Hitotsubashi University
Institute of Innovation Research

Patent quality, cumulative innovation and market value : Evidence from Japanese firm level panel data

Sadao Nagaoka

Presented to the Conference on IT Innovation, Tokyo, December 2004.
Revised : March 2005.

IIR Working Paper WP#05-06

February 2005

Institute of Innovation Research
Hitotsubashi University
Tokyo, Japan
http://www.iir.hit-u.ac.jp

Patent quality, cumulative innovation and market value
: Evidence from Japanese firm level panel data[#]

Sadao Nagaoka[*]

Institute of Innovation Research, Hitotsubashi University

March 2005

Summary

We examine empirically how patent quality in terms of forward citation and science linkage affect the market value of a firm. We find that both indicators affect the market value of a firm significantly even if we extensively control the effects of the other major determinants of the market value, including R&D investment and current return on asset. In addition, the forward citation affects the market value more in cumulative innovation area such as in IT, consistent with a theoretical proposition that the value of having a dependent patent is larger in the industry where innovation is cumulative among firms.

Key words: patent quality; market value; forward citation; science linkage ; cumulative innovation

JEL classification: L10, O32

[#] I would like to thank for very helpful comments made by Shigeru Asaba, Kenneth Flamm and the other participants in the Conference on IT innovation at Hitotsubashi University.

[*] Professor, Institute of Innovation Research, Hitotsubashi University, 2-1 Naka Kunitachi Tokyo 186 Japan Fax 81-42-580-8410, Email nagaoka@iir.hit-u.ac.jp

1. Introduction

It is widely recognized that patent quality such as the forward citation (the average number of citations made by subsequent patents) and the science linkage (the average number of cited science literature) of a patent are significantly related to the value of a patent. Studies by Harhoff, Scherer and Vopel (2003) and Jaffe, Trajtenberg and Fogarty (2002) suggest that patent forward citation at an individual patent level is significantly correlated with the assessment by the inventor of the economic or technical importance of his/her patent. There is also a case study which suggests that the forward citation of a patent is correlated significantly with the social value of the invention incorporated into such patent(see Trajtenberg (1990)). Moreover, there are recent firm level studies mainly focusing on the US firms, which show that the patent quality of a firm is significantly correlated with its market value (Deng, Lev and Narin (1999), Hall, Jaffe and Trajtenberg (2000) and Hirschey and Richardson (2001, 2004)).

One of the interesting unaddressed questions is whether patent quality affects the market value of a firm differently, depending on the characteristics of innovation at industry level. A theoretical consideration suggests that the value of having a dependent patent would be larger in the industry where innovation is cumulative among firms, since a patent with the same number of dependent patents will control a larger number of patents indirectly in such industry. This suggests that the forward citation of the patents of a firm has a stronger effect on its market value in the industry where innovation is more cumulative. On the other hand, the science linkage of the patents of a firm may not have a similar amplified effect in the industry where science linkage is generally high, since high science linkage of an industry does not indicate high appropriation possibility for the patents with high science linkage. This paper analyzes whether we can find empirical evidence for these differential effects of patent quality on the market value of a firm, depending on the nature of innovation of the industry, based on the Japanese firm level data. Such inquiry would contribute to our understanding of the mechanism by which patent quality affects the value of a firm.

Our study can also be distinguished from existing studies on the linkage between patent quality and market value in the extent to which it controls the effects of the major determinants of the market value of a firm other than patent quality, including R&D investment, the number and the degree of diversification of the patent portfolio of a firm, the current accounting return on asset, and industry by year exogenous changes. The past studies, including Hall, Jaffe and Trajtenberg (2000) and Hirschey and Richardson (2001, 2004), control only some of these variables, and in particular do not control the effects of industry by year exogenous changes. Controlling these factors is crucial to avoid finding spurious correlation between patent quality measures and market value. For an example, improved technological opportunities of an industry would increase the market value of a firm belonging to the industry as well as the forward citation frequency and science linkage of the firm.

For this objective, we have matched the patent information from the

US patent database developed by Chi research[1] and the business and financial information of the Japanese firms listed in the Japanese stock exchanges available from NEEDS (Nikkei Economic Electronic Database System) and the stock price information of the World scope. The matching has enabled us to develop the balanced panel database covering 221 firms for three periods from 1988-1992, 1993-1997, and 1998-2002. The database covers a relatively small share (12 %) of the number of the listed firms in manufacturing, utilities (telecommunications, electricity and gas) and electronics-related wholesale trading. However, these firms account for 78 % of the sum of the R&D reported by the listed firms in these sectors, and 40 % of the total industrial R&D in Japan. Thus, we may say the sample covers substantially the R&D performing large firms in Japan (see Table 1 in section 2 for industry distribution of our sample firms).

The organization of the rest of the paper is the following. In section 2, we present analytical framework and two major hypotheses to be tested. In section 3 we describe the econometric specification for testing these hypotheses, the construction of the variables as well as estimation methods, and in section 4 we present the results. Section 5 concludes.

2. Patent quality, cumulative innovation and market value: analytical framework and hypotheses

We extend the following conventional specification of the market value of a firm, which is due to Griliches (1981) and extensively used thereafter (see

Hall (1999) for a comprehensive review of the literature[2]):

$$V = \theta(K + \lambda IK) \qquad (1)$$

where K is the value of the tangible capital stock of a firm, IK is the value of the intangible capital stock, λ is the relative shadow price of the intangible asset, and θ represents the divergence between the market value of a firm and the sum of its tangible and intangible capital stocks. θ is supposed to reflect the market power or the competitive advantage of a firm. We assume that the quality and the size of the patent portfolio of a firm affect this parameter θ. There are two determinants θ: how advanced the proprietary technology of a firm is relative to the existing technology (we call it *backward protection* provided by patent protection) and how long it will remain un-superceded by competing technologies (we call it *forward protection* provided by patent protection). Denoting the extent of the current technological advantage by ϕ (the parameter of backward protection) and the expected length of such advantage by T (the parameter of forward protection), we can have the following specification.

$$\theta = 1 + T\phi \qquad (2)$$

Here $T\phi$ represents the degree of divergence from a competitive situation and represents the quality of R&D investment.

