COE-RES Discussion Paper Series
Center of Excellence Project
The Normative Evaluation and Social Choice of
Contemporary Economic Systems

Graduate School of Economics and Institute of Economic Research
Hitotsubashi University

COE/RES Discussion Paper Series, No.17
January 15, 2004

R&D Boundaries of the Firm: An Estimation of the Double-Hurdle Model on Commissioned R&D, Joint R&D, and Licensing in Japan

Kenta Nakamura
(Graduate School of Economics, Hitotsubashi University)
Hiroyuki Odagiri
(Graduate School of Economics, Hitotsubashi University)

Naka 2-1, Kunitachi, Tokyo 186-8603, Japan
Phone: +81-42-580-8350   Fax: +81-42-580-8351
URL: http://wakame.econ.hit-u.ac.jp/~koho/1intro/COE/index.htm
E-mail: COE-RES@econ.hit-u.ac.jp

R&D Boundaries of the Firm:

An Estimation of the Double-Hurdle Model on Commissioned R&D, Joint R&D, and Licensing in Japan

Kenta Nakamura* and Hiroyuki Odagiri†

January 2004

This is a shortened and slightly revised version of "Determinants of R&D Boundaries of the Firm: An Empirical Study of Commissioned R&D, Joint R&D, and Licensing with Japanese Company Data," Discussion Paper No. 32, National Institute of Science and Technology Policy (NISTEP). We thank NISTEP for allowing us to reproduce the paper in this manner. This study is also partly based on the research conducted as a part of the 21st Century Center-of-Excellence Project on "the Normative Evaluation and Social Choice of Contemporary Economic Systems" and we are grateful to the financial support from the Ministry of Education, Culture, Sports, Science and Technology of Japan. Any opinion expressed in this paper is the authors' alone and does not represent any of the institutions the authors are affiliated with. An earlier version was presented at Hitotsubashi University and the 2003 Spring Meeting of Japanese Economic Association. We thank R. Wakasugi and other participants for helpful comments.

*Ph.D. candidate, Graduate School of Economics, Hitotsubashi University, Rsearch Fellow, the Japan Society for the Promotion of Science

†Professor, Graduate School of Economics, Hitotsubashi University, Director of Research, National Institute of Science and Technology Policy

Abstract

This paper studies the determinants of R&D boundaries of the firm, namely, the firm's choice between performing R&D in-house *versus* procuring it from outside. We separate three modes of *procured R&D* – commissioned R&D, joint R&D, and technology acquisitions (i.e., licensing-in) – and, using the data of about 14,000 manufacturing firms in Japan, estimate the determinants of each mode. Two novelties are incorporated in this analysis. First, because the majority of the sample firms do not perform any R&D activity at all, we estimate a modified double-hurdle model in which the first hurdle determines whether the firm performs any R&D at all and the second hurdle determines whether (and how much) it performs each mode of procured R&D. Second, we employ both firm variables and industry variables (weighted with the firm's sales composition) to test the two major theories of the boundary of the firm, that is, the transaction cost theory and the capability theory. The results generally support these two theories: the estimated positive effects of firm size, in-house R&D intensity, diversification, and vertical integration support the hypothesis that capability is needed for procured R&D, while the estimated positive effect of the index of appropriability with patents supports the hypothesis that appropriability reduces transaction costs.

JEL Classification Code: D23, L22, O31

Keywords:

Firm Boundaries, Commissioned R&D, Joint R&D, Licensing, Transaction Cost, Capability

1 Introduction

Traditionally, the issue of the boundary of the firm has been discussed in relation to make-or-buy decisions in a vertical chain of production. How much supply of materials and parts is (and should be) integrated has been at the center of both theoretical and empirical studies on the boundary of the firm.

This issue has become a critical decision in the firm's research and development (R&D) strategy as well. With technologies becoming more science-based and complex, and with competition becoming more intensive on a global scale, it is now difficult for any firm to develop all the technologies by themselves. More and more, firms depend on scientific knowledge generated in universities, technologies acquired from other firms, and alliances with other firms, universities, and research institutes. This tendency is particularly strong with such high-tech industries as pharmaceuticals, chemicals, electronics, and automobile (e.g., Hagedoorn, 1993, 2002). In these and other industries, how much R&D should be performed in-house and how much R&D should be procured from outside have become one of the central strategic decisions in R&D management.

Following the pioneering work of Teece (1986), a number of studies have investigated the determinants of R&D boundaries of the firm. The results have been mixed. For example, regarding the effect of in-house R&D or patents (or its intensity) on variables of R&D alliances or collaborations, Arora and Gambardella (1990, 1994) and Veugelers (1997) found a significant positive effect, suggesting complementarity between internal R&D and external R&D. However, Kleinknecht and Reijnen (1992) found the effect to be insignificant except for collaborations with foreign research institutes. Furthermore Rocha (1999) found the effect of R&D intensity on the ratio of joint patent applications to be negative, though insignificant.

