000	0.249	0.036	534
	Machinery	3288	0.024	0.020	0.000	0.246	0.021	492
	Pharm	668	0.111	0.095	0.008	0.250	0.059	202
	all	10532	0.039	0.030	0.000	0.250	0.037	1671
Process	Chemical	3367	1.509	1.147	0.002	12.733	1.438	266
	Electric	3567	1.301	0.795	0.002	12.148	1.502	353
	Machinery	3579	1.360	0.903	0.008	12.461	1.522	201
	Pharm	674	2.319	1.286	0.002	12.530	2.425	196
	all	11187	1.444	0.948	0.002	12.733	1.578	1016
Leverage	Chemical	3375	0.476	0.482	0.048	1.007	0.191	258
	Electric	3579	0.464	0.468	0.036	2.232	0.215	341
	Machinery	3589	0.487	0.494	0.043	1.342	0.198	191
	Pharm	704	0.352	0.310	0.016	0.946	0.186	166
	all	11247	0.468	0.471	0.016	2.232	0.203	956
Capint	Chemical	3344	25.275	22.788	0.562	93.800	13.400	289
	Electric	3561	11.374	9.916	0.334	99.574	8.637	359
	Machinery	3555	16.332	14.224	0.831	99.887	9.746	225
	Pharm	677	29.519	26.300	1.100	99.733	15.169	193
	all	11137	18.234	14.799	0.334	99.887	12.683	1066
Inter	Chemical	2430	0.274	0.250	0.001	0.877	0.162	1203
	Electric	2807	0.411	0.405	0.000	0.999	0.233	1113
	Machinery	2861	0.363	0.324	0.000	1.000	0.230	919
	Pharm	380	0.182	0.121	0.000	1.000	0.176	490
	all	8478	0.345	0.309	0.000	1.000	0.221	3725
ROA	Chemical	3375	0.044	0.041	-0.286	0.252	0.048	258
	Electric	3601	0.041	0.040	-0.316	0.254	0.063	319
	Machinery	3589	0.040	0.040	-0.287	0.239	0.057	191
	Pharm	704	0.064	0.067	-0.318	0.224	0.064	166
	all	11269	0.043	0.042	-0.318	0.254	0.057	934

Table 4 includes the industrial breakdown for the analyzed data-set. Whereas Chemical, Electric and Machinery industries have roughly the same number of observations, Pharmaceuticals contain far fewer companies. According to this industrial summary, Pharmaceutical patents have been the most influential, judged by the median value of “citations per patent” that equals on average to 2.63. On the other hand, this industry has a lowest ratio of “patents per yen of R&D”. According to Griliches (1998) (p. 308), these two opposite trends are connected. Owing to the highest industrial appropriability of patents (qualitative aspect), Pharm companies are not urged to convert their R&D expenses into numerous patented inventions (quantitative aspect).

Another reason for the relatively low value of “patents per yen of R&D” for Pharmaceuticals is the significance of R&D expenditures *per se* (see the highest industrial mean of 0.095 from the Table 4). As the Table 6 confirms, R&D intensity goes hand in hand with the capital intensity. While the former points to a higher share of R&D relative to sales, the latter is associated with a larger proportion of fixed assets per employee. Both metrics signify strategic corporate goals that prioritize long-term growth over immediate competitive yields. In a nutshell, owing to the fact that Pharmaceutical inventions are more time-consuming than, for example Machinery and Electronics (see Figure 1), one can clearly describe this industry as R&D- and capital-intensive. As we will see in the Section 4.2.1, for the Drugs companies, the R&D expenditure alone is responsible for their higher market valuation.

As the Table 4 demonstrates, the “Capital Intensity” values for the Chemical industry are also relatively high, logically linking it to the Pharmaceuticals. Considering the fact that both of these industries are responsible for the provision of medical services, one may conclude that Japanese tertiary-sector companies are relatively more capital-intensive than their manufacturing counterparts. This tendency corresponds to the picture presented by Brouwer and Kleinknecht (1997) (see Section 2). On the other hand, while these authors argued that manufacturing sector in Europe has a more pronounced R&D spending pattern, this does not appear to be the case for the Electric and Machinery industries in Japan, judged by their R&D intensity levels provided in the Table 4.

The fact that companies tend to either directly capitalize their R&D spending or concentrate on patents’ protection is also evident on the example of the Machinery industry (see the Table 4). Contrary to Pharmaceuticals, it has the highest rate of “patents per yen of R&D”, being at the same time ranked as the lowest in terms of “citations per patent”.

Fig. 5 Density plots for numeric variables

As the Figure 5 demonstrates, the values of different knowledge stocks such as Patents, Citations and R&D are noticeably right-skewed. This means that there is a tiny fraction of companies receiving large number of patents and citations, as well as having relatively high levels of R&D expenditure. Based on such pattern of distribution, I apply the log-transformation for these variables when running a regression, as the Equation (5) suggests.