Defining Tobin's q as the market value relative to the tangible capital stock ($q = V / K$), we have

$$q = \theta(1 + \lambda IK / K) \qquad (3)$$

Taking the logarithms of both sides, and assuming that both $T\phi$ and $\lambda IK / K$ are significantly less than 1, we have the following basic equation for our empirical analysis:

$$\ln q = \ln(V / K) \cong T\phi + \lambda IK / K \qquad (4)$$

Equation (4) suggests that the market value of a firm is high when its patent portfolio is of high quality (that is, high $T\phi$) for a given level of R&D investment due to either or both of more backward protection and more forward protection. We postulate that $T\phi$ increases with both the forward citation and the science linkage of the patents of a firm. Both measures of patent quality would indicate that such firm has significant inventions in its patent portfolio in terms of large ϕ, T or both. Note that a firm with high ϕ is likely to have large T, since a larger technological advantage gained today implies more time for being caught-up. In addition, high forward citation may directly indicate strong forward protection, since the future patents by other firms are more likely to be dependent on the patent portfolio of the firm. Thus, we have the following hypothesis for empirical testing.

Hypothesis 1

Both the number of the forward citations of the patent portfolio of a firm and its science linkage frequency enhance the market value of a firm for a given level of R&D investment, to the extent that they indicate both or either of stronger forward protection and stronger backward protection.

Let us turn to the interaction between patent quality and the nature of innovation at industry level. In the industry where innovation is cumulative among firms, forward citation is likely to be high for the industry as a whole. Indeed, IT industry (computers & peripherals, electronics, semiconductors and telecommunications) has high forward citations, as seen in Table 1. A patent in computer industry or in telecommunications industry is cited by the subsequent patents almost three times more than a patent in pharmaceutical industry. In the industry with strong cumulative nature of innovation, a patent with the same number of dependent patents can control a larger number of patents indirectly. As a result, we can expect that a highly cited patent tends to be more valued in such industry than in the rest. Assuming that forward citation reflects partly the dependency relationship, we can then postulate that a firm with a given amount of forward citation can enjoy stronger forward protection in the industry with high average forward citation and therefore higher market value (See appendix 1 for more details).

(Table1)

On the other hand, such complementary relationship between a firm level variable and an industry level variable may not hold for science linkage. Pharmaceutical and food industries have the highest science linkage indicators, with a patent in pharmaceutical industry having the number of science references 15 times more than a patent in automobile industry (see Table 1). High science linkage of an industry, however, does not enable the patent of a given science linkage to enjoy more forward protection in such industry. In fact, the quite opposite may be true, since more radical innovations supplanting the incumbent may be more likely to happen in such industry. In summary, we have the following proposition on how patent quality indicators affect market value of a firm, dependent on the forward citation and science linkage of the patents in each industry:

Hypothesis 2

The number of the forward citations of the patent portfolio of a firm has a larger effect on its market value in the industry with high average forward citation, due to the stronger forward protection in such industry, while the number of its science linkage may not have such interactive effect in the industry with high science linkage.

3. Specification, construction of basic variables and estimation methods

We use the following specifications, corresponding to equation (4), the one without the interaction terms between firm level and industry level patent variables and the one with such interaction terms:

$$\ln q_{i,t} = \beta_0 \ln(asset)_{i,t} + \beta_1 roa_{i,t} + \beta_2 (rda)_{i,t} + \beta_3 (\ln spats)_{i,t} + \beta_4 (\ln 1cits)_{i,t} + \beta_5 (\ln 1sci)_{i,t} + \beta_6 (adva)_{i,t} +$$

$$\sum \beta_{7.l} (debtasset)_{i,t-l} + \sum \beta_{8,l} (land)_{i,t-l} + \sum \beta_9 (inva)_{i,t-l} + (IndByPeriodDummies) + \alpha + \varepsilon_{i,t} \quad (5)$$

and

$$\ln q_{i,t} = \beta_0 \ln(asset)_{i,t} + \beta_1 roa_{i,t} + \beta_2 (rda)_{i,t} + \beta_3 (\ln spats)_{i,t} + \beta_{4,1} (\ln 1cits)_{i,t} + \beta_{4,2} (cit)_k \times (\ln 1cits)_{i,t}$$

$$+ \beta_{5,1} (\ln 1sci)_{i,t} + \beta_{5,1} (sci)_k \times (\ln 1sci)_{i,t} + \beta_6 (avda)_{i,t} +$$

$$\sum \beta_{7.l} (debtasset)_{i,t-l} + \sum \beta_{8,l} (land)_{i,t-l} + \sum \beta_9 (inva)_{i,t-l} + (IndByPeriodDummies) + \alpha + \varepsilon_{i,t} \quad (6)$$

Here, i indicates a firm and t indicates a period, and $\varepsilon_{i,t}$ is the error term.

The dependent variable $(lnq)_{i,t}$ is the logarithm of Tobin' q which is defined as the ratio of the market value relative to the book value of firm i in period t: (book value of the debt+ market capitalization of equity)/book value of the asset. We measure it for each firm at the end of the last year of three five-year periods (1998-1992, 1993-1997 and 1998-2002).