In this paper, we aim to analyze such determinants, using a comprehensive data of manufacturing firms in Japan. Our study is unique in several respects. First, whereas most of the existing studies were confined to particular industries, such as biotechnology, or to a small number of firms, we use a large data set of approximately 14,000 Japanese firms that cover all the manufacturing industries. Second, we apply a double-hurdle model (Cragg, 1971) in order to investigate the two-step decisions, that is, whether or not the firm should perform any R&D at all and, if it should, how much it should spend for external R&D resources. Third, to take into account the fact that firms procure external R&D resources through diverse means, we separate commissioned R&D, joint R&D, and technology acquisitions. Fourth, as the determinants of these means of procured R&D, we examine not only firm characteristics, such as firm size, R&D intensity, diversification, vertical integration, ownership, and cash flow, but also technological and industrial characteristics. The latter are represented by the indexes of appropriability by patents, the extent of information flow, and innovation speed, which were derived from the questionnaire study of Goto and Nagata (1996). These variables, we will argue, are closely related to the two major theories of the boundary of the firm – the transaction cost theory of Williamson (1975, 1985) and the capability theory of Penrose (1959), Nelson and Winter (1982), and others.

The remainder of the paper is organized as follows. In the next section, we define three major forms of procured R&D – commissioned R&D, joint R&D, and technology acquisitions – and discuss the fundamental differences among them. In Section 3, we discuss the above-mentioned two theories of the R&D boundary of the firm. In Section 4, we explain the data source and the variables on procured R&D to be used as the dependent variables in our regression. In Section 5, we explain our double-hurdle model. In Section 6, we explain the independent variables, together with the hypotheses on their effects to procured R&D and, in Section 7, we present the estimation results. Finally, Section 8 summarizes our findings and concludes the paper.

2 In-house R&D versus Procured R&D

'In-house R&D' (IRD) refers to the activity of the firm whereby it sets up and fulfills a research project within itself, by employing necessary resources, such as researchers, research materials, and equipment. Alternatively, it can procure a part of the R&D activity from outside. In this paper, we investigate three modes of what we call 'procured R&D'; commissioned R&D, joint R&D, and technology acquisitions. They differ in important ways[1].

[1]Some authors (e.g., Audretsch et al., 1996, and Bönte, 2003) used the terms, 'internal R&D' and 'external R&D', in place of 'in-house R&D' and 'procured R&D'. We prefer the term, 'procured R&D', to 'external R&D' because external R&D can be worthless unless the firm makes deliberate efforts to *procure* it, by making sacrifices in the form of payments or the allocation of its human and other resources. Although the firm may also enjoy the benefit of external R&D without payment through *spillovers*, they are not the subject of the present study: for spillovers, see Griliches (1992).

Technology acquisitions (TA) are the purchase of technologies through, most commonly, licensing-in of patents. Non-patented technologies, such as knowhow and consultancy, may be also purchased. A salient feature of TA is that the technology to be traded has been already invented by the time the contract is made; therefore, uncertainty is lower as to the outcome from the contract and the object of the contract can be more clearly defined. That is, the 'predictability' of the outcome is higher and so is the 'definability' of the work to be conducted to fulfill the contract (Odagiri, 2003)[2].

Both predictability and definability are lower in *commissioned R&D* (CRD) and *joint R&D* (JRD), or *R&D alliances* as they are collectively called, because they need to be contracted before the actual R&D process is to be started. CRD and JRD differ with regard to the way the participants are involved and the outcomes are shared between them. In JRD, the R&D work is to be shared by the participants, each of them contributing R&D funds and/or R&D personnel whereas, in CRD, the R&D work is basically the responsibility of the commissioned party. The commissioning party provides R&D funds as stipulated by the CRD contract and usually receives the entire right to the R&D outcome.

Incomplete definability and low predictability can cause information asymmetry among the partners and, thereby, moral hazards. For instance, in CRD, the commissioned party (say, Firm B) may realize during the course of the commissioned research that the chance of coming up with an invention is actually much lower than was predicted at the time of the contract. However, if the commissioning party (Firm A) is unaware of this fact, B may be tempted to conceal it so that it can keep receiving the research fund from A. In JRD, a free-rider problem may occur because each participating firm has an incentive to minimize its contribution and yet to receive information fully from the project and the partners. Thus, neither CRD nor JRD is free from transaction costs as broadly defined. Yet, both are useful in the exploitation of capabilities held by other partners, which provide a strong motivation for procured R&D. It is thus suggested that we need to inquire into the issues of transaction costs and capabilities to study the R&D boundaries of the firm.

[2]Technology may be also acquired through acquisition of patent-holding companies (Huber, 1991; Ahuja and Katila, 2001). In the US, M&As for this purpose are common particularly in IT industries (Inkpen, Sundaram, and Rockwood, 2000). In this paper, we do not discuss such M&As because, firstly, M&As may be motivated by organizational as well as technological purposes and, secondly, M&As are relatively infrequent in Japan (Odagiri, 1992).