The dependent variable, *Tobin's q* is also distributed unevenly, with most of the values concentrated towards the left end of the plot. However, rather than being simply right-skewed, it resembles the Poisson distribution. As a consequence, I apply the square root transformation to this variable. According to Wood (2017), this reflects the fact that the econometric model “is biased downwards on the response scale itself” (p. 331).

Table 5 Model Fits

Square-root-transformed Tobin's q	0.994
4th-root-transformed Tobin's q	0.9109
Log-transformed Tobin's q	-0.3927
Original Tobin's q	1.159

Table 5 shows that for the Equation (5), the dependent variable *Tobin's q* modified via the square-root transformation is a better fit than the hypothetical log-transformation or 4th-root transformation when comparing the fitted values to the original one.

Differently from the regression components discussed so far, the variable “Citations per Patent” that controls for innovative quality is distributed more evenly, thus pointing at the significant size effects. A relatively

even distribution of values also characterizes the group of the control variables that include: *Capital Intensity*, *Internationalization*, *Leverage*, *Process Innovation*, *R&D Intensity* and *ROA*. I incorporate the *Capital Intensity* variable based on the suggestion of Czarnitzki and Kraft (2010). Calculated “as fixed assets divided by the number of employees” (p. 1945), it reflects the capital requirements needed to successfully wage industrial competition.

Fig. 6 Visualized correlation matrix

Next, I move to the overview of the links existing between the variables. The Figure 6 demonstrates high correlation between Patents’ and Citations’ stocks. This, once again points at the fact that having on average one citation per patent is much more likely than accumulating multiple citations. In other words, influential patents constitute exception. The significantly positive level of correlation (0.83) between Patents’ and Citations’ stocks is confirmed by the Table 6.

Table 6 Correlation Matrix

	RDA	PSA	CSA	PatRD	CitPat	Q	RDInt	Process	Leverage	Capint	Inter	ROA
RDA	1											
PSA	0.28	1										
CSA	0.17	0.83	1									
PatRD	-0.15	0.52	0.51	1								
CitPat	-0.06	-0.02	0.22	0.06	1							
Q	0.21	0.05	0.06	-0.07	0.07	1						
RDInt	0.23	0.01	0	-0.17	0.04	0.12	1					
Process	-0.02	-0.09	-0.07	-0.09	0.03	0.06	0.11	1				
Leverage	0.33	0.2	0.18	-0.01	0.06	0.27	-0.26	-0.01	1			
Capint	-0.14	-0.11	-0.08	-0.02	0.08	0.09	0.15	0.29	-0.14	1		
Inter	0.15	0.04	0	-0.1	-0.11	0.09	0.07	-0.02	-0.07	-0.16	1	
ROA	-0.16	-0.1	-0.08	-0.03	-0.01	0.08	0.03	0.1	-0.32	0.07	0.1	1

It is known that the presence of correlation in linear regression models leads to the multicollinearity problem. Based, on this consideration, and also following the past literature, I run separate regressions for the R&D, Patents and Citation stocks. In addition, I have computed the variance inflation factor (VIF) scores for the main regression model that follows the Equation (5). The results of the test are presented in the Table 7. Since none of the VIF coefficients exceeds 2, we can confirm the absence of multicollinearity issues.

Table 7 Variance Inflation Factors

	GVIF	1/VIF
Year	1.238	1.006
Industry	1.561	1.077
R&D Assets	1.277	1.13
Patents/R&D	1.084	1.041
Citations/Patents	1.082	1.04
R&D Intensity	1.408	1.187
Process Innovation	1.08	1.039
Leverage	1.344	1.159
ROA	1.186	1.089
Capital Intensity	1.236	1.112

4.2 Regressions

In most of the existing studies on the market perception of innovation, authors use the “nonlinear least squares” (NLS) model. Notably, this technique requires an ex-ante knowledge of function(s) describing relationships between the chosen variables (Venables and Ripley 2002, Chapter 8). As a result, NLS are mostly used in natural sciences. On contrary, social and economic processes are less likely to follow blueprints. Due to this, the aforementioned authors have probably relied on self-starting models provided by statistical software without further explanation. Thus, neither their camouflaged choices, nor the obtained coefficients are well-interpretable.

In the face of this, current article starts with the linear “ordinary least squares” (OLS) model estimation that includes time- and industry-effects. As a next step, I additionally estimate nonlinear patterns via the application of the “generalized additive models” (GAMs) which I discuss in the Section 4.2.2.2. Following the advice from the related studies (see Section 2), I do not apply the fixed effects.

4.2.1 OLS models

My first estimations of the market valuation of knowledge assets are the so-called “horse race” regressions – the term coined by Hall, Jaffe, and Trajtenberg (2005). The name derives from the fact that every equation contains only one knowledge asset as an independent variable. Thereby it becomes possible to tentatively identify the innovation metric attracting most of the market’s attention. As the Table 8 demonstrates, overall premiums related to R&D stock are more significant than the ones associated with patents and citations. This finding corresponds to Cohen et al. (2002) who stressed the importance of R&D spillovers for the Japanese firms.