Let us move to independent variables. The logarithm of the value of tangible assets ($\ln(asset)_{i,t}$) controls the economy of scale. Its coefficient is positive if economy of scale exits. The variables $(rda)_{i,t}$ and $(adva)_{i,t}$ *are respectively* the R&D asset ratio and the advertisement asset ratio of a firm, and we expect positive coefficients for these variables. We use four patent based variables (*spats, cits*, *sci*, and *focus*), which are the flow variables based on the US patents granted in each period. Given the significant variation of the economic significance of an individual patent across industries (as shown in Table 1, the number of US patents per R&D is more than 10 times larger for a firm in instrument industry than for a firm in pharmaceutical industry), we use the normalized number of patents as a variable representing the size of patent portfolio in each period. In particular, we use the share of the number of the patents granted to a firm $(spats)_{i,t}$ in the entire Chi database. More precisely, it is the average of the shares of the US patents granted to each firm in 30 technology areas for a given period, with the weight being the number of the patents held by a firm in each technology area. We expect a positive coefficient for *spats*, since the appropriability of R&D investments are enhanced by more patents.

The variable $(cits)_{i,t}$ indicates the normalized number of forward citations made by the subsequent US patents up to the end of 2002 with respect to the patents granted to firm *i* in period *t*[3], after being adjusted for the biases due to the truncation of the length of citing period for recent periods. It is important to note that the fact that the forward citation frequency per patent for the patents granted in the first period (1988-1992) is significantly larger than that for the patents granted in the last period (1988-2002) does not mean that the quality of the patent deteriorated over time. The variable *cits* adjusts the number of citations per patent (*cit*) in period *t*, based on the ratio of the average number of citations per patent granted in the period from 1983 to 1988 and that granted in period *t* in each technology area. More specifically, it is given by the following formula:

$$cits_{i,t} = \sum_{j} (cit_{i,j,t})(\overline{cit_{j,1988}} / \overline{cit_{j,t}})(npat_{i,j} / npat_{i}) \quad (7)$$

Here, $cit_{i,j,t}$ is the average number of forward citations per patent of firm *i* in technology area *j* out of 30 technology areas in period *t* (See Appendix 1 for technology classification), $\overline{cit_{j,t}}$ is the average number of forward citations of the patents granted to all firms (including non-Japanese firms) in the CHI database which have at lease one patent in technology area *j*. On the other hand, *sci* indicates the average science linkage per patent of each firm, which is given by the average number of (backward) citations of science papers by the US patents granted to a firm (*science linkage* indicator constructed by Chi research)[4]. We expect positive coefficients for these variables: *cits* and *sci*.

In order to test hypothesis 2, we will implement the following two estimations. First we will estimate above equation (6), in which the effects of forward citation and science linkage of a firm can depend on industry variables of these measures. Here, cit_k is the average number of forward citations of industry k to which firm i belongs, and sci_k is the average science linkage of industry k in a logarithmic term, to which firm i belongs. We use the Japanese industry values for the period from 1993 to 1997. We expect that $\beta_{4,2}$ has a positive coefficient, given the complementary relationship between the forward citation at firm level and the forward citation at industry level (see hypothesis 2). Second, we will estimate above equation (5) focusing on IT industry (computers & peripherals, electronics, semiconductors and telecommunications) and pharmaceutical and food industry. We expect that the forward citation has a larger coefficient in the IT industry which has high average forward citation (see Table 1).

The variable *focus* is the HHI of the patent portfolio. It is defined as the sum of the squared shares of each of 30 technology areas in terms of the numbers of patents in, which is a broad classifications of technology (see appendix 1), in the total patents granted to a firm for a given period. Since we control the quality of the patents held by a firm, the *focus* variable is meant to measure the effect of focusing (or conversely diversification) of research or business in appropriating the research results.

Since all of these variables on the intangible assets (*rda*, *adva* and the patent based variables) refer to the investments in the current period, we use the return on asset $(roa)_{i,t}$ as an additional explanatory variable, which would reflect the effects of the investments on intangible assets made in the past. In addition we expect that *roa* could substantially control the effects of firm level missing variables (see the following discussions on estimation strategy), including firm fixed effects. We expect that the past investment performance in intangible assets are reflected in the current return on asset, which determines the market value of a firm, together with the current performance of the investments in intangible assets. Thus, it has a positive coefficient.

We introduce the following measures of the asset composition as the variables to control the divergence between the book value and the replacement value of the assets, which is very likely to vary across types of assets: the proportion of the investments in securities and affiliated companies (*inva* =financial investments, including investments in subsidiaries/total assets), the proportion of land (*land* =land/total assets), and debt asset ratio (*debtasset*). These balance sheet variables represent the information of the fiscal year just before the last year of three five-year periods, that is, the year which mainly ends on the march 31st of 1992, 1997, and 2002. We also introduce industry by time dummies to control the differentials of industry level demand growth and technological opportunities as well as the effects of macroeconomic changes over time. It is important to control them since R&D and market value move together in responding to the variations of these industry level missing variables and to macroeconomic shocks.

We also estimate an equation, using the return on asset (*roa*) as a dependent variable instead of Tobin' *q*, in order to check the robustness of our findings with respect to the two hypotheses. This specification evaluates the effects of patent quality on the accounting profitability of a firm. It is important to note that since R&D investment takes time for its effect to be realized while it is immediately expensed, we may have a negative coefficient of *rda* in *roa* equation. In addition, we also estimate the following market value equation using the unadjusted forward citation variable while allowing the coefficient of *cit* to vary over time. The logarithm of the Tobin's q is then specified as follows:

$$\ln q_{i,t} = \beta_0 \ln(asset)_{i,t} + \beta_1 roa_{i,t} + \beta_2 (rda)_{i,t} + \beta_3 (\ln spats)_{i,t} +$$

$$\beta_{4,1}(y1992cit)_{i,t} + \beta_{4,2}(y1997cit)_{i,t} + \beta_{4,3}(y2002cit)_{i,t} + \beta_5 (\ln 1sci)_{i,t} + \beta_6 (adva)_{i,t} +$$

$$\sum \beta_{7.l}(debtasset)_{i,t-l} + \sum \beta_{8,l}(land)_{i,t-l} + \sum \beta_9 (inva)_{i,t-l} + (IndByPeriodDummies) + \alpha + \varepsilon_{i,t} \quad (8)$$

Here $y1992cit$ takes the value of *cit* for the first period (1988-1992) and zero for the other two periods. $y1997cit$ and $y2002cit$ are similarly defined.