3 Two Main Theories

3.1 The Transaction Cost Theory

Probably the best-known theory on the boundary of the firm is the *transaction cost theory*, advocated by Williamson (1975, 1985). Under uncertainty, complexity, asymmetric distribution of information, and bounded rationality, market transactions can be costly particularly because the participants may behave in an opportunistic fashion. Generally, these transaction costs make in-house activities more advantageous than market transactions. However, integrating the activities in-house can be also costly because of reduced competitive pressure from the market, influence costs, and other costs from integration. The firm, therefore, needs to determine the best allocation between in-house R&D and procured R&D with due consideration for the balance between integration costs and transaction costs.

This balance depends on many factors but here we focus on two. The first is the extent that required tangible and intangible assets are transaction-specific. If they are transaction-specific, few suppliers are willing to invest in them for fear of the buyer's hold-up. Thus, in the case of R&D, if the R&D project requires investment in specific equipment or other assets, it is more likely carried out in-house than being procured from outside.

The second is the extent of definability and enforceability of property rights. In the transaction of intangible assets such as technology, it is not easy to specify in the contract the range of technology to be transacted and each party would be tempted to interpret it in a way more favorable to them. In a commissioned R&D, for instance, the commissioned party (a firm, university, research institute, etc.) would try to limit the range of technology to be handed over to the commissioning party. The intellectual property rights system, such as the patent system, helps the parties to resolve this difficulty because the range of relevant technology is specified in a patent. This tendency is most applicable in the licensing of patents because, by referring to patent numbers, contracts can be well defined. Even in commissioned or joint R&D, one can more easily write and enforce contracts by stipulating that the outcome be patented and handed over from the commissioned party to the commissioning party or to be shared among the partners.

In reality, however, patents may not allow the inventor to appropriate profits fully from the invention and the extent of such appropriability is known to vary across industries (Cohen et al., 2000; Goto and Nagata, 1996; Levin et al., 1987). We will therefore investigate the effect of appropriability in our empirical analysis.

3.2 The Capability Theory

The second theory to explain the boundary of the firm is the *capability theory*, which originates from Penrose (1959) who stressed the importance of viewing the firm as a collection of physical and human resources, and was developed further by Nelson and Winter (1982) and others, and applied, for instance, to discuss the Japanese industrial development by Odagiri and Goto (1996). The theory has been also called the theory of a resource-based view of the firm (Wernerfelt, 1984), organizational capability (Chandler, 1990), dynamic capability (Teece, Pisano, and Shuen, 1997), or core competence (Prahalad and Hamel, 1990), with slightly different emphases and purposes.

It takes time and costs for the firm to create and enhance its tangible and intangible assets and hence its capabilities. The firm can of course develop its capabilities through investment and learning. Yet, the speed and direction of this development are constrained and influenced by not only the firm's social and economic environment but also the volume and composition of its current assets and the history of its development. As a result, the development is bound to be path-dependent.

The firm can fulfill a certain task cheaper and faster if it procures it from an outside party who possesses more of the necessary capabilities than when it conducts it within itself. In other words, the decision on the firm's boundary is dependent on the relative level of in-house capabilities versus outside capabilities.

However, one needs to note the dynamics of capabilities. If a certain activity is performed within the firm, it can learn from the experience and enhance its capabilities. Hence, even if the cost of doing so is higher in the short run, the expected long-run cost reduction may be large enough to offset the short-run cost. On the contrary, if the firm depends on outside resources, its capabilities will gradually become obsolete, causing the firm to lose not only the capability to perform the activity in-house but also the capability needed to evaluate the procured goods and services, monitor the suppliers, and bargain with potential suppliers and partners. In short, it will also lose its 'absorptive capacity'.

Therefore, it is indispensable for the firm to maintain a certain level of capability through in-house R&D[3]. It should not consider the relative merit of in-house R&D versus procured R&D merely from the comparison of current costs. The capability theory teaches us that it needs to take a dynamic and broad view in determining the boundary of the firm.

4 Data and the R&D Variables

We now proceed to our empirical analysis on the determinants of commissioned R&D, joint R&D, and technology acquisitions (i.e., licensing-in). We use unpublished firm-level data from the *Basic Survey of Business Structure and Activities* (hereafter BSA): see Appendix 1 for the detail of this data. BSA is unique in that it asked the firms to provide information on not only their in-house R&D but also the above-mentioned three modes of procured R&D. However, the question items vary slightly from year to year and all of the three were included only in the 1998 BSA report, which covers the data of 1997 accounting year (April 1997 to March 1998 for most firms). Thus, the data for all the variables in the following are taken from this 1998 report, except those variables representing industrial and technological characteristics to be discussed later.

