CIRJE-F-66

The Role of R&D Diversification in Exploiting Knowledge Spillover:

Evidence from Major Japanese Chemical and Pharmaceutical Companies

Tsuyoshi Nakamura
University of Tokyo

December 1999

Preliminary

The Role of R&D Diversification in Exploiting Knowledge Spillover: Evidence from Major Japanese Chemical and Pharmaceutical Companies

First version 99/04/09
This version 99/12/24

Tsuyoshi Nakamura *

JEL classification: O31, O33, L16

*Center for International Research on the Japanese Economy, Faculty of Economics, University of Tokyo.
E-mail address: nkmr@e.u-tokyo.ac.jp

Abstract

The role of R&D diversification has been explained by the fact that there exist knowledge spillovers among research programs within a firm. However, the benefit from diversification is not ensured by internal spillovers only. Knowledge can spill over not only within a firm, but also across firms. If it costs a firm to exploit external knowledge spillovers no more than to internal knowledge spillovers, the firm need not keep wide variety of research menu inside.

The aim of this paper is to clarify the role of R&D diversification with taking care of knowledge spillovers both within a firm and across firms. The main hypothesis is that the impact of diversification is exaggerated when a firm confronts barrier to exploit knowledge stock outside. Based on the patent data of about 30 major companies in Japanese chemical and pharmaceutical industries, this paper finds the evidence in favor of this hypothesis. Robustness of the results is held for several indices of R&D diversification. The impact of R&D diversification depends on both quantitative and qualitative measures of accessibility to external knowledge.

1 Introduction

Non-exclusiveness of utilizing knowledge is important for a firm to diversify its R&D activity. The literature on corporate diversification, like Teece (1980) and Panzer and Willig (1981), has emphasized that the existence of shareable inputs is important for diversification to be beneficial. Since knowledge created in a sector of a firm is shareable in other parts of the firm, diversification of R&D can realize what Panzer and Willig called the 'economies of scope' due to knowledge spillover within a firm.

The existence of knowledge spillover within a firm is examined by Henderson and Cockburn (1996). They analyzed the in-depth data on research activities of ten major pharmaceutical firms, which provides the information on patents and R&D expenditure at research program level. Their empirical results showed that the productivity of a research program is raised by R&D activity of other related programs in the firm. This implies that once knowledge is created in one research program of a firm, it contributes to other programs in the firm through spillover.

However, spillover within a firm does not ensure diversification of R&D activities to be beneficial by itself. Its relative size to spillover from outside is another key factor. Jovanovic (1993) described the condition of beneficial R&D diversification by a simple model. In his model, Jovanovic pointed out that spillovers of within a firm should exceed those across firms for productivity enhancing diversification. If the knowledge appropriated by other firms is available at least as low cost as the internal knowledge, it is not necessary for a firm to keep various research programs inside, because it can exploit the external knowledge instead.

That is, the low availability of the external knowledge is also important for R&D diversification. As Teece (1980) summerized, there are a lot of transactional difficulties of the external knowledge. The transfer of knowledge is severely exposed to the problems of asymmetric information and uncertainty. These problems prevent the contract of technology transfer from being complete, hence it becomes difficult to utilize the knowledge of other firms at low costs. In this situation, it is more efficient to diversify R&D activity and keep various research menus than to receive the external knowledge.

Thus, the purpose of this paper is to clarify the effect of R&D diversification on the efficiency of inventive activity, in association with the availability of external knowledge. Although Henderson and Cockburn (1996) found out the significant spillover within a firm, they left the problem whether its impact exceeds the spillover from outside. Therefore, the exact role of R&D diversification has not been examined. This paper aims at covering this point by assessing how the importance of R&D diversification is influenced by the availability of external knowledge. [1]

To this end, some measures of diversification and accessibility to external knowledge should be constructed. All of these variables are derived from the detailed information on firms' patents, including the breakdown of patent applications by research category and the names of joint applicants. The former is used for constructing several indices of diversification. The latter corresponds to accessibility to external knowledge, which is qualified by not only the number of joint applicants, but also their contents. Since it is difficult to measure the degree of diversification exactly, multiple indices are examined to check the robustness of results.

This paper gives an insight of the relationship between internal and external knowledge as well. If the significance of R&D diversification decreases as the use of external knowledge becomes less difficult, it implies that a firm with wide variety of R&D activities can circumvent the difficulty of knowledge transfer. On the other hand, the ease-up on the difficulty in utilizing external knowledge allow the firm to focus on its research menu. Accordingly, the empirical results of this paper provide the implication on the specialization of R&D activities which is prevalent among Japanese firms of recent years.

[1] Some empirical studies including Lang and Stulz (1994) pointed out that corporate diversification decreases the firm's value measured by Tobin's q, etc. This paper does not discuss on this point. The focus of this paper is not on the total firm's value, but on the efficiency in the firm's R&D sector.

Our sample is major chemical companies in Japan. [2] Chemical companies are likely to have multiple research areas. The R&D activities in Chemical industry are highly diversified as compared with others. Statistics Bureau (1998) reports 28.9% of chemical firms doing research activities have R&D spending in more than two industrial classes in 1996, while the share of multi-inventing firms is 14.6% in machinery industry, 17.4% in electric machinery industry, 18.9% in transporting equipment industry, and 28.4% in precision instrument industry. The average of manufacturing is 23.3%. [3]

In addition, some case studies show the linkage among research areas in chemical companies. For example, Fukushima (1998) dealt with the case with one of our sample companies and described that new invention such as detergent was achieved by combining the research activities, such as biochemistry and polymer processing, which have been developed independently. Chemical engineering creates some materials, which can develop into other intermediate (e.g., chemical textile) and final goods (e.g., detergent). Moreover, the several development processes may interact.

The spillover within firms plays an important role in inventions of chemical companies. Nevertheless, the technologies invented by other firms can similarly contribute to R&D activities. Kagono et al. (1999) presented several cases with Japanese firms where the technology transfers enhance developing new business lines. Considering these effects of both internal and external knowledge spillovers, this paper tackles the problem how important R&D diversification is when the external knowledge can help new inventions.

This paper proceeds as follows. Section 2 provides the framework of the analysis, including the estimation methods. Two types of estimation methods are examined. One is the ordinary linear model with fixed effect. The other is the fixed-effect version

[2] Strictly speaking, some of the sample firms are not classified in checmial industry. They are paper and pulp, textile manufacturing, and ceramic engineering companies in the narrow classification, but they adopt chemical production lines broadly, so can be regarded as chemical companies practically.