More specifically, R&D expenditures are valued for the Chemical and Pharmaceutical industries. In other words, similarly to Cohen et al. (2002), my results confirm that amid the general tendency of the Japanese stock market to regard R&D as the most important innovative proxy, it is also known for singling out Drugs industry in terms of its intangible output.

Table 8 Horse Race Regressions of Knowledge Assets

	<i>Dependent variable:</i>		
	Tobin's Q		
	(1)	(2)	(3)
Intercept	1.231*** (0.037)	0.808*** (0.038)	0.828*** (0.036)
<i>D (Electric)</i>	−0.029 (0.021)	0.345*** (0.043)	0.341*** (0.037)
<i>D (Machinery)</i>	−0.002 (0.025)	0.027 (0.041)	0.039 (0.037)
<i>D (Pharm)</i>	0.698*** (0.062)	0.459*** (0.130)	0.429*** (0.099)
<i>RE/D/Assets</i>	0.131*** (0.010)		
Electric	−0.016 (0.012)		
Machinery	0.001 (0.013)		
Pharm	0.100*** (0.027)		
<i>Patents/Assets</i>		−0.013*** (0.003)	
Electric		0.047*** (0.006)	
Machinery		0.010* (0.005)	
Pharm		0.013 (0.016)	
<i>Citations/Assets</i>			−0.011*** (0.003)
Electric			0.052*** (0.006)
Machinery			0.013** (0.005)
Pharm			0.010 (0.013)
Yearly dummies	Yes	Yes	Yes
Observations	7,813	7,586	7,535
R ²	0.154	0.095	0.099
Adjusted R ²	0.152	0.092	0.096
F Statistic	56.839*** (df = 25; 7787)	31.720*** (df = 25; 7560)	32.855*** (df = 25; 7509)

Note:

*p<0.1; **p<0.05; ***p<0.01

To make sure that the industrial effects are significant, I performed an *analysis of variance* (ANOVA, using the F-test) on the two models, which sole difference is the presence of interactions with industries in the second one. As seen from the Table 9, the industrial interactions present in the second model are indeed significant, judged by the small p-value. In addition, the lower value of the RSS indicates a better fit.

Table 9 Analysis of Variance (Industrial Effects)

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
7132	1174	NA	NA	NA	NA
7123	1156	9	18.29	12.52	6.157e-20

Table 10 Market Value as a Function of R&D, Patents, and Citations Stocks

	<i>Dependent variable:</i>	
	Tobin's Q	
	(1)	(2)
Intercept	−0.081* (0.042)	−0.023 (0.044)
<i>D (Electric)</i>	0.083*** (0.014)	0.062 (0.038)
<i>D (Machinery)</i>	0.018 (0.013)	−0.094*** (0.034)
<i>D (Pharm)</i>	0.378*** (0.023)	0.396*** (0.093)
<i>R&D/Assets</i>	0.117*** (0.010)	0.128*** (0.026)
Electric		−0.004 (0.026)
Machinery		0.056* (0.033)
Pharm		1.027*** (0.200)
<i>Patents/R&D</i>	−0.007*** (0.002)	−0.009*** (0.003)
Electric		0.020*** (0.005)
Machinery		−0.007 (0.005)
Pharm		0.016 (0.015)
<i>Citations/Patents</i>	0.012*** (0.003)	−0.005 (0.005)
Electric		0.056*** (0.008)
Machinery		0.028*** (0.008)
Pharm		−0.008 (0.011)
<i>R&D Intensity</i>	0.985*** (0.184)	0.660*** (0.193)
<i>Process Innovation</i>	0.006* (0.003)	0.005 (0.003)
<i>Leverage</i>	1.131*** (0.032)	1.091*** (0.033)
<i>ROA</i>	2.746*** (0.102)	2.728*** (0.101)
<i>Capital Intensity</i>	0.005*** (0.0004)	0.005*** (0.0004)
Yearly dummies	Yes	Yes
Observations	7,162	7,162
R ²	0.312	0.322
Adjusted R ²	0.309	0.319
F Statistic	111.359*** (df = 29; 7132)	89.185*** (df = 38; 7123)

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 10 includes several variables that control for the quality of knowledge stocks. Similarly to Hall, Jaffe, and Trajtenberg (2005), it demonstrates that belonging to the Pharmaceutical industry implies a premium in terms of market valuation of corporate knowledge assets. Furthermore, Drugs are notable for the marketability of their R&D expenditures (R&D stock weighted by Assets). In other words, alone the larger share of the R&D spending is a credible enough signal for investors to allocate premium value to the respective Pharm companies.

On the other hand, while being more R&D-intensive *per se*, Pharm industry does not strike as an efficient user of the R&D funds (the “Patents per R&D” variable). Only Electric industry stands out for its R&D spending that translates into more inventions valued by market. For it, one may say that the amount of the R&D-directed funds is contingent upon the quantity of patented innovations.