Our sample consists of 14,070 manufacturing firms in the survey. Among these, 6,281 (44.6%) reported to have expended for in-house R&D and, including commissioned and joint R&D and licensing-in, 6,648 (47.2%) reported to have made some form of R&D activities. We will use this information in the following analyses.

There are four R&D-related variables:

IRD_i = the amount (in million yen) of R&D expenditures firm i used in-house.

CRD_i = the amount (in million yen) of R&D expenditures that firm i commissioned to any of the organizations outside the firm.

JRD_i = the number of partners (besides the firm in question) in joint R&D projects that firm i participated[4].

---

[3] In fact, in the questionnaire study, we conducted earlier (Odagiri, Koga, and Nakamura, 2002), many Japanese bio firms raised the full utilization of internal resources and the need to nurture them as the major reasons for not making R&D alliances even when there are opportunities for such alliances.

[4] BSA requires that, in any joint R&D, the participants share R&D activities, share the outcomes, and exchange a contract. It is thus separate from commissioned R&D or subcontracting.

TA_i = the amount of payment firm i made for licensing-in of patents (whether the licensing contract was made during the year or earlier)[5].

CRD_i includes R&D expenditures commissioned to the firm's affiliates (of which the firm owns more than 20 percent of the share), its majority-controlling parent (who owns more than 50 percent of the firm's share), other firms, universities, and government research institutes, inside or outside of Japan. BSA reports the proportion of CRD_i to its affiliates and the parent; hence, we can calculate the following two:

$CRDI_i$ = R&D expenditures that firm i commissioned to its affiliates or the parent (to be called 'in-group commissioned R&D').

$CRDN_i$ = R&D expenditures that firm i commissioned to any party outside of the group ('non-group commissioned R&D')[6],

where, of course, $CRD_i = CRDI_i + CRDN_i$.

In-group commissioned R&D, one may hypothesize, is a form of *quasi*-internal R&D rather than procured R&D and, hence, the determinants can differ from those for non-group commissioned R&D. We will test this hypothesis.

Table 1 gives the descriptive statistics for the variables. Among the 14,070 sample firms, 1,315 (9%) commissioned R&D. Of the total commissioned R&D, 79 percent went to non-group. Only 296 (23%) of the 1,315 firms commissioned R&D to in-group. Among these firms, however, nearly two thirds of the commissioned R&D went to in-group (not reported in the table). These facts imply two things. First, more than three quarters of the firms making commissioned R&D are commissioning only to non-group and more than three quarters of commissioned R&D expenditures went to non-group. However, there are a number of firms who are heavily commissioning R&D to their affiliates or the parents. Presumably, many of these firms have hived-off their R&D departments as separate subsidiaries[7], or they may have a specialist R&D company within the group.

The proportion of firms conducting joint R&D or technology acquisition is smaller than that of commissioned R&D at, respectively, 6.8 percent and 5.9 percent. However,

---

[5] TA_i also includes payments for knowhows that accompany patents.

[6] $CRDN_i$ still includes R&D commissioned to firms for which firm i owns less than 20 percent of the share.

[7] That many Japanese firms hive-off some of their divisions has been discussed in Odagiri (1992).

as a proportion to the 6,648 firms with positive in-house and/or procured R&D, the percentages (not reported in the table) increase to 20 (CRD_i), 14 (JRD_i), and 13 (TA_i).

5 The Double-Hurdle Model

CRD_i, JRD_i, and TA_i equal to zero for more than ninety percent of the sample. As is well known, when the dependent variable is constrained to be non-negative and takes the value of zero in a large portion of sample, the OLS estimates are biased and the common research strategy is to apply a left-censored (at 0) Tobit model (Tobin, 1958). We basically follow this strategy. In addition, since we know that a number of firms not only have not made the procured R&D but, actually, have not made any R&D activity at all inside or outside of the firm, we wish to utilize this information.

Put differently, we may approach the issue of the R&D boundaries of the firm as a sequence of R&D decisions. Firstly, should the firm expend for any R&D activity at all? If the answer is yes, then, secondly, how much should it expend in-house and how much by procurement and, if procurement, how much should it expend for each mode of procured R&D?

Let us, for the moment, consider only commissioned R&D as a means of procured R&D. Then, one can estimate a two-stage model. The first stage determines if $IRD_i + CRD_i > 0$. If it is, then the second stage determines if $CRD_i > 0$ and, if it is, how much expenditures should be made for it. These two stages have to be estimated jointly. For this purpose, we apply a double-hurdle model, which was originally suggested by Cragg (1971) as a generalized form of the Tobit model. In Cragg's original model, two hurdles refer to the following: "First, a positive amount has to be desired. Second, favorable circumstances have to arise for the positive desire to be carried out" (Cragg, 1971, p. 831). In our case, the first hurdle is to have positive $IRD_i + CRD_i$ or, equivalently, to have $ICRDD_i = 1$ where $ICRDD_i$ is a dummy variable that equals one if and only if $IRD_i + CRD_i > 0$, and the second hurdle is to have a positive CRD_i. Figure 1 shows the flow chart for this model.