[3] *Report on the Survey of Research and Development* figures corporate R&D expenditures by industry and product field. Product fields are classified into 31. The values reported here mean the proportion of firms with R&D spending in more than two fields only among those with intramural R&D expenditure, not all the surveyed firms.

of Poisson estimation to take care of count data property. Section 3 describes the brief sketch of the data. The detail of the data is mentioned in appendix. The results of main estimations are presented in section 4. This section consists of four parts. The first part reports the effect of R&D diversification on the efficiency of invention, without taking the availability of the external knowledge into consideration. The subsequent two parts test the key hypothesis that the extent of external knowledge spillover lowers the significance of R&D diversification. In the last part of this section, the premise on knowledge generating function, on which the estimations are based, is discussed. Section 5 summerizes the results and argues the remaining questions.

2 Framework

The purpose of the following sections is to investigate the role of R&D diversification by measuring its impact on the efficiency of research activity. The economies of scope in R&D have been already examined in Henderson and Cockburn (1996). They presented that the knowledge stock accumulated in a research program of a company enhance the productivity of the related programs in the company. As the first step, this paper also explores the existence of knowledge spillovers within a firm, though the data set used here is more aggregated than Henderson and Cockburn's. However, the analysis goes further to consider the relative intensity of internal knowledge spillovers to external knowledge spillovers. Hereby the effect of R&D diversification can be examined as an alternate for exploiting knowledge stock outside the company, which is likely to be less accessible.

The idea of this analysis comes down to the knowledge generating function defined below. New invention, that is, generating new knowledge requires current R&D

spending and available knowledge stock accumulated in past times. So the knowledge generating function at firm level is shown as:

$$N = f(R, K), \tag{1}$$

where N is new inventions, R is R&D spending, and K is the accumulated knowledge stock.

K includes the impact of knowledge spillovers both within a firm and across firms. Hence the 'effective' size of K depends on not only the 'raw' size of accumulated stock, but also the intensity of spillovers. Since two types of knowledge spillover are considered in this paper, there are two factors determining the intensity of spillovers. The degree of R&D diversification is one factor, which relates to knowledge spillover within a firm, as implied by Henderson and Cockburn (1996). When a firm has the wide range of its research menu, the effective size of internal knowledge increases because it is shared in multiple research programs.

Figure 1 describes this situation. In all the three cases, the total amount of knowledge stock is 10 in raw level. If the knowledge stock is devoted to one field, its effective size does not change from its raw level. Then suppose the case where the knowledge stock is divided equally into two fields. In case (b), the total size of knowledge stock in raw level remains the same as in the fully concentrated case, (a). However, this is not the case in effective level. When some part, say half, of the stock accumulated in one field is also costlessly available in another field, each of two fields exercises their research projects, as if they had half times larger knowledge in additional. Thus, the total size of knowledge in effect becomes one and a half time as large as in case (a). As the knowledge stock is diversified into various programs, its total extent for a given 'raw' level virtually becomes larger due to spillovers. In case (c), knowledge in one field is shared in the other four fields, hence the effective size of knowledge in this firm increases up to three times as large as in case (a).

The other factor is collaborated research, which influences the intensity of external knowledge spillovers. Branstetter and Sakakibara (1998) examined the effect of Japanese research consortia and found that it enhances fims' R&D efficiency. This suggests that joint research enables its participants to share their knowledge stock mutually. Thus, the extent of research collaboration is one of the proxies for accessibility to external knowledge. The relevance of this proxy will be discussed in the next section.

When the external knowledge substitutes to the internal knowledge at low cost, the effect of R&D diversification should be impaired. Firms need not keep various kinds of knowledge inside them, because the knowledge spilled over from outside can help inventions instead. Thus, the effective size of knowledge stock is defined as:

$$K = \phi(z)K^I + \psi(a)K^E, \tag{2}$$

where K^I and K^E are the size of internal and external knowledge, respectively. Their effective size is influenced by the degree of R&D diversification, z, and the proxy of accessibility to the external knowledge, a. The two types of spillover are regarded as substitutes. This assumption is discussed in section 4. The relation described so far is illustrated in Figure 2.

Combining (1) and (2), the knowledge generating function is redefined as

$$N = f(R, \phi(z)K^I + \psi(a)K^E). \tag{3}$$

The function f satisfies the ordinary assumptions such as $f_1 > 0$, $f_2 > 0$, $f_{11} < 0$, and $f_{22} < 0$. $\phi' > 0$ and $\psi' > 0$ are also held, since the more diversified firm exploits internal knowledge more efficiently and the firm with higher accessibility to external knowledge takes advantage of more enormous spillovers.

On the basis of these assumptions on (3), two main hypotheses are posited. First, from the assumptions that $f_2 > 0$ and $\phi' > 0$, the more diversified firm increases its efficiency of R&D activities. Second, the impact of R&D diversification is reduced as the external knowledge becomes more available, because $f_{22} < 0$ and $\psi' > 0$.

The framework of estimation to test these hypotheses follows simplified versions of (3). Estimation methods take two forms. One is the ordinary linear model with fixed effects to deal with firm-specific factors like the efficiency of innovation and the propensity to patenting. The other is the fixed-effect version of Poisson regression proposed by Hausman, Hall, and Griliches (1984). The dependent variable here is the number of patent applications, which is restricted to some integer. Poisson regression is one of the widely used methods to take care of the count data property.

The estimation equations in the linear model are specified as in Branstetter and Sakakibara (1998), who were based on the knowledge generating function similar to (3), used for the study on Japanese research consortia. K^I and K^E are not exactly observed, [4] so the estimate equation has only R&D expenditure, the degree of R&D diversification, and the proxy of accessibility to the external knowledge as explanatory variables. For simplicity, (3) is assumed to be in the form of Cobb-Douglas type. The equations are basically specified as its logarithmic form.

The first hypothesis focuses on the impact of R&D diversification only, so the corresponding equation is:

$$\log N_{it} = \beta \log R_{it} + \gamma \log z_{it} + u_{it}. \tag{4}$$

i and t indicate individual and time, respectively. Note that there is some unobservable elements such as the research ability of each firm. Thus, u_{it} takes the fixed-effect form:

$$u_{it} = \alpha_i + \epsilon_{it}. \tag{5}$$

If $\gamma > 0$, the first hypothesis is supported.