Indeed, qualitative aspect matters here as well. Electric industry, along with the Machinery, is peculiar for receiving market premium for more influential patents, judged by the number of citations. On the other hand, the fact that Pharm patents are well-cited compared to other industries is not reflected in their higher market valuation. The same holds true for the related Chemical industry. This, however, should not misleadingly imply the irrelevance of patents’ quality for these sectors. Rather, one may say that for them, the R&D-related effects “eclipse” the impact of patents’ quality.

4.2.2 Adding nonlinear effects

4.2.2.1 Horse race regressions Since most of the previous studies have stressed the nonlinear relationship between knowledge stocks (independent variables) and the market reaction (dependent variable), my next step is to test this assumption. First, since the horse race regression includes only one quantitative explanatory variable per model, it is easy to test the potential nonlinear responses against the baseline linear model by comparing the polynomial transformations.

Table 11 Polynomial Transformations of R&D Stock

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.077	0.01313	81.97	0
RDA	0.1877	0.01459	12.86	1.876e-37
RDA 2nd degree polynomial	0.08411	0.00984	8.547	1.526e-17
RDA 3rd degree polynomial	0.01616	0.005223	3.093	0.001989
RDA 4th degree polynomial	0.0007456	0.0007862	0.9484	0.3429

Table 11 demonstrates the significance of such changes until the 3rd degree (also called a “cubic” transformation) for the R&D Stock variable. Further transformations often lead to overfitting, whereby predicted values do not form any generalizable pattern, rendering the very definition of a “model” obsolete. In this case, moreover, the hypothetical 4th-degree transformation would not be justified due to the large corresponding p-value. Let us visually explore the rationale behind transforming the R&D stocks variable. Since we already determined the significance of industrial effects (see Table 9), it would make sense to plot the respective industrial breakdown.

Fig. 7 Modeling the relationship between R&D stocks and Tobin's q

Figure 7 contains several assumptions about the relationship between the two variables. These possibilities are captured by lines of different colors. The red curve representing “thin plate spline” is the trajectory that most closely reproduces the shape of the correlation. Not following any predefined model, it cannot be therefore expressed as a certain function. However, it provides the reliable benchmark for comparing all the subsequent assumptions to this realistic link between the two variables.

As we can see, the green curve representing the linear model does not accurately replicate the actual correlation, especially in the cases of Electric and Pharm industries. Therefore, our next guess is that further transformations would perform better. Indeed, the quadratic transformation (yellow curve) of the dependent variable RDA resembles the “thin plate spline” much clearer, thus being a more precise fit. The cubic transformation (blue curve) is even closer to the real picture. However, the difference with the quadratic shape is too little to justify the consumption of an additional degree of freedom. Based on these considerations, I chose to implement the 2nd degree (quadratic) modification of the R&D stock variable.

Following the similar logic of the marginal improvement of the fit, as well as of the visual proximity, I modified the remaining innovation-related variables in a following way. I added the cubic transformations for the Patent and Citations Stocks within the respective horse race regressions. The underlying tests that compare the polynomial transformations are included into the Table 12 for Patents’ Stock, and into the Table 13 – for Citations’ Stock.

Table 12 Polynomial Transformations of Patents Stock

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.877	0.1517	12.37	8.594e-35
PSA	0.3713	0.06677	5.561	2.777e-08
PSA 2nd degree polynomial	0.04883	0.01016	4.805	1.583e-06
PSA 3rd degree polynomial	0.002476	0.0006298	3.932	8.504e-05
PSA 4th degree polynomial	4.242e-05	1.341e-05	3.163	0.001568

Table 13 Polynomial Transformations of Citations Stock

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.454	0.1012	14.36	4.148e-46
CSA	0.1968	0.04992	3.942	8.156e-05
CSA 2nd degree polynomial	0.02412	0.008389	2.876	0.004046
CSA 3rd degree polynomial	0.001077	0.0005653	1.904	0.05689
CSA 4th degree polynomial	1.515e-05	1.292e-05	1.173	0.2408

In quantitative terms, the polynomial coefficients of the weighted Patents’ Stocks represent linear *orthogonal* combinations of variables PSA , $(PSA)^2$ and $(PSA)^3$. In other words, although each of these values are the modifications of a same variable, they are nevertheless uncorrelated, which is an essential condition for a linear regression model (James et al. 2013, Chapter 7). This, in turn, means that the new coefficients reflect the transformed variables and therefore cannot be interpreted as coefficients of a multiple linear equation.

4.2.2.2 Market value as a function of knowledge assets As the “horse race” regressions include only one explanatory variable (plus time- and industry-effects), it was enough to visualize the thin-plate-spline relationship between one of the knowledge assets and Tobin’s q . However, in case of the multiple regression model, like the one presented in the Table 10, adding nonlinear effects requires more sophisticated approach.

One of the techniques geared for such purpose is called “generalized additive models” (GAMs). Its main advantage is the ability to find a trade-off between closely capturing the actual relationship present in the data (i.e. being predictive), and not overfitting the model (i.e. not replicating the noise) (Pedersen et al.