These two equations, that is, the first-hurdle equation (that determines $ICRDD_i$) and the second-hurdle equation (that determines CRD_i), are jointly estimated by means of the maximum likelihood method (see Appendix 2 for the derivation of the likelihood function). In this maximum likelihood estimation, the maximand includes the terms related to the second-hurdle equation only for samples with $ICRDD_i = 1$. Also, the covariance between the residuals of the two equations are taken into account .

When other modes of procured R&D, namely, joint R&D and technology acquisitions, need to be considered, it is ideal to have the choice among the three modes simultaneously incorporated into the second hurdle. Unfortunately, we cannot do so for want of a 'total' R&D variable, in addition to the complexity of maximum likelihood computation. As mentioned above, IRD_i, CRD_i, and TA_i are in monetary units but JRD_i is the number of participants. Hence, the sum of these numbers is meaningless. Besides, TA_i is the amount the firm paid for acquired technologies during the year. Since this payment is usually composed of fixed initial payment and running royalty, with the latter being commonly determined as a fixed percentage of sales, the amount can fluctuate violently from year to year. Also, the firm may keep paying for many years after the actual technology acquisition took place. Consequently, adding TA_i to IRD_i and CRD_i need not provide a good measure of the current R&D activity.

For these reasons, we estimate the double-hurdle model separately for each mode of procured R&D. For the JRD_i and TA_i equations, however, a new variable RDD_i is used in place of $ICRDD_i$ in the first-hurdle, where RDD_i equals one if any of IRD_i, CRD_i, JRD_i, and TA_i is positive and zero otherwise (i.e., when none of these four is non-zero), in order to include several firms who have not expended for in-house or commissioned R&D and yet participated in joint R&D or made payments for technology acquisitions.

6 Independent Variables and Hypotheses

There are two types of independent variables, those for firm characteristics and those for industrial characteristics. We now discuss them together with our hypothesis on the signs of the coefficients. The definitions of independent variables and their basic statistics are shown in Table 2.

6.1 Firm Characteristics

R&D intensity

This variable appears as an explanatory variable only in the second hurdle. Cohen and Levinthal (1989) argued that, with R&D, the firm can enhance its absorptive capacity that is needed to exploit external knowledge efficiently. Also, a more R&D-intensive firm will be more alert to outside R&D opportunities and will have more knowledge on potential alliance partners and the technologies to license. It is thus hypothesized that R&D intensity has a positive impact on procured R&D, where R&D intensity ($RDINT_i$) is defined by the ratio of in-house R&D expenditure to sales. Earlier empirical results of Arora and Gambardella (1990, 1994) and Veugelers (1997), and more recently, Bayona et al. (2001) support this hypothesis[8].

Some authors, on the contrary, suggested that in-house R&D and procured R&D are substitutes because the firm can fulfill a certain R&D task either by making it by itself or by commissioning it from outside. Pisano (1990), for instance, found that biotechnology firms that have accumulated technical knowledge in-house are less likely to rely on external knowledge. If this were the case, then, the firm with active in-house R&D would rather not procure R&D from outside and, consequently, we should expect a negative coefficient for $RDINT_i$. As will be shown later, however, our estimated coefficient is positive. Hence, in the sense that procured R&D increases with in-house R&D, the two appear complementary than substituting, although we have not rigorously tested the causality as in Colombo and Garrone (1996).

Size

The relationship between firm size and R&D investment has been studied by many, often in conjunction with Schumpeter (1942). Although their results disagree as to whether a larger firm expends for R&D more than proportionally, they agree that “the

[8] A good question is whether only in-house R&D contributes to absorptive capacity or procured R&D also contributes. If it is the process of R&D being made within the firm that contributes to the formation of absorptive capacity, then, in-house R&D intensity ($RDINT_i$) is more likely to matter. If, on the other hand, invented technologies, whether invented in-house or not, are the sources of absorptive capacity, then, total R&D intensity ($TRDINT_i$) is more likely to matter. Actually, this choice hardly matters because, on average, in-house R&D expenditure accounts for 94 percent of total R&D expenditure and the correlation coefficient between $RDINT_i$ and $TRDINT_i$ reaches 0.98. We confirmed this fact by using $RDINT_i$ and $TRDINT_i$ as alternative explanatory variables and obtaining basically the same estimation results. It may be also argued that R&D stock, that is, an accumulated value of R&D expenditures with obsolescence taken into account, is a more accurate measure of the firm’s technological capabilities than R&D expenditure of a single year. The major reason that we did not use R&D stock is the lack of continuous time-series R&D data for many of the sample firms. Rather than reducing the sample size by restricting the sample to those for which the time-series R&D data is available, we decided to use R&D flow data and maintain the sample size as large as possible.

likelihood of a firm reporting positive R&D effort rises with firm size" (Cohen and Klepper, 1996). Therefore, we hypothesize that the effect of size on the dummy variable, $ICRDD_i$ or RDD_i, in the first hurdle is positive.