We use pool estimation as our preferred estimation, although we show the results of random-effects estimation as well as within (or fixed effects) estimation (Hall, Jaffe and Trajtenberg (2000) and Hirschey and Richardson (2001, 2004) also use pool estimations). There are two important sources of estimation biases for the above specifications. One is errors in variables. We expect that our indicators of the quality of patent are very noisy measures of true quality. Within (fixed-effects) estimation tends to amplify the impacts of errors in variables by reducing the variation of independent variables significantly, so that it causes the downward bias in estimation. Between estimation, on the other hand, reduces the effects of the errors in variables since it is based on the averages of the observations over time. The second source of estimation biases is a potential bias due to the correlation between a firm level missing variable and independent variables.

Within (firm fixed-effects) estimation can avoid the biases due to the presence of such correlation, while between estimation is subject to such bias. The pool estimation and random effects estimation, being the averages of within estimation and between estimation, balances these considerations. The pool estimation has a smaller weight on within estimation and is more efficient when firm fixed effects are negligible. The fact that our specifications extensively control firm level effects (for an example, the effect of managerial capability on profitability of a firm could be significantly captured by the return on asset *roa*) may provide a justification for us to use the pool estimation as our preferred estimation.

Table 1 presents the summary statistics of the key variables for the US patents granted during period from 1993 to 1997 (see Table Appendix_2 for entire summary statistics). Figure 1 shows that there exist positive correlations between two measures of patent quality and the market value for the Japanese firms. Tobin's q rises with forward citation and science linkage. The questions we address econometrically in next section are to analyze whether such correlation survives, controlling the other major determinants of the market value of firms, and whether their effects differ depending on whether innovations are cumulative or science-based.

(Figure 1)

4. Estimation results

Table 2 and 3 show the main estimation results, while Appendix Table A-3 presents the results of the supplementary regressions for robustness checks.

All estimations have the logarithm of Tobin's *q* as a dependent variable, except for estimations 8 and 9 which have *roa* as a dependent variable. All estimations use the aggregate balanced panel data, except for estimation 6 using the sub-sample of IT industry and for estimation 7 using that of pharmaceutical and food industry (which has the highest science linkage). We exclude the upper 5% and the lower 5% of the sample in terms of Tobin'*q* in estimation 2 in the appendix A-3 to see whether the results are not driven by a small number of outliers. We use the un-balanced panel data in estimation 3 in the appendix A-3.

First, let us take a look at the estimations based on the aggregate sample in Table 2. The first three estimations provide basic results, based on pool or random effect estimations, while estimation 4 provides the result of estimation without the accounting profitability (*roa)* as an explanatory variable. Pool and random effects estimations produced very similar coefficient estimates. We start with the patent quality variables. The forward citation (*ln1cits*) has a positive and significant coefficient (5% level) in estimation 1, even controlling the other major determinants of market value of a firm, including *roa*. The estimated coefficient of estimation 1 suggests that 10% increase of forward citation results in about 0.9 % improvement of the market value of a firm. Estimation 2 and 3 suggest that, once the forward citation interacted with the industry level forward citation (*citind*ln1cits*) is introduced, only the latter term has a positive coefficient (5% or 1% level). This implies that forward citation has a significantly larger effect in the industry with high forward citation. According to the coefficients in estimation 2, the marginal increase of forward citation has the effect amounting to 0.35 (=0.066*8.90-0.242) in telecommunications sector, while it has only 0.12 (=0.066*5.56-0.242) in automobile industry, thus three times difference between the industry with high forward citation and that with low forward citation. If we drop the *roa* variable, the significance of these variables increases, as shown in estimation 4.

Science linkage (*sci*) is also significant (5% level) in pool estimation 1, according to which 10% improvement of science linkage results in 0.8% improvement of the market value of a firm. On the other hand, the interaction term between the firm level value and the industry level value (*sciind*ln1sci*)) is not significant in both pool and random effect estimations (estimation 2 and 3 respectively), although it has a positive coefficient in both estimations. Thus, there is no strong evidence for the significance of the interaction term for science linkage.

Let us turn to the rest of the variables. They have the coefficients with the expected sigh, although the level of significance varies significantly across firms. As expected, the accounting profitability (*roa*) has a highly significant positive coefficient in all equations. R&D intensity (*rda*) and the patent share (*spats*) have positive coefficients, but their significance levels are low. The variable *rda* has a significance of 5% only in random effects estimation (estimation 3), while the patent share variable has only 10% significance in the same estimation. This low significance of the number of patents in market value equation seems to be partly due to multi-colinearity but may also indicate the presence of a strongly diminishing return from the number of patents in enhancing the appropriability of R&D[5].

Advertisement investment (*adva*) has a positive coefficient but not significant, although it becomes significant if we drop the accounting profitability *roa* (See estimation 4). The size of the estimated coefficient is much lower than for R&D intensity in estimations 1 to 3, which may be explained by the fact that advertisement investment may have more a near-term effect on the market value of a firm than R&D investment, so that its effect is more caught by the *roa* variable. Finally, focus variable (*lnfocus*) does not have a significant coefficient in any market value equation, although it has a positive coefficient and the size of coefficient becomes larger when *roa* variable is dropped (see estimation 4). The focus of R&D has a favorable effect on the accounting profitability, controlling firm size among others (estimations 8 and 9).