2019). In this respect, fitted smooth functions provided by GAMs present a clear insight into the relationships between variables.

Similarly to the polynomial coefficients discussed above, the GAM values are not interpretable in a linear fashion. Instead, all the coefficients except for the industrial and yearly dummies as well as the intercept, represent the so-called “effective degrees of freedom” (EDF). EDFs show the complexity of the smooth (expressed as the number of basis functions), with the value “1” being an equivalent of a straight line, “2” – of a quadratic curve and so on (Pedersen et al. 2019). As a result, while in the linear regression each predictor has only one coefficient, the GAM contains as many coefficients for every independent variables as there are basis functions (plus the intercept). For this reason, in addition to reporting the coefficients, I include the visual representations of GAMs, which, together with diagnostics plots, is a legitimate way to report the results (Wood 2017, Chapter 7).

Table 14 contains the summary of the model. EDFs exceeding 2 combined with low p-values (beyond the 1% threshold) point at the fact that using GAM instead of a linear model is justified. This condition holds true for the most of the predictors. On the other hand, “Patents per R&D yen” and “Citations per Patent” show almost linear trends with EDF values less than 2. Hence, their relationships with Tobin’s q is better captured by the Table 10 coefficients.

As the Table 15 demonstrates, the proposed model explains 42.7 % of the variance in the Tobin’s q, which represents a robust fit and a considerable improvement from the 32.2 % value of the respective linear model (see Table 10). Furthermore, similarly to the linear models, accounting for industrial effects in GAMs improves the fit.

Table 14 GAM Summary

	edf	Ref.df	F	p-value
R&D/Assets	8.16	9.00	7.20	0.00
Patents/R&D	0.87	9.00	0.98	0.00
Citations per Patent	0.93	9.00	1.19	0.00
R&D Intensity	5.22	9.00	2.67	0.00
Process Innovation	7.28	9.00	5.15	0.00
Leverage	7.17	9.00	153.04	0.00
ROA	5.48	9.00	136.71	0.00
Capital Intensity	5.50	9.00	15.80	0.00

Table 15 GAMs Diagnostics

	GAM	Industrial GAM
AIC	6425	6263
BIC	6863	6963
Log Likelihood	-3149	-3030
Deviance	1010	977.2
Deviance explained	0.408	0.427
Dispersion	0.142	0.138
R²	0.403	0.419
GCV score	0.144	0.14
Num. obs.	7162	7162
Num. smooth terms	8	17

Figure 8 displays several diagnostics of the GAM fit. The Figure 8a, also known as a “Q-Q plot”, shows the distribution of residuals relative to a straight diagonal line that represents an ideal trajectory. With a few observations towards the higher end being outliers, the values mostly satisfy the normality demand. The Figure 8b provides the illustration of residuals following Gaussian, i.e. normal distribution. In the Figure

8c, the residuals generally appear to fall randomly around the zero x-axis value, thus not resembling any potentially overlooked function. Finally, the linear pattern present in the Figure 8d points at the close alignment of the predicted estimates of the dependent variable with its actual values.

Fig. 8 GAM Tests

In addition, since yearly and industrial effects are imputed in a regression as dummy variables, their characteristics within the GAM also reverberate those from the respective linear model. For example Figure 9b confirms that market allocates the largest premium to Pharmaceutical companies, while not rewarding other industries as consistently. Except for the Drugs companies, only the Electric firms reap the benefits of market attention thanks to their industrial class.

Since the presented linear regressions did not explicitly mention chronological developments, it is time to bring them up here. According to the Figure 9a, the lowest Tobin's q values corresponding to 2008 echo the effects of the Global Financial Crisis that broke out a year before. On the other hand, a few years preceding this event were marked by a pronounced growth trend reaching its peak in 2005. Finally, the recent years including 2016 also show a slow but steady recovery of the stock market.

Fig. 9 GAMs visualizations of Yearly and Industrial effects

So far I have discussed the partial resemblance of GAMs with the linear models. However, most importantly, applying GAM provides us with key insights on how to expand and/or fine-tune the initial linear predictions. First, I will discuss the general GAM and then will move to industry-wise effects. Here, I do not include the Pharmaceuticals due to the relatively small size of the sample, which may not be enough to formulate nonlinear assumptions.

Fig. 10 Visualizations of GAMs relationships

The pattern of premium allocation for R&D stocks has some notable nonlinear trends. According to the Figure 5, most of the observations have the R&D stock values ranging from 0 to 3. Within this scope, Figure 10 shows that market positively evaluates the surge in the R&D stock. On the industrial level, the similar pattern can be observed for the Electronics industry (Figure 12). Notably, this tendency is overlooked by the linear model that nevertheless recorded the positive and significant coefficients for the remaining industries (see Table 10).