In the second hurdle, it is difficult to predict the effect of firm size on the frequency of procured R&D. On the one hand, as Granstrand et al. (1997) and Patel and Pavitt (1997) emphasized, large firms may be technologically diversified and well-endowed, thereby having a better knowledge and better access to potential external partners. On the other, they may be able to achieve scale and scope economies with their in-house R&D, thus feeling a lesser need for procuring R&D resources. For instance, Veugelers (1997) and Veugelers and Cassiman (1999) found a negative relationship between firm size and R&D cooperation to argue that, because small firms can neither undertake a large-scale research nor undertake a number of research projects simultaneously, economies of scale and scope cannot be achieved, making R&D alliances more attractive to these firms.

A cursory look at our data suggests that the first hypothesis is more likely to hold; for instance, the proportion of firms making commissioned R&D is 20 percent among the firms with 300 employees or more but only 6 percent among smaller firms. We thus predict a positive relationship between $LSALE_i$, the natural logarithm of sales (including oversea sales), and procured R&D. This positive relationship may also come from the fact that CRD_i, JRD_i, and TA_i are all measured as numbers, such as expenditures and the number of partners, and not ratios. Hence, a positive association between $LSALE_i$ and these variables need not mean that a larger firm expend on procured R&D more than proportionately[9].

Vertical Integration

The extent of vertical integration (VI_i) is measured by the ratio of value-added to sales, on the presumption that a less integrated firm will expend a larger proportion of sales in the procurement of parts and components, thus having a smaller value-

[9] We found, however, that the estimated coefficients remain positive and significant (except that for JRD_i) even when CRD_i, JRD_i, and TA_i are measured as intensities (i.e., as ratios to sales): see Nakamura and Odagiri (2003). Therefore, commissioned R&D and technology acquisitions in fact tend to increase more than proportionately with size.

added/sales ratio.

A more vertically integrated firm, one may hypothesize, should feel a stronger need for R&D because it has to maintain technological competence in all stages of the vertical chain. Then the effect of VI_i on the probability of conducting R&D would be positive in the first hurdle. The effect of double-counting has to be also taken into account. Because costs of R&D personnel and capital are included in the firm's value added, an R&D-performing firm is likely to have a higher value-added/sales ratio. Again, we would expect a positive effect of VI_i in the first hurdle.

Such double-counting is unlikely to occur in the second hurdle because commissioned R&D will be carried out within the commissioned party utilizing its employees and capital. It may occur in the case of joint R&D as long as the firm's researchers participate in the joint R&D but, since the expenditure on joint R&D is tiny in comparison to in-house R&D, the effect of double-counting must be too small to change the sign of the coefficient of VI_i in the second hurdle.

Other effects of VI_i on procured R&D can be mixed. On the one hand, VI_i may be interpreted as a proxy variable indicating that the firm's environment is more favorable to integration than market transactions. Transaction costs may be higher because of higher asset specificity, a larger sunk cost, or a larger probability of hold-ups. Ideally, a direct measure of these costs is more desirable than VI_i. For instance, Ulset (1996) found that potential sunk costs in R&D are positively related to vertical integration (i.e., in-sourcing) of R&D by using R&D project-level data for the IT industry. Such data is unavailable in Japan and, assuming that VI_i is positively correlated with the extent of transaction costs, we may hypothesize that firms with higher VI_i are more likely to undertake R&D internally.

On the other hand, from the viewpoint of absorptive capacity, a vertically integrated firm may have a higher capability to perform alliances and to absorb their outcomes, owing to its experience of having had business relations with firms of vertically diverse activities and culture, its understanding of technologies at vertically different stages, and its wider knowledge of potential partners. Then, firms with higher VI_i would be more likely to engage in procured R&D.

In consequence, the transaction cost theory and the capability theory would predict the sign of the coefficient of VI_i on CRD_i, JRD_i or TA_i differently.

Diversification

The extent of diversification, DIV_i, is measured by one minus the square root of Herfindahl index (the sum of the squares of the proportions of the firm's sales of three-digit products). In the first hurdle, its effect on the probability of R&D is expected to be positive. Nelson has earlier argued in his now classic paper (Nelson, 1959) that the outcome of R&D is inherently uncertain and this uncertainty makes diversified firms more advantageous in the commercialization of invented technologies. Hence, a more diversified firm will likely undertake R&D with a higher probability, although empirical results on the relationship between diversification and the level of R&D investment are not unanimous: McEachern and Romeo (1978) and Jovanovic (1993) found a positive correlation but a series of studies by Hoskisson and others (e.g., Hoskisson et al., 1993) found a negative relation.