Let us move to the results of within estimation (see estimation 5 in Table 2). Within estimation removes the estimation bias due to any remaining firm fixed effect which is not controlled by *roa* variable, while it tends to worsen the errors in variables problem. As for the patent quality variables, the coefficients of the forward citation variables (*ln1cits* and *citind*ln1cits*) are not much affected, although they become insignificant (compare estimations 2 and 5). The coefficients of the science linkage variables, on the other hand, declined further. On the other hand, *roa* remains highly significant, and *rda* and *spats* variables increased their significance.

We then turn to the estimation results in Table 3. Estimations 6 focus on IT industry and estimation 7 on pharmaceutical and food industries. The estimated coefficient of forward citation for IT industry, which is significant at 10% level. The coefficient of *cits* is 0.235 for IT industry (estimation 6 in Table 3), while it is 0.093 for the aggregate sample (estimation 1 in Table 2). That is, a firm with a highly cited patent can enhance its market value significantly more in IT industry. That is, forward citation matters more in cumulative technology area, consistent with our hypothesis.

(Table 3)

On the other hand, science linkage is no more significant in affecting the market value of a firm in pharmaceutical and food industry. The size of the coefficient for *ln1sci* estimated for the sub-sample of pharmaceutical and food industry (although not significant) is similar to that for the aggregate sample (compare estimation 1 and 7). In addition, the share of the patents granted (*spats*) has a positive coefficient (significant at 10%) in the pharmaceutical and food industry (see estimation 7), unlike the finding to the contrary based on the aggregate sample (see estimation 1). R&D itself (*rda*) is significant for this subset of industry. These findings suggest that a patent has a significantly higher value in this industry.

Estimations 8 to 9 use the accounting rate of return (*roa*) as the dependent variable, using the aggregate sample. The estimations results for forward citation variables are significantly consistent with those found for market value of a firm. Since the coefficient of the interaction term (*citind**ln1cits*) is positive and significant, forward citation has a larger positive effect on accounting profitability of a firm in the industry with high average forward citation. Science linkage is found to be not significant for accounting profitability (see estimation 8). The negative coefficient of *rda* may not be surprising, since R&D investment takes time for its effect to be realized while it is immediately expensed.

Let us finally turn to the robustness checks, shown in Table Appendix-3. Estimation A-1 uses the un-normalized forward citation variable of the patents of a firm, while the equation allows the coefficient to vary over time. This raw forward citation variable (*cit*) has a positive and highly significant coefficient (1% or 5% level) for both as the most recent period (1998-2002) and for the second period (1993-1997). The science linkage variable has a positive and significant coefficient at 5% level. Thus, these results support hypothesis 1, even if we use the raw citation variable. Estimation A-2 in Table Appendix-3 shows that excluding the upper 5% and the lower 5% of the sample in terms of Tobin'*q* makes the forward citation variables (*lncit* and *citindln1cits*) more significant (5% and 1% level significance respectively). Estimation A-3 uses the sample of an unbalanced panel. This sample is larger, since it covers the firms which appeared only in the second or third period or which exited in one of these two periods, so that it may have sample selection problem in pool or random effects estimations. As shown in the table, focusing on the unbalanced panel does not significantly affect the results either, although the significance of the forward citation variables is weaker.

5. Conclusions and discussions

This paper has examined whether the patent quality of a firm significantly affects the market value of a firm and whether the frequency of forward citation (by subsequent patents) matters more in the industry where innovation is cumulative. We have two major findings. First, we have found that both forward citation and science linkage of the patents granted to a firm significantly increases its market value of the Japanese firms, even if we extensively control the effects of the other major determinants of market value. They include R&D and advertisement investments, the current return on asset as a variable representing firm level missing variables and industry level exogenous changes. Thus, these two patent quality variables are the significant indicators of the R&D quality of a firm.

Second, we have found that the forward citation of the patents of a firm affects its market value more in the industry where innovation is cumulative among firms. This finding is consistent with a theoretical proposition that the value of having a dependent patent is larger in the industry where innovation is cumulative among firms. On the other hand, science linkage of a patent does not have a similar amplified effect in the industry where science linkage is generally high. Such difference between forward citation and science linkage may not be surprising, given that a patent with high science linkage does not necessarily imply that such patent enables more appropriation of the return from the follow-up innovations, unlike a patent with high forward citation. Although both are the measures of cumulative innovation in a broad sense, science linkage mainly indicates the linkage between industry research and academic research, while forward citation in patent terms indicates the linkage within industrial research.

Appendix 1. A simple model of the value of a patent in the industry of cumulative innovation with tree structure

Let us consider the following three-period model of innovation competition. Initially the patent of firm *A* has value ϕ, which represents the degree of the divergence from the competitive situation (see equation (2) in the main text). We ignore discounting. A competing invention occurs in the end of the first two periods. Such invention may be dependent directly or indirectly on the patent of firm A or it may be independent from the patent of firm *A*. We assume that the emergence of the dependent patent changes the value of the patent of firm A by a factor ρ (>0) and the emergence of an independent patent makes it totally obsolete. In addition, the patent of firm *A* loses its exclusive right in the end of the third period due to the statutory limitation on the patent term.

We assume that the probability that the patent of firm *A* becomes obsolete by an competing invention in the end of the first period declines with the number of the forward citations of its patent (cit_A):

$$\Pr_1 = 1 - \delta cit_A . \qquad \text{(a.1)}$$

This reflects the fact that a patent with more forward citations is more likely to control the subsequent patents. When the invention in the first period is dependent on the patent of firm *A*, the firm *A* can sustain its market power after the first period, and the value of the patent changes from ϕ to $\rho\phi$.