The variable “Patents per R&D yen” that controls for of the R&D output follows a linear trend on a general level according to the Figure 10. However, this perception changes with the introduction of the industrial breakdown. According to the Figure 5, most of the “Patents per R&D yen” values range from 0 to 0.15. Furthermore, according to the Table 3, its median is only 0.006. The Figure 11b shows that within this data segment, R&D spending efficiency results in higher market evaluation of the Chemical companies – the aspect not captured by the linear model (see Table 10).

Furthermore, when disentangling “Citations per Patent” (that seem linear according to the Figure 10) on the industrial basis, one can also distinguish peculiar nonlinear trends. In the case of Electronics and Machinery, market positively evaluates the growth of citations per patent of up to 6 (Figure 12c, Figure 13c). In turn, the reaction appears more vague when patents become heavily cited. This most likely reflects reputational inertia, whereby patents are often cited due to hitherto accumulated credits rather than their present-day marketability (Duguet and MacGarvie 2005). At any rate, with the median level of “Citations per Patent” being roughly two (Figure 5), and with overwhelming majority of cases falling within the range between 0 and 6, one can safely conclude that influential patents within the manufacturing sector contribute to the surge in Tobin’s q. While the previously conducted linear regression provided a similar insight (see Table 10), GAM added important nuances.

Among other notable nonlinear trajectories is the Tobin’s q correlation with ROA. According to the Table 3,

ROA's minimal value is -0.318, while the median is 0.042. This means that the significant portion of Tokyo Stock Exchange companies have negative ROA values. As the Figure 10 clearly demonstrates, the market perception of ROA growth within this negative range is marked by distrust. On contrary, within the positive range, the market has been, as expected, rewarding the surge in ROA. Indeed, as seen from the Table 10, simple linear model does not capture this feature.

Fig. 11 Visualizations of GAMs relationships for the Chemicals

Fig. 12 Visualizations of GAMs relationships for the Electronics

Fig. 13 Visualizations of GAMs relationships for the Machinery

5 Conclusion

This paper provided insights on how corporate innovative activities among Japanese listed companies are reflected in their market value. “Market value” itself belongs to the world of perceptions rather than that of a productive economic activity. However, despite its rather arbitrary nature, it is a key reputational and therefore an economic asset for any listed company. While being volatile in the short-run, corporate stock provides valuable insights to present and potential investors, plus serves as an effective self-evaluating mechanism for corporate management. To better incorporate physical assets into the research question, this paper chose “Tobin’s q ” (calculated via dividing firm’s market capitalization by its replacement value) as a dependent variable.

In spite of the long legacy of the related scholarship, my research stands out at least for two reasons. First, it fills the gap of consistently under-representing Tokyo Stock Exchange, the third-largest stock market in the world. Second, it implements novel computational techniques for demonstrating nonlinear tendencies existing between the innovation-related predictors and the Tobin’s q .

Among the notable findings of this paper are the industrial patterns of innovative activities. Higher market evaluation appears to be inseparable from technological innovations within the Pharmaceutical industry. Notably, Drugs are the most R&D-intensive sector analyzed, which points at the relative importance that market places on its R&D stocks. It also allocates a relatively high premium for the surge in the R&D spending for the Chemicals that share similar characteristics.

Although the initial linear model did not detect any growth in Tobin’s q related to the efficiency of R&D spending (measured as “Patents per yen spent on R&D”) within the Chemical industry, the generalized additive model (GAM) discovered such a pattern. The same linear model, however, captured a small but significant market bonus allocated to the Electric Equipment companies. In contrast, R&D output does not appear to be significant for the Pharmaceutical and Machinery industries.

To add to this story, “Citations per Patent”, which is a measurement of patents’ quality, play a significant role for the market perception of technological innovations within Japan’s manufacturing sector. On the other hand, citations accruing to patents within Pharmaceutical and Chemical industries do not attract higher market premiums. In this respect, one observes the dividing line between putting more emphasis on R&D spending for capital-intensive industries like Chemicals and Drugs on one hand, and concentrating on the quality of the patented output for relatively more labor-intensive manufacturing sector on the other.

However, it is important to remind ourselves that Drugs’ and Chemicals’ patents are the most cited ones. Thus, being aware of it in advance, market rewards these industries by default, while additionally appreciating their R&D commitments. In a nutshell, market encourages the manufacturing sector, with its room for

patents' quality improvement, to concentrate on innovative *output*. As for the remaining industries, such as Drugs and Chemicals, market value reflects the innovative *input* efforts such as R&D investment.

Regarding the future research on this topic, I concur with Czarnitzki, Hall, and Oriani (2006) on the need to further explore nonlinear methods for better grasping the market assessment of innovative activities. While the previous studies have mostly relied on the “nonlinear least squares”, their computational methods have been obscure. In this regard, generalized additive models provide a viable alternative for further studies.

Finally, there exist other notable researches such as the one by Hall and Oriani (2006), who included corporate ownership structures into the their predictive model. Certainly, the innovation- and finance-related variables used in the current article do not fully capture the fluctuations in firms' market value. This provides a rich avenue for seeking further links between technological innovation and, among others, corporate governance, which I intend to explore in future.