In the second hurdle, the capability theory can suggest either a positive or negative effect of DIV_i. On the one hand, a more diversified firm may have a lesser need to depend on outside partners in pursuing R&D in non-core fields, suggesting a negative effect of DIV_i on procured R&D. On the other hand, a more diversified firm may have a broader absorptive capacity, which helps the firm to procure R&D efficiently, suggesting a positive effect of DIV_i. For instance, by applying Nelson's argument to procured R&D, we may say that a more diversified firm should be able to utilize the uncertain outcome from joint R&D more effectively. Assuming that the latter effect is dominant, we hypothesize that DIV_i will have a positive coefficient in the second hurdle. This assumption, we trust, is plausible because our measures of procured R&D are the expenditures or numbers and not intensity, and because, even though the first argument suggests that a more diversified firm will procure proportionally fewer of its R&D from outside, it need not imply that such a firm will procure a smaller amount of R&D from outside.

Cash Flow

R&D investment is usually riskier than other forms of investment. As a result, under information asymmetries between firms and investors (or lenders), the firm can more easily invest in R&D when it has an abundant cash flow. Thus, Himmelberg and Petersen (1994) found in small high-tech firms that R&D expenditure is sensitive to cash flow. Also, Goto et al. (2002) found that the ratio of cash flow to assets has a positive effect on the R&D-asset ratio in both large and small Japanese manufacturing firms. We thus expect that a firm with more abundant cash flow is more likely to perform R&D; that is, the ratio of cash flow to sales, CFS_i, will have a positive coefficient in the first hurdle.

The same argument is applicable to commissioned and joint R&D provided that, because of the low predictability and definability as discussed in Section 2, such investment is riskier than investment in tangible assets and is also less suitable for collateral. This argument may be less applicable to technology acquisitions (TA_i) because, usually, technologies to be licensed have been already invented and hence their predictability and definability need not be low. We thus predict CFS_i to have a positive impact on CRD_i and JRD_i but not necessarily on TA_i.

One may alternatively argue that joint R&D is preferred to in-house R&D when the firm wishes to share the cost and risk of the R&D project with the partners. According to this hypothesis of cost- and risk-sharing motivations for joint R&D, the firm may opt for joint R&D when it is short of cash flow, implying a negative coefficient of CFS_i on JRD_i and, perhaps to a lesser extent, CRD_i. Kleinknecht and Reijnen's (1992) finding supported this argument among joint research between domestic firms.

Parent Control

PC_i is a dummy variable indicating the presence of a parent company. It equals one if and only if the firm is owned a majority share by its parent company. Wakasugi (1999) found a higher R&D intensity among firms owning subsidiaries and argued that there is a division of labor between parent companies and subsidiaries, with the former conducting most of R&D. According to this argument, if the firm is a subsidiary (and unless it is a subsidiary established for the purpose of R&D like Honda's Honda R&D Co., Ltd.), it is presumably less likely to undertake R&D activity at all (in the first hurdle) and less likely to undertake procured R&D (in the second hurdle) because the decisions on alliances and licensing will be also concentrated in the parent company. One may therefore hypothesize that PC_i should have a negative coefficient in both the first and second hurdles.

However, regarding $CRDI_i$ (in-group commission of R&D), the discussion is more complex because subsidiaries may commission R&D to the parents or other in-group firms. In fact, we found that 46 percent of the firms commissioning R&D in-group are parent-controlled (i.e., $PC_i = 1$) and the percentage is significantly higher than that of firms not commissioning R&D in-group. Hence, the aforementioned hypothesis of a negative effect of PC_i is unlikely to apply to $CRDI_i$ and we would instead expect a positive effect.

6.2 Industrial and Technological Characteristics

R&D strategies depend not only on firm characteristics but also on industrial and technological characteristics. In this paper, we employ three variables representing these characteristics – appropriability, information flow, and innovation speed. They are derived from Goto and Nagata (1996), who sent questionnaires to 1,219 Japanese R&D-performing manufacturing firms with capitalization over one billion yen, with 643 responses, and reported industrial averages in their paper[10]. These variables, therefore, are constructed from the responses of big firms and there remains a possibility that they do not accurately describe the environment of smaller firms that occupy the majority of our sample.

Since our empirical study is made at the firm level, we may use either the data of each firm's responses or industrial averages. One may argue that firm-level data are more appropriate because even if firms belong to the same industry, the environment can be heterogeneous among firms. One may, on the other hand, argue that firm-level data are susceptible to the firms' subjective evaluation and prefer industry-level data, which is less dependent on each firm's individual opinion. Partly for this reason and partly for a practical reason that the sample size of Goto and Nagata's survey is much smaller than that of BSA and hence firm specific data are not available to all the BSA sample firms, we use industrial data. However, to take into account the effects of inter-firm differences, we generated firm-specific variables by computing the weighted averages of industrial data (basically at the three-digit Japanese SIC level) with the sales composition of each firm as the weights. Thus, even though these variables are

[10] This survey was conducted in 1994 in conjunction with the Carnegie Mellon survey, which is an expansion and update of the Yale Survey (Levin et al., 1987): see Cohen, Goto, Nagata, Nelson, and Walsh (2002) and Cohen, Nelson, and Walsh (2000, 2002). Some of the data are not reported in the report and we thank A. Nagata for providing unpublished data for us.

for industrial characteristics, they are firm-level variables and, we believe, reflect the technological and market environment of each firm accurately.