In the second period, the patent of firm *A* has no value if the independent patent is granted in the first period. Unless, it sustains its value in the second period but may become obsolete due to the emergence of a new independent invention in the end of the second period. For simplicity, we assume that the second period patent is not directly dependent on the patent of firm A. However, it can be indirectly dependent through its dependence on the first period patent which is dependent on the patent of firm *A*. In such case, the patent of firm *A* still sustains its value in the third period. We assume that the probability that the second period invention is independent of the first period patent declines with the average number of the forward citations of the industry patents:

$$\Pr_2 = 1 - \delta cit_{ind} \qquad \text{(a.2)}$$

Given these assumptions, the value of the patent of firm *A* expected at the beginning of the first period is given by the sum of the expected values of the following three contingencies: the patent of firm A has its value for one period, two periods or three periods:

$$y = \phi \Pr_1 + \phi(1+\rho)(1-\Pr_1)\Pr_2 + \phi(1+\rho+\rho^2)(1-\Pr_1)(1-\Pr_2)$$

$$= \phi\{1 + \rho\delta cit_A + (\rho\delta)^2 cit_A cit_{ind}\} \qquad \text{(a.3)}$$

Thus, the derivate of the patent value with respect to its forward citation is given by

$$\partial y / \partial(cit_A) = \phi\{\rho\delta + (\rho\delta)^2 cit_{ind}\} \qquad \text{(a.4)}$$

That is, the effect of forward citation on the value of the patent increases with the average forward citations of the industry patents.

Reference

Branstetter L. 2003, “Is Academic Science Driving a Surge in Industrial Innovation? Evidence from Patent Citations,” Mimeo

Deng Z. B. Lev, and F. Narin, 1999, “Science and technology as predictors of stock performance, “*Financial analysts Journal*, Vo. 55, no. 3, pp.20-32

Griliches Zvi, 1981, "Market value, R&D and patents," *Economic letters*, 7,183-187

Hall Bronwyn H. A. Jaffe and M. Trajtenberg, 2000, “Market Value and Patent Citations: A First Look,” *NBER Working Paper* 7741

Hall, B. H, 1999, “Innovation and Market Value,” in Barrell, Ray, Geoffrey Mason, and Mary O'Mahoney (eds.), *Productivity, Innovation and Economic Performance*, Cambridge: Cambridge University Press.

Harhoff D., F. M. Scherer and K. Vopel, 2003, “Citations, family size, opposition and the value of patent rights,” Research Policy, Volume 32, Issue 8, 1343-1363

Hirschey Mark and Vernon J. Richardson, 2004, “Are scientific indicators of patent quality useful to investors?” *Journal of Empirical Finance*, 11, 91-107

Hirschey Mark and Vernon J. Richardson, 2001, “Valuation effects of patent quality: A comparison for Japanese and U. S. firms”, *Pacific-Basin Finance Journal*, 9, 65-82

Narin F. 2000, “Tech-Line Background paper”, in *From Knowledge Management to Strategic Competence, Series on Technology Management--Vol. 3*, 155-195, Joe Tidd, ed., Imperial College Press

Jaffe, A. and M. Trajtenberg and M.S. Fogarty, 2002, “The Meaning of Patent Citations: Report on the NBER/Case-Western Reserve Survey of Patentees,” in *Patents, Citations, and Innovations: A Window on the*

Knowledge Economy (edited by Jaffe, A. and M. Trajtenberg), MIT Press.

Nagaoka Sadao, 2004, "R&D and market value: appropriability vs. preemption," Institute of Innovation Research, Hitotsubashi University, *presented at the meeting at the 10th meeting of International Schumpeter Society*, (http://www.schumpeter2004.uni-bocconi.it/papers.php)

Meyer, M., 2000, "Does Science Push Technology? Patents Citing Scientific Literature," *Research Policy*, 29, pp. 409-434.

Scherer, F.M. and D. Harhoff, 2000, Technology policy for a world of skew-distributed outcomes*, Research Policy,* 29: 559-566

Trajtenberg M, 1990, "A Penny for Your Quotes: Patent Citations and the Value of Innovation." *Rand Journal of Economics,* pp.172-187

[1] See Narin (2000) for an explanation of the database. It covers the major patentees of the US patents.

[2] See Nagaoka (2004) for a non-conventional approach, emphasizing the non-rival nature of technology within a firm.

[3] Our data consist of the patent data in 30 technology areas. We calculated the average forward citation for each firm for a given period based on the number of patents granted in each technology area.

[4] See Branstetter (2004) for an analysis of the causes of the recent rise of science linkage and Meyer (2000) for various meanings of the science paper references.

[5] This low significance of the number of patents is not consistent with the finding of Hall et al. (2000) for the US firms, although it is consistent with Hirschey et al. (2001). The difference between our study and Hall et al. (2000) may be accounted for by our better control of exogenous and firm level factors in our study, and by the difference of the firm coverage.