References

- Acharya, Viral V., Ramin P. Baghai, and Krishnamurthy V. Subramanian. 2013. “Labor Laws and Innovation.” *The Journal of Law and Economics* 56 (4): 997–1037. doi:[10.1086/674106](https://doi.org/10.1086/674106).
- Brouwer, Erik, and Alfred Kleinknecht. 1997. “Measuring the Unmeasurable: A Country's Non-R&D Expenditure on Product and Service Innovation.” *Research Policy* 25 (8): 1235–1242.
- Cockburn, Iain, and Zvi Griliches. 1988. “Industry Effects and Appropriability Measures in the Stock Market's Valuation of R&D and Patents.” *American Economic Review* 78 (2): 419–423.
- Cohen, Wesley M., Akira Goto, Akiya Nagata, Richard R. Nelson, and John P. Walsh. 2002. “R&D Spillovers, Patents and the Incentives to Innovate in Japan and the United States.” *Research Policy* 31 (8-9): 1349–1367.
- Croissant, Yves, and Giovanni Millo. 2018. *Panel Data Econometrics with R: The plm Package*. Wiley.
- Czarnitzki, Dirk, Bronwyn H Hall, and Raffaele Oriani. 2006. “The Market Valuation of Knowledge Assets in US and European Firms.” *The Management of Intellectual Property*, 111–131.
- Czarnitzki, Dirk, and Kornelius Kraft. 2010. “On the Profitability of Innovative Assets.” *Applied Economics* 42 (15): 1941–1953. doi:[10.1080/00036840701749019](https://doi.org/10.1080/00036840701749019).
- Daróczi, Gergely. 2014. *Pander: An R Pandoc Writer*. <http://cran.r-project.org/package=pander>.
- Duguet, Emmanuel, and Megan MacGarvie. 2005. “How Well Do Patent Citations Measure Flows of Technology? Evidence from French Innovation Surveys.” *Economics of Innovation and New Technology* 14 (5): 375–393. doi:[10.1080/1043859042000307347](https://doi.org/10.1080/1043859042000307347).
- Goto, Akira, and Kazuyuki Motohashi. 2007. “Construction of a Japanese Patent Database and a First Look at Japanese Patenting Activities.” *Research Policy* 36 (9): 1431–1442. doi:[10.1016/j.respol.2007.06.005](https://doi.org/10.1016/j.respol.2007.06.005).
- Griliches, Zvi. 1984. *R&D, Patents, and Productivity*. Chicago: University of Chicago Press, National Bureau of Economic Research.
- Griliches, Zvi. 1998. “Patent Statistics as Economic Indicators: A Survey.” In *R&D and Productivity: The Econometric Evidence*, 287–343. Chicago: University of Chicago Press, National Bureau of Economic Research.
- Griliches, Zvi. 1981. “Market Value, R&D, and Patents.” *Economics Letters* 7 (2): 183–187. doi:[10.1016/0165-1765\(87\)90114-5](https://doi.org/10.1016/0165-1765(87)90114-5).
- Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg. 2005. “Market Value and Patent Citations.” *RAND Journal of Economics*, 16–38.