The second hypothesis indicates that gamma varies with the availability of the external knowledge, a. The general form of estimate equation for the second hypothesis can be written as

$$\log N_{it} = \beta \log R_{it} + \gamma(a_{it}) \log z_{it} + u_{it}, \tag{6}$$

[4] Since the data on firms' patents are available by research field, the external knowledge stock can be numerically calculated after Jaffe (1986)'s method. as long as the knowledge of the sample firms are concerned. However, the sample size is so small that the calculated knowledge stock is not considered to reflect the external spillover exactly.

where u_{it} satisfies (5) again. Two variants of (6) are estimated to test the second hypothesis depending on how to measure a. One measure of a reflects just the quantative aspect of accessiblity to external knowledge, denoted by s. Accessibility to external knowledge is assessed by its size, i.e., frequency of joint research only. Since there is no information of the functional form of $\gamma(\bullet)$, the actual estimation postulates a linear form as

$$\gamma(s_{it}) = \gamma_0 + \gamma_1 s_{it}. \tag{7}$$

The main interest is in the sign of γ_1. If the second hypothesis is true, then $\gamma' < 0$, so negative γ_1 is expected. The sign of $\gamma(s_{it})$ is also examined on the basis of the estimated γ_0 and γ_1.

However, the impact of external knowledge may depend on not only its size, but also its contents. If internal knowledge is substitutable for external knowledge, the impact of external knowledge accessed may become weak when it overlaps largely the firm's R&D menu. In this situation, a should be measured by two dimensions, that is, quantity and quality. Thus, another specification of $\gamma(\bullet)$ is introduced as

$$\gamma(s_{it}, d_{it}) = \gamma_0 + \gamma_1 s_{it} + \gamma_2 d_{it}, \tag{8}$$

where s denotes again the size of available external knowledge, while d means its degree of difference from internal knowledge stock. If the hypothesis mentioned above is true, $\gamma_1 < 0$ and $\gamma_2 < 0$. The estimated value of γ is also calculated similarly.

The Poisson fixed-effect model imposes the multiplicative effect on the ordinary Poisson regression model. Thus, conditional mean of dependent variable is described as

$$E[N_{it}|\boldsymbol{X}_{it}, \alpha_i] = \alpha_i \exp(\boldsymbol{X}'_{it}\boldsymbol{\beta}), \tag{9}$$

where α_i is again firm-specific and time invariant. $\boldsymbol{X}_{it}$ is the set of independent variables and beta is its coefficient vector. In estimations, they are specified as some variants of (4) or (6). The parameters, $\boldsymbol{\beta}$, are estimated by maximizing the log likelihood function conditional on the sufficient statistics for α_i's, that is, the individual-specific totals of dependent variables. The conditional log likelihood function becomes

$$\mathcal{L}_c(\boldsymbol{\beta}) = \sum_i \left[\ln \left(\sum_t N_{it} \right)! - \sum_t \ln(N_{it}!) + \sum_t N_{it} \ln \left(\frac{\exp(\boldsymbol{X}_{it}'\boldsymbol{\beta})}{\sum_{s=1}^{T} \exp(\boldsymbol{X}_{is}'\boldsymbol{\beta})} \right) \right]. \quad (10)$$

Statistical inferences on coefficients are as in the linear model.

3 Data

The sample firms are 32 Japanese major chemical companies and the sample period is from 1985 to 1995. Some of key variables are derived from the data on patent applications of these companies. The patent data are provided by "*Kagaku Kigyō no Dōkō to Senryaku*" (The Performance and Strategy of Chemical Firms), edited by Kagaku Gijutsu Tokkyo Chōsa-kai (survey group of patents by chemical technology). This book reports the details of patent applied by approximately 30 major companies in Japanese chemical industry. [5] This data source reports the number of patent application in each research field and appearance of jointly applied patents. The former is used to construct the indices of R&D diversification and the latter provides information on collaborated research. The sample period is defined as starting at the time when the reported classification of patents became common to all the sample firms.

Patent application

New invention, N, is measured by the number of patent applications in each year. It is often pointed out that the patent application is not an exact measure for the results from corporate research activities, since the applied patents do not always have 'novelty' as new inventions. Accordingly, some empirical studies adopt the number

[5] Two of the reported firms are omitted in this analysis because of an experience of merger in the sample period.

of granted patents as the indicator of new inventions instead of patent applications. [6] However, as for major chemical companies analyzed here, most of applied patents become granted later. [7] Therefore, as long as this case is concerned, the number of applications is considered to reflect the true results of R&D fairly well. Apart from this, the patent does not necessarily coincide with N due to the difference in the propensity to patenting among firms. It is involved in the fixed effect term here.

Indices of diversification

As for the indices of diversification, three types of them are adopted among several candidates. Note that these indices are derived from the patent data mentioned above. So they do not reflect the distribution of R&D spending in the current year, but of the knowledge stock accumulated in past years.

The first index of R&D diversification is the share of patents in 'side lines,' which is denoted by $SIDE$ henceforth. Kagaku Kigyō no Dōkō to Senryaku classifies patents into 16 fields based on International Patent Classification. Among 16 fields, [8] that with maximum share is define as 'main line,' and the others are regarded as side lines. [9] That is,

$$SIDE = 1 - \max_{j} \sigma_j, \tag{11}$$

where σ_j is the share of the field i. When a firm sets up a new research field or extends a formerly minor field, the share of the greatest field is likely to drop. Thus, $SIDE$ is getting larger, as a firm becomes more diversified.

[6] Henderson and Cockburn (1996) used the more narrowly defined proxy. Thy measured the research results by the number of 'important' patents granted in two of Japan, Europe, and the United States.

[7] Japanese Patent Office published the names of firms with over 70% registration rate (that is, the ratio of registrated patents to those requested for examination). All the sample firms except for two are included in the publication, and 16 of them record over 80% registration rate.

[8] The fields are: (1) agricultural product, (2) food product, (3) biotechnology, (4) medical equipment and pharmaceuticals, (5) cosmetics and perfumery, (6) inorganic chemical, (7) organic chemical, (8) high polymer chemistry, (9) metals, (10) processing and manipulation, (11) textile. paper, and pulp, (12) machinery, (13) construction and civil engineering, (14) information technology and electronics, (15) analysis and measurement, and (16) otherwise.

[9] By definition, the 'main line' is not necessarily fixed during the sample period. It is defined just as the dominant field in each year.

Second, Herfindahl index is also often the measure of diversification. Herfindahl index is defined as

$$HI = \sum_j \sigma_j^2. \tag{12}$$

This value becomes smaller when patenting in minor fields goes up and that in major fields falls down. When research activities are concentrated in one field, HI increases toward one. For the lucidity of expression, the second index of diversification is presented like

$$HERF = 1 - HI. \tag{13}$$

This is positively related to the degree of diversification and converges to one as the diversification proceeds.