Table 1 The average characteristics US patents granted to Japanese firms for 93-97 by sectors

No.	Industry	No. of firms	q	*annual rd (Million Yen)*	number of patents	npats/rd (semiconductor=100)	cit per patent	sci per patent	focus	IT (Cumulative technology)	Hihg science linkage
1	AUTO	21	1.26	5,066	201	54	5.56	0.13	3,068		
2	CHEM	47	1.34	9,281	198	29	4.32	0.53	2,587		
3	COMP	8	1.56	90,015	1,357	20	9.52	0.37	2,924	●	
4	CONS	5	1.19	3,671	83	31	4.10	0.15	3,705		
5	ELEC	10	1.59	33,476	707	29	5.59	0.42	2,246		
6	ELTR	27	1.55	26,163	792	41	6.64	0.26	2,749	●	
7	ENGY	4	1.23	5,949	83	19	4.18	0.95	1,349		
8	FOOD	9	1.38	5,255	68	18	2.82	2.53	3,309		●
9	FORP	3	1.51	2,303	81	48	6.40	0.08	3,733		
10	HLTH	3	1.90	17,488	127	10	5.03	1.03	3,230		
11	INST	15	1.62	22,037	1,070	66	5.46	0.43	2,906		
12	MACH	29	1.44	7,759	114	20	4.43	0.18	2,818		
13	MATL	5	1.82	6,814	134	27	5.59	0.37	2,410		
14	METL	14	1.41	8,397	260	42	4.38	0.53	1,532		
15	PHAR	11	1.70	28,208	138	7	2.91	2.10	3,664		●
16	SEMI	4	2.15	2,303	170	100	5.94	0.75	2,342	●	
17	TCOM	2	1.39	82,405	283	5	8.90	1.03	4,064	●	
18	TEXT	5	1.20	5,587	70	17	3.71	1.25	1,647		
	Total	222									

Note. Citation is up to the end of year 2002. *q is* for the end of 1997 *and* R&D is for 1996 Fiscal year (mostly March 31st in 1997).

Table 2 Estimation results (Patent quality and Tobin's q)

	Estimation 1 (3 periods)			Estimation 2 (3 periods)			Estimation 3 (3 periods)			Estimation 4 (3 periods)			Estimation 5 (3 periods)		
	Pool estimate			Pool estimate			Random-effects			Random-effects			Within estimation		
lnq	Coef.	Robust Std. Err.		Coef.	Std. Err.		Coef.	Std. Err.		Coef.	Std. Err.		Coef.	Std. Err.	
lnasset	0.025	0.014	*	0.022	0.014		0.025	0.016		0.048	0.018	***	0.178	0.067	***
roa	3.732	0.678	***	3.646	0.678	***	3.220	0.337	***				2.411	0.415	***
rda	1.028	0.659		1.230	0.663	*	1.254	0.633	**	0.681	0.696		2.304	1.282	*
spats	0.015	0.012		0.014	0.012		0.021	0.013	*	0.009	0.014		0.053	0.023	**
ln1cits	0.093	0.037	**	−0.242	0.129	*	−0.279	0.131	**	−0.361	0.143	**	−0.349	0.222	
citindln1cits				0.066	0.026	**	0.069	0.025	***	0.087	0.027	***	0.071	0.045	
ln1sci	0.079	0.037	**	0.013	0.059		0.002	0.070		0.029	0.075		−0.004	0.099	
sciindln1sci				0.047	0.039		0.040	0.047		0.016	0.051		0.011	0.066	
lnfocus	0.014	0.026		0.003	0.026		0.004	0.028		0.041	0.031		0.003	0.052	
adva	0.282	0.793		0.020	0.799		0.287	1.144		2.548	1.249	**	−0.536	2.856	
industry by time dummies	yes			yes			yes			yes			yes		
structure of assets	yes			yes			yes			yes			yes		
	Number of obs = 666 R-squared = 0.4709 Root MSE = .24204			Number of obs = 666 R-squared = 0.4790 Root MSE = .24059			Number of obs = 666 Number of groups = 222 R-sq: within = 0.3709 between = 0.5500 overall = 0.4757 sigma_u = .10737546 sigma_e =.20172801			Number of obs = 666 Number of groups = 222 R-sq: within = 0.3336 between = 0.4012 overall = 0.3735 sigma_u = .1341138 sigma_e = .21004082			Number of obs = 666 Number of groups = 222 R-sq: within = 0.4066 between = 0.1696 overall = 0.1861 sigma_u = .35405676 sigma_e = .20172801		

***: significant at 1%, **: significant at 5%, *: significant at 10 %

Table 3 Estimation results (Patent quality and Tobin's q or ROA)

	Estimation 6 (3 periods)			Estimation 7 (3 periods)			Estimation 8 (4 periods)			Estimation 9 (4 periods)		
	Pool estimate for lnq in IT			Pool estimate for lnq in phar. & food			Pool-effects for roa			Pool-effects for roa		
lnq /roa	Coef.	Robust Std. Err.		Coef.	Std. Err.		Coef.	Std. Err.		Coef.	Std. Err.	
lnasset	0.031	0.057		0.057	0.058		0.005	0.002	***	0.005	0.002	***
roa	3.371	0.797	***	4.075	1.537	**						
rda	−0.733	1.208		−0.307	2.403		−0.135	0.100		−0.129	0.098	
spats	0.038	0.036		0.093	0.052	*	−0.003	0.001	**	−0.003	0.001	**
ln1cits	0.235	0.121	*	−0.117	0.151		0.008	0.005	*	−0.038	0.020	*
citindln1cits										0.009	0.004	**
ln1sci	−0.272	0.175		0.062	0.092		0.008	0.006		0.017	0.009	*
sciindln1sci										−0.010	0.006	*
lnfocus	−0.009	0.068		−0.002	0.117		0.011	0.003	***	0.010	0.003	***
adva	−0.673	3.464		2.842	2.004		0.554	0.129	***	0.596	0.131	***
industry by time dummies	yes			yes			yes			yes		
structure of assets	yes			yes			yes			yes		
	Number of obs = 123 R-squared = 0.5399 Root MSE = .27009			Number of obs = 60 R-squared = 0.5990 Root MSE = .22934			Number of obs = 832 R-squared = 0.4399 Root MSE = .03261			Number of obs = 832 R-squared = 0.4536 Root MSE = .03225		