- Hall, Bronwyn H., and Raffaele Oriani. 2006. "Does the Market Value R&D Investment by European Firms? Evidence from a Panel of Manufacturing Firms in France, Germany, and Italy." *International Journal of Industrial Organization* 24 (5): 971–993. doi:[10.1016/j.ijindorg.2005.12.001](https://doi.org/10.1016/j.ijindorg.2005.12.001).
- Hall, Bronwyn H., Grid Thoma, and Salvatore Torrisi. 2007. "The Market Value of Patents and R&D: Evidence from European Firms." *NBER Working Papers* 1 (13426). doi:[10.3386/w13426](https://doi.org/10.3386/w13426).
- Haneda, Shoko, and Hiroyuki Odagiri. 1998. "Appropriation of Returns from Technological Assets and the Values of Patents and R&D in Japanese High-Tech Firms." *Economics of Innovation and New Technology* 7 (4): 303–321. doi:[10.1080/104385998000000038](https://doi.org/10.1080/104385998000000038).
- Hlavac, Marek. 2018. *Stargazer: Well-Formatted Regression and Summary Statistics Tables*. Bratislava, Slovakia: Central European Labour Studies Institute (CELSI). <https://CRAN.R-project.org/package=stargazer>.
- Ishii, Yasuyuki. 2010. "The Role of Patent, Citation and Objection Stocks in the Productivity Analysis of R&D Using Japanese Company Data." *International Journal of Innovation Management* 14 (05): 947–963. doi:[10.1142/S1363919610002945](https://doi.org/10.1142/S1363919610002945).
- James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An Introduction to Statistical Learning*. Vol. 103. Springer Texts in Statistics. New York, NY: Springer New York. doi:[10.1007/978-1-4614-7138-7](https://doi.org/10.1007/978-1-4614-7138-7).
- Kassambara, Alboukadel. 2019. *Ggcorrplot: Visualization of a Correlation Matrix Using 'Ggplot2'*. <https://CRAN.R-project.org/package=ggcorrplot>.
- Kleinknecht, Alfred, Kees van Montfort, and Erik Brouwer. 2002. "The Non-Trivial Choice Between Innovation Indicators." *Economics of Innovation and New Technology* 11 (2): 109–121.
- Mahlich, Jörg C. 2010. "Patents and Performance in the Japanese Pharmaceutical Industry: An Institution-Based View." *Asia Pacific Journal of Management* 27 (1): 99–113.
- Motohashi, Kazuyuki. 2008. "Licensing or Not Licensing? An Empirical Analysis of the Strategic Use of Patents by Japanese Firms." *Research Policy* 37 (9): 1548–1555.
- Nadiri, M. Ishaq, and Ingmar R. Prucha. 1996. "Estimation of the Depreciation Rate of Physical and R&d Capital in the U.S. Total Manufacturing Sector." *Economic Inquiry* 34 (1): 43–56. doi:[10.1111/j.1465-7295.1996.tb01363.x](https://doi.org/10.1111/j.1465-7295.1996.tb01363.x).
- Nakanishi, Yasuo, and Setsuo Yamada. 2007. "Market Value and Patent Quality in Japanese Manufacturing Firms." *MPRA, Munich Personal RePEc Archive*, 17.
- Nakanishi, Yasuo, and Setsuo Yamada. 2008. "Measuring the Rate of Obsolescence of Patents in Japanese Manufacturing Firms." *MPRA, Munich Personal RePEc Archive*, 11.
- Pakes, Ariel, and Zvi Griliches. 1984. "Patents and R&D at the Firm Level: A First Look." In *R&D, Patents, and Productivity*, 55–72. Chicago: University of Chicago Press, National Bureau of Economic Research.
- Pakes, Ariel, and Mark Schankerman. 1984. "The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of Return to Research Resources." In *R&D, Patents, and Productivity*, 73–88. Chicago: University of Chicago Press, National Bureau of Economic Research.
- Pakes, Ariel, Margaret Simpson, Kenneth Judd, and Edwin Mansfield. 1989. "Patent Renewal Data." *Brookings Papers on Economic Activity. Microeconomics* 1989: 331. doi:[10.2307/2534724](https://doi.org/10.2307/2534724).
- Pedersen, Eric J., David L. Miller, Gavin L. Simpson, and Noam Ross. 2019. "Hierarchical Generalized Additive Models in Ecology: An Introduction with MgcV." *PeerJ* 7 (May). PeerJ Inc.: e6876. doi:[10.7717/peerj.6876](https://doi.org/10.7717/peerj.6876).
- Piketty, Thomas. 2017. *Capital in the Twenty-First Century*. Harvard University Press.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

- Rufibach, Kaspar. 2009. “reporttools: R Functions to Generate LaTeX Tables of Descriptive Statistics.” *Journal of Statistical Software, Code Snippets* 31 (1). <http://www.jstatsoft.org/v31/c01/>.
- Sandner, Philipp. 2010. *The Valuation of Intangible Assets: An Exploration of Patent and Trademark Portfolios*. Innovation Und Entrepreneurship. Gabler Verlag. doi:[10.1007/978-3-8349-8393-0](https://doi.org/10.1007/978-3-8349-8393-0).
- Schankerman, Mark. 1998. “How Valuable Is Patent Protection? Estimates by Technology Field.” *The RAND Journal of Economics* 29 (1). [RAND Corporation, Wiley]: 77–107. doi:[10.2307/2555817](https://doi.org/10.2307/2555817).
- Scherer, F. M. 1965. “Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions.” *The American Economic Review* 55 (5): 1097–1125.
- Schmookler, Jacob. 2013. *Invention and Economic Growth*. 2nd printing 1976. Reprint 2014. Cambridge: Harvard University Press. doi:[10.4159/harvard.9780674432833](https://doi.org/10.4159/harvard.9780674432833).
- Trajtenberg, Manuel. 1990. “A Penny for Your Quotes: Patent Citations and the Value of Innovations.” *The RAND Journal of Economics* 21 (1): 172–187. doi:[10.2307/2555502](https://doi.org/10.2307/2555502).
- Venables, W. N., and B. D. Ripley. 2002. *Modern Applied Statistics with S*. 4th edition. New York: Springer.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. doi:[10.21105/joss.01686](https://doi.org/10.21105/joss.01686).
- Wood, Simon N. 2017. *Generalized Additive Models: An Introduction with R, Second Edition*. Chapman and Hall/CRC.
- Xie, Yihui. 2016. *Bookdown: Authoring Books and Technical Documents with R Markdown*. Boca Raton, Florida: Chapman; Hall/CRC. <https://github.com/rstudio/bookdown>.
- Xie, Yihui. 2020. *Knitr: A General-Purpose Package for Dynamic Report Generation in R*. <https://yihui.org/knitr/>.
- Zhu, Hao. 2019. *KableExtra: Construct Complex Table with ‘Kable’ and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>.