Although these two indices are often used in the studies on diversification, [10] their values are not free from how to classify research fields. Thereby this paper introduces the third index which is less subject to classification. Biotechnology and electronics are newly developed technologies since 1980s and are different from the traditional field of chemical companies. Therefore, more diversified firms are likely to extend R&D activities in these fields. The third index of R&D diversification is defined as the share of patents in the two fields and denoted by $BIO\&ELEC$. [11]

Joint applicants

Since the accessibility to the external knowledge is not observed directly, it is necessary to introduce its proxy. In this paper, the proxy is derived from the data on joint applicants including firms, organizations, and individuals. Although there are various kinds of the ways to exploit external knowledge either implicitly or explicitly, the joint applicant is favorable to our analysis. First, the extent of implicit spillover is difficult to measure numerically. Moreover, it may be reflected in the fluctuation of R&D ex-

[10] See Goto and Suzuki (1987), for example.

[11] The patents in biotechnology are defined as those coded with C12 in terms of IPC. Electronics addresses the code of B41, F21, G02-G12, and H01-H05. The latter definition seems broader than usually considered.

penditure. [12] Second, technical collaboration and technology transfer agreement are strong candidates as explicit ways of spillovers. They are reported in firms' financial statements. However, the reported cases are highly limited to what is considered to be 'important.' Accordingly, their fluctuation among times and firms are not so reasonable in this analysis.

Thus, the appearance of joint applicants is more appropriate than others. Branstetter and Sakakibara (1998) suggested that collaborated research between firms promotes the knowledge spillover between them. As the variety of joint applicants expands, the firm can exploit highly rich stock of the external knowledge more easily. One of the indices used in estimations is the logarithmic number of joint applicants, which is denoted by $JOINT$. [13]

$JOINT$ reflects the information on the size of accessible external knowledge, which is denoted as s in the previous section. Nevertheless, it does not qualify the contents of externally spilled-over knowledge. Since names of joint applicants are available, [14] the measure of accessibility to external knowledge can be modified to consider qualitative aspect of joint applications. It is reflected in the degree of technological difference between a firm's research activities and external knowledge accessed. In this paper, technological difference between firms, denoted by d in the previous section, is measured in a simple way. If a firm collaborates with non-chemical companies, the spilled over knowledge is quite different from what it has internally, so the benefit of joint research may become larger than that in the case of collaboration with chemical companies. Thus, the proportion of non-chemical companies in joint applicants raises the quality of collaborated research. To see this, $NONC$, which is defined as the logarithmic ratio

[12] According to Cohen and Levinthal (1989), the amount of R&D expenditure indicates the firm's capacity to absorb the external knowledge.

[13] It is one critical problem that joint application is not come out before the collaborated research is successfully realized. If the collaborated research has not brought patents, or if the collaborator(s) restrain to file its (their) name(s) as applicant(s), the data fail to observe the existence of collaboration. This sometimes happens to the collaborations with academia. In this paper, such errors of omission are caused at common rate for all the sample firms.

[14] One of the sample firms is dropped in constructing this measure due to lack of the data.

of non-chemical joint applicants, is added to explanatory variables in estimating (6). [15]

R&D expenditure

R&D expenditure is derived mainly from *Kaisha Shiki Hō*. The data is available only at firm level, so the spillovers within a firm cannot be examined as directly as in Henderson and Cockburn (1996).

The descriptive statistics of main variables are presented in Table 1. The detailed information on the data is provided in the Appendix.

4 Results

This section is for testing the hypotheses presented in the previous section. The hypotheses are arranged as below.

H1: A firm with more diversified R&D activities is more productive in its R&D. That is, in (4) and (6), $\gamma > 0$.

H2: The effect of R&D diversification is reduced when external knowledge is easily available. That is, in (6) with either (7) or (8), $\gamma_1 < 0$.

H3: The effect of R&D diversification is reduced when external knowledge a firm exploit is derived largely from non-chemical companies. That is, in (6) with (8), $\gamma_2 < 0$.

By examining H1-H3, the exact role of R&D diversification is clarified and we have a viewpoint of the relation between internal and external knowledge. The robustness

[15] Some of the sample firms are classified as textile or pulp producing companies. Thus, *NONC* is defined as the ratio of joint applicants in the industries where the sample firms do not belong.

of results is checked by applying two estimation methods, the linear model and the Poisson model, for essentially similar equations.

Note that one-year lag is taken for explanatory variables. This is because the source of those variables is the patent data which provides the dependent variable as well. Hence the common shock on the efficiency of R&D may cause spurious correlation between patent applications and the explanatory variables, such as R&D diversification and joint applications. Taking lag of explanatory variables deals with this problem. The one-year lagged distribution of patent applications reflects that of knowledge stock accumulated up to the previous year. In other words, the estimation here postulates that a firm determines the menu of R&D, depending on the variety of internal knowledge and the availability of external knowledge realized in the previous year.

4.1 The effect of R&D diversification

Table 2 shows the results from the linear model estimation of (est1) to test H1. [16]

The impact of R&D diversification is explored by using three types of indices. In all the columns from (a) to (c), the coefficients of R&D diversification, γ, are significantly positive. As the diversification extends, the firm's R&D becomes more productive or efficient. The comparable results are obtained by the Poisson regression with fixed effects as shown in Table 3. All of coefficients are significantly positive again, and take almost close values to what is reported in Table 2. Accordingly, we can observe the positive influence of R&D diversification, which is considered to be due to promoting the knowledge spillovers within a firm, as Henderson and Cockburn (1996) suggested. [17]

[16] Although the impact of biotechnology is equivalently regarded as that of electronics in these estimations, the equivalence should be checked. Note that the 'log 0' problem occurs by dividing BIO&ELEC into two parts, because some firms applied no patent in either field in some years. When zero is replaced by a tiny value, 0.0000001, F test did not reject the hypothesis of equivalence between two fields in linear estimation and neither Wald test in Poisson regression.

[17] The estimated coefficient on R&D spending is much smaller in all the tables than what has been reported in previous studies. One explanation to this may be the difference of independent variable. This paper measures research outcomes by patents in the stage of application, while most of previous

Comparing the three cases, we find the coefficient of $BIO\&ELEC$ is smaller than that of $SIDE$. This difference is explainable by the definition of the variables. In most observations, biotechnology and electronics are included in side lines. γ is the elasticity of R&D diversification on patenting, that is, the percentage increase in patent applications when the index of R&D diversification increases at 1%. Thereby the mean value of BIO&ELEC is about 29% of that of $SIDE$ in Table 1, so the 1% increase of $BIO\&ELEC$ is, ceteris paribus, almost equal to 0.29% increase in $SIDE$ on average. Thus, on the basis of the point estimates in Table 2, for example, 0.29% increase in $SIDE$ causes 0.11% increase in patent applications, which is not largely different from the result in column (c).