***: significant at 1%, **: significant at 5%, *: significant at 10 %

Table Appenidix_1

Technology classification

1	Agriculture
2	Oil & Gas, Mining
3	Power Generation & Distribution
4	Food & Tobacco
5	Textiles & Apparel
6	Wood & Paper
7	Chemicals
8	Pharmaceuticals
9	Biotechnology
10	Medical Equipment
11	Medical Electronics
12	Plastics, Polymers, & Rubber
13	Glass, Clay, & Cement
14	Primary Metals
15	Fabricated Metals
16	Industrial Machinery & Tools
17	Industrial Process Equipment
18	Office Equipment & Cameras
19	Heating, Ventilation, Refrigeration
20	Misc. Machinery
21	Computers & Peripherals
22	Telecommunications
23	Semiconductors & Electronics
24	Measurement & Control Equipment
25	Electrical Appliances & Components
26	Motor Vehicles & Parts
27	Aerospace & Parts
28	Other Transport
29	Misc. Manufacturing
30	Other

Industry classification

	Industry code	Industry
1	COMP	Computers
2	ELTR	Electronics
3	SEMI	Semiconductors
4	TCOM	Telecommunications
5	ELEC	Electrical
6	INST	Instrument. & Optical
7	MACH	Machinery
8	AUTO	Automotive
9	CHEM	Chemicals
10	CONS	Consumer Products
11	ENGY	Energy
12	FOOD	Food, Bev. & Tobacco
13	FORP	Forest & Paper Prods.
14	HLTH	Health Care
15	MATL	Materials
16	METL	Metals
17	PHAR	Pharmaceuticals
18	TEXT	Textiles

Note. No firm in aerospace,biotechnology,engrng., oil field svcs. industries exists in our sample.

Table Appendix_2 Summary Statistics

						Mean			Correlation												
Variable	Obs	Mean	Std. Dev.	Min	Max	1992	1997	2002	q	lnasset	roa	rda	spats	cit	cits	sci	focus	adva	debtasset	land	inva
q	666	1.426	0.547	0.546	5.044	1.472	1.461	1.346	1												
lnasset	666	12.45	1.14	9.76	16.23	12.40	12.45	12.50	0.10	1											
roa	666	0.039	0.038	−0.151	0.284	0.044	0.047	0.026	0.47	0.00	1										
rda	666	0.025	0.023	0.000	0.137	0.024	0.024	0.027	0.20	0.22	0.15	1									
spats	666	0.004	0.007	0.000	0.061	0.004	0.004	0.004	0.08	0.38	−0.02	0.38	1								
cit	666	4.667	3.683	0.217	28.036	7.755	5.043	1.202	0.17	0.01	0.26	−0.08	0.10	1							
cits	666	7.774	3.252	1.712	28.669	8.270	7.497	7.555	0.19	0.07	0.12	−0.05	0.15	0.67	1						
sci	666	0.537	0.970	0.000	12.102	0.313	0.575	0.723	0.10	0.18	0.21	0.31	−0.07	−0.22	−0.20	1					
focus	666	2662	1520	638	10000	2661	2728	2596	0.08	−0.43	0.16	0.02	0.06	0.05	0.05	−0.03	1				
adva	666	0.008	0.012	0.000	0.098	0.009	0.008	0.006	0.15	0.05	0.29	0.24	0.13	0.12	0.09	0.25	0.13	1			
debtasset	666	0.542	0.183	0.083	1.082	0.575	0.543	0.508	−0.31	0.22	−0.32	−0.17	0.03	0.03	−0.09	−0.09	−0.22	−0.12	1		
land	666	0.059	0.043	0.003	0.327	0.049	0.058	0.069	−0.19	−0.15	−0.04	−0.11	−0.10	−0.18	−0.11	0.00	0.11	0.08	0.06	1	
inva	666	0.214	0.114	0.019	0.925	0.167	0.190	0.284	−0.10	0.20	−0.20	0.04	0.16	−0.31	−0.02	0.07	−0.03	0.05	−0.17	−0.07	1

Table Appendix-3 Supplementary estimations

	Estimation A-1 (3 periods)				Estimation A-2			Estimation A-3		
	Pool-effects for roa				Pool estimate for trimmed sample			Pool estimate for unbalanced panel		
	Coef.	Std. Err.		lnq /roa	Coef.	Robust Std. Err.		Coef.	Std. Err.	
lnasset	0.024	0.014	*	lnasset	−0.003	0.009		0.043	0.011	***
roa	3.823	0.674	***	roa	3.164	0.488	***	3.831	0.598	***
rda	1.023	0.653		rda	0.937	0.534	*	1.642	0.677	**
lnspats	0.013	0.012		spats	1.066	1.365		−0.998	1.708	
y1992ln1cit	−0.034	0.056		ln1cits	−0.247	0.109	**	−0.156	0.124	
y1997ln1cit	0.110	0.043	**	citindln1cits	0.063	0.022	***	0.047	0.025	*
y2002ln1cit	0.356	0.116	***	ln1sci	0.044	0.055		0.029	0.054	
ln1sci	0.079	0.038	**	sciindln1sci	0.040	0.035		0.020	0.037	
lnfocus	0.012	0.026		lnfocus	0.006	0.020		0.000	0.023	
adva	0.302	0.772		adva	0.879	0.722		0.752	0.634	
industry by time dummies	yes			industry by time dummies	yes			yes		
structure of assets	yes			structure of assets	yes			yes		
	Number of obs = 666				Number of obs = 600			Number of obs = 797		
	R-squared = 0.4803				R-squared = 0.4298			R-squared = 0.4620		
	Root MSE = .24028				Root MSE = .17813			Root MSE = .24652		

***: significant at 1%, **: significant at 5%, *: significant at 10 %