The influence of joint applications on R&D diversification

The central interest of this paper is in how influential the accessibility to external knowledge is on the impact of R&D diversification. The results on this point are reported in Table 4 and Table 5. In both tables, estimate equations include the interaction term between the indices of R&D diversification and the quantitative measure of joint applicants, which allows the elasticity of R&D diversification varying with the frequency of joint applications.

The estimation results support H2 unambiguously. The coefficient of the interaction term, that is, γ_1 in (6), is significantly negative regardless of the types of indices for diversification and applied estimation methods. This implies that R&D diversification affects more intensely on the productivity of research activities when a firm suffers from the difficulty to access external knowledge more severely. On the contrary, if a firm benefits easily from knowledge spillovers across the firms, the firm needs not keep the variety of its R&D menu in order to put spillovers within the firm to good use. [18]

studies use narrower concept, for example, granted patents and 'important' patents. Thus, our measure includes 'lower quality' outcomes than others. If the quality of outcomes is likely to be improved as more R&D spending is devoted for a given (probably firm-specific) innovative ability, patent application is less sensitive to R&D spending than granted patent and 'important' patent.

[18]There can be leads and lags between joint applications and the influence of the knowledge held in joint applicants. The timing of collaborated research may not necessarily correspond with joint applications. Some joint projects may cease before applying patents on the research results. On the

The bottom three rows report the estimated values of the total impact of R&D diversification, $\gamma(s) = \gamma_0 + \gamma_1 s$. They are computed from the point estimates in Table 3, the average, minimum, and maximum values of JOINT. The average values produce the elasticities of R&D diversification quite close to what is reported in Table 2. It is remarkable that the estimated elasticity keeps positive value even at the minimum level, when calculated from the maximum value of $JOINT$. Thus, R&D diversification still has positive impact on the efficiency of research activities, even if the variable elasticity is postulated in the form of (7).

4.2 The qualitative measure of joint applicants

In the previous subsection, every joint applicant is treated equally. They have the same influence on knowledge spillover in spite of their technological characteristics. However, it is natural to consider that spillovers from joint applicants in chemical industry are technologically close to what the sample firms keep internally, while those from non-chemical companies are relatively novel to them.

Considering the difference of technological characteristics, this subsection adds the qualitative measure of joint applicants to estimations. The more non-chemical companies are involved in joint applicants, knowledge spillovers due to joint research include more novelty for the firm. The estimation results are presented in Table 6 and Table 7. [19] Unlike other cases, the results are somewhat inconsistent between two estimation methods. The results from the linear model, reported in Table 6, show insignificant γ_2 with expected sign, while all of other coefficients are significant and have correct sign.

other hand, some of the joint applicants may still keep contact with each other after patent applications. To deal with this problem, the three-year average values of joint applicants are also examined. That is, the proxy of joint research is reconstructed as the average of current, one-year lagged, and one-year led values. This modification provided the results on R&D diversification and joint research similar to what reported in Table 4, though the coefficients of R&D expenditure dropped sharply and lost significance.

[19] The quality of joint applicants can be measured by slightly different indices mentioned in the data appendix. The estimations based on such measures also bring the similar results except for negative lower bound of $\hat{\gamma}$ in some cases.

On the contrary, the results of the Poisson regression, provided in Table 7, have all estimates, including γ_2, with expected sign at the strongly significant level.

Although there is inconsistency between the results of two models, the problem of insignificance is not so severe in Table 6. Especially for column (c), the p-value for γ_2 is 0.107, so it is quite close to the significant level. Thus, as a whole, the hypothesis H3 can be accepted. The technological distance from joint applicants depresses the importance of keeping varied research menu inside because external knowledge accessed is highly novel to the firm.

The range of estimated γ, which is derived from point estimates of Table 6 and Table 7, is provided the bottom three rows in each table. Although the average values are slightly lower than what is obtained from other specifications except for one case, they remain positive even at the lower bound. The positive impact of R&D diversification is robust result.

4.3 External spillover and internal spillover

Before closing this section, it should be checked whether external knowledge is substitutable for internal knowledge. Hypothesis H2 is based on substitution between internal and external knowledge, so it is a critical point for interpreting the estimation results. The results reported above are coherent with substitutability of two types of spillovers, when R&D diversification enhances the productivity of R&D as shown in the tables discussed thus far. If the spilled-over external knowledge is complement to internal knowledge spillover, the coefficient of the interaction term should have been positive, because more diversified firm raises its effect from the external knowledge.

The evidence on substitutability can be seen also in the data of joint applicants. Only one-fifth of joint applicants are clearly classified in the same industry as the sample firms on average. [20] This suggests that the firms are inclined to seek companies with

[20] Even when the ratio is calculated only for the joint applicants which are identified their industry,

different types of technologies as their research collaborators. This is likely to happen if firms consider external knowledge as substitutable for their own knowledge.

In order to clarify the problem, the relation between the two variables, the indices of R&D diversification and $JOINT$, are examined. [21] Note that the levels of these variables are probably subject to the spurious correlation, since both of them are derived from the common patent data. To avoid this, the check is to analyze the impact of joint applications on the variation of diversification. As mentioned above, the indices of diversification reflect the distribution of knowledge stock, so their variations correspond to the current flow of newly accumulated knowledge. Table 8 presents the results from the fixed-effect estimations. In all cases, the coefficient of $JOINT$ is negative. Thus, the more joint applicants hold down spreading inside R&D menu, because the external knowledge becomes more accessible. This implies that the external knowledge is substitutable for the spillovers of internal knowledge.

Another specification was tested in the form with $NONC$ added as well. The coefficient of $JOINT$ was negative in all cases, moreover significant in some cases. However, the results on $NONC$ were quite unstable. The sign of its coefficient reversed depending on specification of estimation equations. The ratio of non-chemical companies in joint applicants has two effects on R&D diversification opposite to each other. It reduces the necessity to keep varied research menu within a firm, while it expands the variety of research output derived from collaborated research. As reported in Table 8, the goodness of fit is extremely low, thus minor modification of model may cause drastic change of resultant sign. In any case, the large size of joint research with non-chemical

it is at most a quarter on average.

[21] Another specification was tested in the form with $NONC$ as well. The coefficient of $JOINT$ was negative in all cases, moreover significant in some cases. However, the results on $NONC$ were quite unstable. The sign of its coefficient reversed depending on specification of estimation equations. The ratio of non-chemical companies in joint applicants has two effects on R&D diversification opposite to each other. It reduces the necessity to keep varied research menu within a firm, while it expands the variety of research output derived from collaborated research (Mowery et al. (1998) show that research collaboration induces the technological overlap among collaborators). As reported in Table 8, the goodness of fit is extremely low, thus minor modification of model may cause drastic change of results. In any case, the large size of joint research with non-chemical companies is inclined to restrain diversifying R&D activities.

companies is inclined to restrain diversifying R&D activities.

Some empirical studies, such as Cohen and Levinthal (1989), emphasize the complementarity of internal effort on R&D and the external knowledge. This does not contradict the discussion above. In this paper, the coefficient of R&D spending is significantly positive in all cases. From the functional forms of the estimate equations, this result suggests that internal R&D effort can be complement to the spilled-over knowledge stock, either internal or external. R&D spending raises the productivity of the knowledge stock and it does not matter whether the knowledge stock is accumulated inside or outside the firm.

5 Conclusion

This paper empirically inquires into the role of R&D diversification, in association with the accessibility to external knowledge. The results are summerized as:

- R&D diversification enhances the productivity of R&D due to the knowledge spillover within a firm.
- This effect depends on the accessibility to external knowledge measured by the number of joint patent-applicants. That is, the effect is more significant when it is difficult to exploit the knowledge spillovers across firms.
- The impact of R&D diversification depends on not only the size of external knowledge, but also its contents as shown at least in the Poisson regression. If a firm collaborates largely with non-chemical companies, it is less necessary to keep wide variety of R&D menu inside, because novel knowledge is easily available.

Inventive activity is the combination of the current research effort with the knowledge stock accumulated in past years. When there is some barrier to utilize knowledge accumulated externally, a firm can circumvent the difficulty to keep the wide range of R&D menu inside. On the contrary, if it takes low costs to exploit the external knowledge, the firm needs not rely on internal spillover, so the importance of R&D diversification decreases.

There are three remaining problems. First, since this paper shows that externally spilled-over knowledge is substitutable for the internally accumulated knowledge, the next question is what determines a firm's choice between them. Furthermore, even if a firm chooses keeping the knowledge accumulated internally, the direction of diversification is the next problem. In fact, the sample firms show various aspects of R&D diversification. For example, the share of patents in biotechnology is negatively correlated with that in electronics (the correlation coefficient is about -0.33). The determinants of these choices should be examined.

Second, the contents of external knowledge can be qualified in different ways from what this paper treats. This paper takes care of technological distance among joint applicants in quite simple way. The sole criterion is whether they belong to chemical (or textiles, or pulp producing) industry or not. Instead of such discrete choice, Jaffe (1986) proposed 'technological distance' as the distance between innovation menus of firms. Using such a measure may bring clearer insight on the relation between internal and external knowledge spillover. Moreover, apart from technological aspect, business intimacy may influence the significance of external spillover. The contact with firms in the same business group may be stronger than others. To assess the extent of knowledge spillover precisely, the heterogeneity of joint applicants should be taken into consideration more carefully.

Finally, it is interesting to explore how technological diffusion process is affected by the knowledge flow among firms. Intense knowledge flow contributes to rapid diffusion, while it may reduce firms' incentive to incorporate the new technology, since the high accessibility enables a firm to do R&D without holding knowledge inside. Such a macroeconomic problem remains as a task for future.

A Data Appendix

A.1 Patent

Japanese patent system obligates that the applied patent has to be disclosed one and a half year after the filing date. *Kagaku Kigyō no Dōkō to Senryaku* makes use of this system and defines the number of patent applications in a year. For example, the patents disclosed from July 1991 to June 1992 can be regarded as applied during the year of 1990. The number of patent and joint applicants and three indices of R&D diversification are all based on this definition.

The patent applications are classified in terms of the IPC code reported first in the patent information. The classification is owe to Kagaku Kigyo no Doko to Senryaku.

A.2 Joint applicants

Kagaku Kigyō no Dōkō to Senryaku provides the names of joint applicants as well. On the basis of this information, they are classified into two groups, that is, those in the same industry as the sample firms and others. The latter contains individuals and some organizations like universities and national research institutes. Classification is according to *Kaisha Shiki Hō* and firm list for declared income ranking published by Tōyō Keizai Shinpō-sha. However, this data source does not cover all joint applicants. For more than 10% of joint applicants, no information their industry is available. $NONC$ used in Table 6 and Table 7 neglects such unidentified joint applicants. That is, it is defined as the logarithm of

$$1 - \frac{\text{\#joint applicants clearly classified in chemical/textile/pulp}}{\text{\# total joint applicants}}.$$

Another definition of quality measure for joint applicants was also constructed as the logarithm of

$$1 - \frac{\text{\#joint applicants clearly classified in chemical/textile/pulp}}{(\text{\#total joint applicants - \# unidentified})}.$$

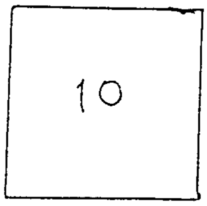
The background of this definition is the assumption that unidentified firms have the same distribution of industries as the identified. This modified variable, denoted by $NONC2$, however, did not change the estimation results essentially.

A.3 R&D expenditure

The basic source of R&D expenditure is Kaisha Shiki Hō, published by Tōyō Keizai Shinpō-sha. The deficit data in this book are fixed up by the data from *Nikkei NEEDS* and Kagaku Kigyo no Doko to Senryaku. The estimated R&D expenditure is the figures from these complementary sources multiplied by the average ratio of Kaisha Shiki Hō's data and the complementary data. The nominal values are converted into the real value with the R&D spending deflator (Companies, etc.), which is derived from *White Paper on Science and Technology* in each year.

Figure 1. Effective size of knowledge and spillovers within a firm

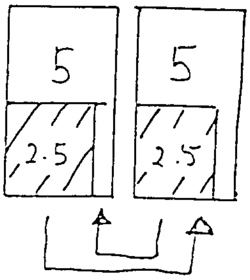
a) Fully concentrated case



effective size = 10

field 1

b) two-field diversified case



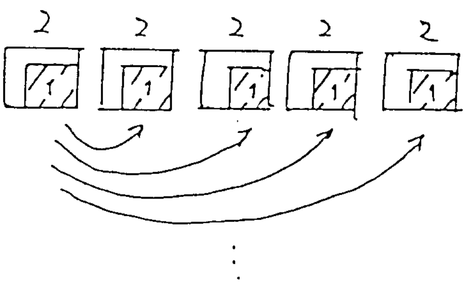
Half of knowledge stock in each field is shared in another field.

effective size $= 5 + 2.5 + 5 + 2.5$

$= 15$

field 1 field 2

c) five-field diversified case



Half of knowledge stock in each field is shared in another field.

effective size $= (2 + 1 \times 4) \times 5$

$= 30$

Figure 2. The internal and external knowledge flow

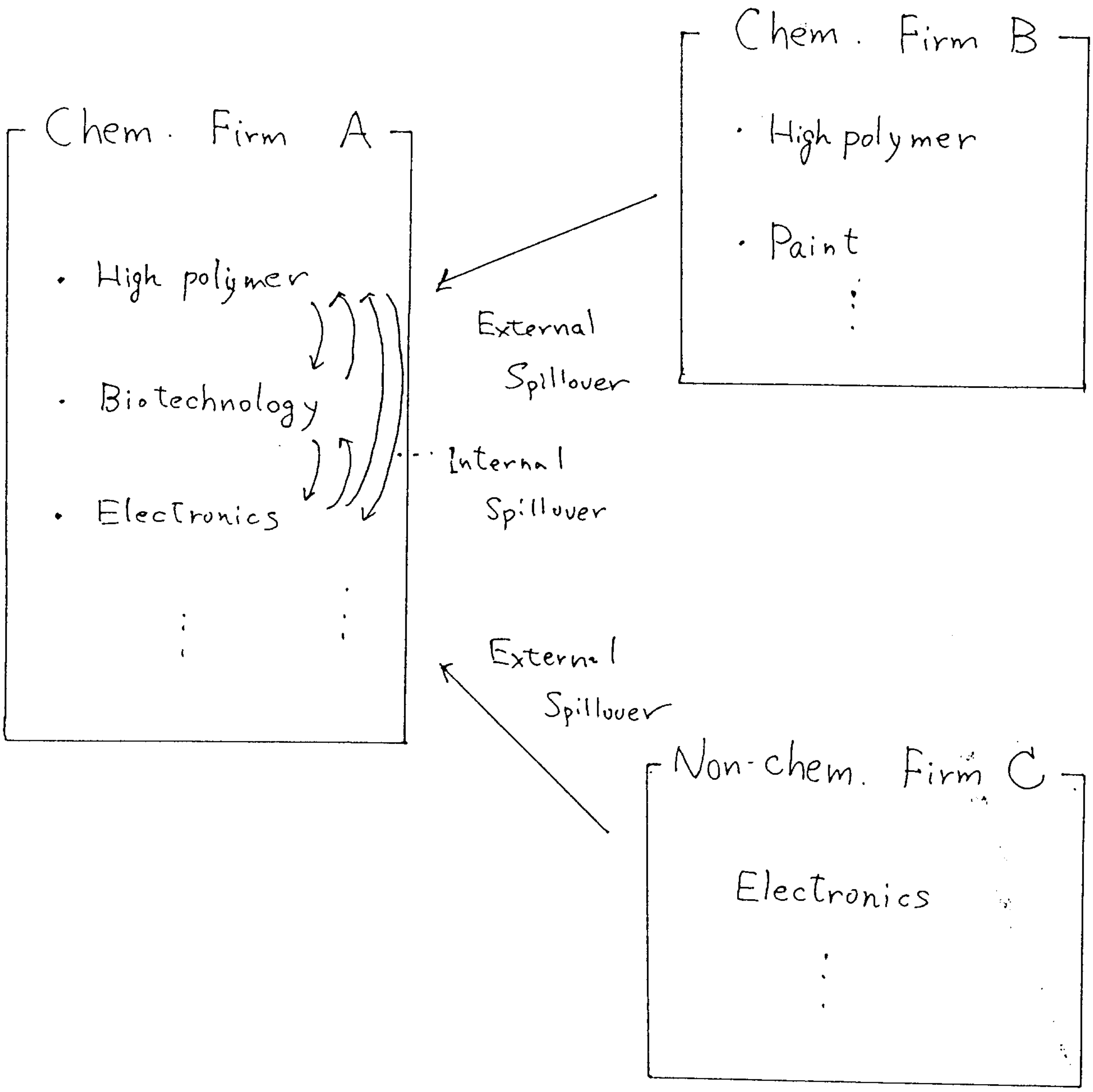


Table 1. Descriptive statistics of main variables

	mean	min	max	s.d.
patent application	477.35	59	1542	300.36
SIDE	0.6137	0.1336	0.8322	0.1179
HERF	0.7709	0.4698	0.8986	0.0711
BIO&ELEC	0.1758	0.0297	0.4672	0.0812
joint applicant	24.49	4	93	14.33
the share of non-chemical joint applicant	0.7879	0.2941	1.0000	0.1474

number of obs. =352

Table 2. The impact of R&D diversification on the productivity of R&D: the linear regression model

	(a)	(b)	(c)
measure of Div.	*SIDE*	*HERF*	*BIO&ELEC*
$\log R\&D$	0.2106 ***(0.0548)	0.2162 ***(0.0549)	0.1784 ***(0.0557)
z	0.3797 ***(0.1210)	0.5583 ***(0.2058)	0.1493 ***(0.0407
adjusted R^2	0.895	0.894	0.896

number of obs. =352

The figures in parentheses are standard deviations.

***: Significant at 1% level

Table 3. The impact of R&D diversification on the productivity of R&D: the Poisson regression model

	(a)	(b)	(c)
measure of Div.	*SIDE*	*HERF*	*BIO&ELEC*
$\log R\&D$	0.2423 ***(0.0117)	0.2419 ***(0.0118)	0.1823 ***(0.0121)
z	0.4584 ***(0.0296)	0.6843 ***(0.0527)	0.1732 ***(0.0102)
$\log L$	-4473.76	-4509.67	-4450.51

number of obs. =341 ($N = 31$, $T = 11$)

The figures in parentheses are standard deviations.

***: Significant at 1% level

Table 4. The effect of R&D diversification varying with joint research: the linear regression model

	(a)	(b)	(c)
measure of Div.	*SIDE*	*HERF*	*BIO&ELEC*
$\log R\&D$	0.1821 ***(0.0544)	0.1902 ***(0.0545)	0.1580 ***(0.0550)
z	0.9234 ***(0.1923)	1.4908 ***(0.3332)	0.2791 ***(0.0541)
$z \times JOINT$	-0.1790 ***(0.0498)	-0.3140 ***(0.0892)	-0.0544 ***(0.0153)
ajusted R^2	0.899	0.898	0.900
Estimated elasticity of diversification:			
average	0.3388	0.4816	0.0896
min	0.0856	0.0206	0.0152
max	0.7843	1.3448	0.2272

number of obs. =352

The figures in parentheses are standard deviations.

***: Significant at 1% level

Table 5. The effect of R&D diversification varying with joint research: the Poisson regression model

	(a)	(b)	(c)
measure of Div.	*SIDE*	*HERF*	*BIO&ELEC*
$\log R\&D$	0.2141 ***(0.0118)	0.2201 ***(0.0118)	0.1638 ***(0.0120)
z	1.0695 ***(0.0483)	1.7671 ***(0.0852)	0.3232 ***(0.0138)
$z \times JOINT$	-0.1908 ***(0.0119)	-0.3467 ***(0.0212)	-0.0586 ***(0.0037)
$\log L$	-4344.34	-4374.77	-4324.11
Estimated elasticity of diversification:			
average	0.4594	0.6582	0.1359
min	0.2049	0.1956	0.0578
max	0.8051	1.2865	0.2420

number of obs. =341 (N = 31, T = 11)

The figures in parentheses are standard deviations.

***: Significant at 1% level

Table 6. The quantity and quality effect of joint research: the linear regression model

	(a)	(b)	(c)
measure of Div.	*SIDE*	*HERF*	*BIO&ELEC*
$\log R\&D$	0.1987 ***(0.0595)	0.2102 ***(0.0595)	0.1739 ***(0.0603)
z	0.8214 ***(0.2076)	1.2968 ***(0.3489)	0.2232 ***(0.0607)
$z \times JOINT$	-0.1623 ***(0.0504)	-0.2816 ***(0.0901)	-0.0459 ***(0.0158)
$z \times NONC$	-0.1536 (0.1402)	-0.3582 (0.2550)	-0.0553 (0.0342)
adjusted R^2	0.901	0.900	0.902
Estimated elasticity of diversification:			
average	0.3388	0.4816	0.0896
min	0.0856	0.0206	0.0152
max	0.7843	1.3448	0.2272

number of obs. =341 ($N = 31$, $T = 11$)

The figures in parentheses are standard deviations.

***: Significant at 1% level

Table 7. The quantity and quality effect of joint research: the Poisson regression model

	(a)	(b)	(c)
measure of Div.	*SIDE*	*HERF*	*BIO&ELEC*
$\log R\&D$	0.2442 ***(0.0133)	0.2541 ***(0.0133)	0.1931 ***(0.0137)
z	0.8674 ***(0.0524)	1.4344 ***(0.0896)	0.2133 ***(0.0156)
$z \times JOINT$	-0.1609 ***(0.0122)	-0.2953 ***(0.0218)	-0.0431 ***(0.0038)
$z \times NONC$	-0.3936 ***(0.0413)	-0.7897 ***(0.0745)	-0.1337 ***(0.0104)
$\log L$	-4049.73	-4075.31	-4023.38
Estimated elasticity of diversification:			
average	0.4467	0.6781	0.1074
min	0.1383	0.0959	0.0180
max	1.1261	1.9915	0.3173

number of obs. =341 ($N = 31$, $T = 11$)

The figures in parentheses are standard deviations.

***: Significant at 1% level

Table 8. The impact of joint applications on R&D diversification

	(a)	(b)	(c)
Independent variable:	$\Delta SIDE$	$\Delta HERF$	$\Delta BIO\&ELEC$
JOINT	-0.0121 *(0.0070)	-0.0075 (0.0049)	-0.0121 **(0.0061)
adjusted R^2	-0.047	-0.043	-0.056

number of obs. =352

The figures in parentheses are standard deviations.

**: Significant at 5% level

*: Significant at 10% level

References

[1] Branstetter, Lee, "Are Knowledge Spillovers International or Intranational in Scope? Microeconometric Evidence from the U.S. and Japan," National Bureau of Economic Research Working Paper No. 5800, 1996.

[2] ———————— and Mariko Sakakibara, "Japanese Research Consortia: A Microeconometric Analysis of Industrial Policy," *Journal of Industrial Economics,* 46(2), 207-33, 1998.

[3] Cohen, Wesley M. and Daniel A. Levinthal, "Innovation and Learning: The Two Faces of R&D," *Economic Journal,* 99(397), 569-96, 1989.

[4] Fukushima, Michi, "Seijuku Shijō ni okeru Inobēshon — Kao "Attack" no Kaihatsu (Innovation from the Matured Market: the Development of "Attack" by Kao)," (in Japanese), in H. Itami et al. ed., *Inobēshon to Gijutsu Chikuseki* (Innovation and Technology Accumulation), Yuhikaku, 1998.

[5] Goto, Akira and Kazuyuki Suzuki, "The Diversification of R&D and the spillover effect of technology," *The Economic Review* (Keizai Kenkyū) (in Japanese), 38(4), 298-306, 1987.

[6] Hausman, Jerry, Bronwyn H. Hall, and Zvi Griliches, "Econometric Models for Count Data with an Application to Patents-R&D Relationship," *Econometrica,* 52(4), 909-38, 1984.

[7] Helpman, Elhanan ed., *General Purpose Technologies and Economics Growth,* MIT Press, 1998.

[8] Henderson, Rebecca and Iain Cockburn, "Scale, Scope, and Spillovers: the Determinants of Research Productivity in Drug Discovery," *RAND Journal of Economics,* 27(1), 32-59, 1996.

[9] Jaffe, Adam B, "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value," *American Economic Review,* 76(5), 984-1001, 1986.

[10] Jovanovic, Boyan, "The Diversification of Production," *Brookings Paper on Economic Activity: Microeconomics* 1, 1993.

[11] Kagaku Gijutsu Tokkyo Chōsa Kai, *Kagaku Kigyō no DōKō to Senryaku* (in Japanese).

[12] Kagono, Tadao, Kozo Yamada, and Kansai Productivity Center, *Nihon Kigyō no Shin-Jigyō Kaihatsu Taisei* (The System of Japanese Firms for Developing New Business) (in Japanese), Yuhikaku, 1999.

[13] Lang, Larry H. P. and Rene M. Stulz, "Tobin's q, Corporate Diversification, and Firm Performance," *Journal of Political Economy,* 102(6), 1248-80, 1994.

[14] MacDonald, James M., "R&D and the Directions of Diversification," *Review of Economics and Statistics,* 67(4), 583-90, 1985.

[15] Mowery, David C., Joanne E. Oxley, and Brian S. Silverman, "Technological Overlap and Interfirm Cooperation: Implications for the Resource-Based View of the Firm," *Research Policy,* 27(5), 507-23, 1998.

[16] Okada, Yosuke and Shigeru Asaba, "The Patent System and R&D in Japan," in A. Goto ed., *Innovation in Japan*, Oxford University Press, 1997.

[17] Panzar, John C. and Robert D. Willig, "Economies of Scope," *American Economic Review,* 71(2), 268-72, 1981.

[18] Science and Technology Agency, *White Paper on Science and Technology.*

[19] Statistics Bureau, *Report on the Survey of Research and Development*, the 1996 edition, 1998.

[20] Teece, David J., "Economies of Scope and the Scope of Enterprise," *Journal of Economics Behavior and Organization*, 1, 223-47, 1980.