These variables are calculated as weighted averages in one more sense. As regards appropriability and innovation speed (to be explained presently), Goto and Nagata asked the companies to provide responses for each of product innovation and process innovation. Hence, we weighted these responses with the ratio of R&D spending on product innovation versus that on process innovations[11].

Three such firm-specific variables of industrial and technological characteristics will be now explained.

Appropriability

Goto and Nagata asked the respondents the percentage of their projects in the past three years for which each of the following eight means of protecting competitive advantages from innovations was effective – secrecy, patents, other legal protections, lead time, complementary sales and service, complementary manufacturing facilities and knowhow, complexity of production and product design, and others. They then reported the industrial averages of these percentages for each means.

Among these, we concentrate on the role of patents because, as discussed in Section 3, the effectiveness of patent protection is one of the major determinants of transaction costs and also because some of the other means listed above are not purely exogenous to the firms[12]. We define $APPRO_i$ as the extent of appropriability by patents. As shown in Table 2, the mean value of $APPRO_i$ equals 0.322, implying that, on average, the firms reported that patents were effective in about a third of the projects. As in the US (Levin et al., 1987; Cohen et al., 2000), it is highest among pharmaceutical firms:

[11]On the average of all manufacturing industries, 80.9 percent of the R&D cost was for product innovations and 14.7 percent for process innovations, with the rest being for miscellaneous category that was ignored in our analysis.

[12]One may alternatively use the maximum among the seven means (excluding 'others') as an appropriability variable. Such a variable may be appropriate as a determinant of R&D intensity because, whatever the means of appropriation, a higher appropriability is expected to stimulate R&D investment (Goto et al., 2002). However, in procured R&D, it is important that such appropriability is legally secured and patents are the most effective means for this purpose. Besides, in our preliminary regressions, we found that the maximum-based appropriability variable has a poorer explanatory power than the patent-based appropriability variable, presumably because some industries replied with unrealistically high numbers (e.g., 100 percent) for some of the non-patent means.

see Appendix Table 3 in which pharmaceuticals are included in chemicals.

As discussed earlier, according to the transaction cost theory, stronger property rights would enable the firms to engage in inter-firm relations with lower transaction costs; hence, $APPRO_i$ is expected to have a positive impact on procured R&D in the second hurdle. By preventing free-riding, they would also increase the private returns to R&D investment; hence, the incentive for R&D must be enhanced and $APPRO_i$ must have a positive coefficient in the first hurdle as well.

Our prediction of a positive effect of $APPRO_i$ in the second hurdle agrees with the theoretical prediction of Arora and Fosfuri (2003) on licensing. Empically, Gans et al. (2000) confirmed this prediction by showing that a stronger intellectual property protection increases the probability of cooperation between start-ups and incumbents. Hernán et al. (2003), on the other hand, argued that firms in sectors with stronger patent rights do not need to rely on research joint ventures to internalize spillovers and empirically confirmed this prediction. Similarly, Cassiman and Vergelers (2002) found a negative, though insignificant, effect of legal protection (including protection by patents) on the probability of R&D cooperation. If this argument is correct, then we should expect a negative coefficient of $APPRO_i$ on JRD_i. However, if joint R&D is formed to take advantages of the technological capabilities of the partners as the capability theory implies and not to internalize spillovers among the partners, then the transaction-cost reducing effect of $APPRO_i$ must be more prominent. With this view, we predict a positive coefficient.

Information flow

Goto and Nagata asked if each of twelve probable information sources was conducive to the 'proposal of a new project' or the 'completion of an existing project' in the past three years. They reported the percentage of firms who replied affirmatively for each source, each industry, and each of new project proposal versus project completion. Among the 12 sources, 'universities', 'public laboratories', and 'academic associations, etc.' will be hereafter called 'scientific sources', while 'suppliers with share ownership relationship', 'suppliers without share ownership relationship', 'customers', and 'competitors' will be called 'transaction-based sources'[13].

[13]The remaining five are 'joint ventures', 'consultants', 'other external sources', 'other R&D departments within the firm', and 'manufacturing department within the firm'. These were ignored because

We averaged among the three scientific sources and between new project proposal and project completion to get an industrial value of information flow from scientific sources, and then computed the firm-level value as an weighted average of industrial values in the manner discussed earlier. We call this variable 'scientific information flow' and denote it by $FLOWS_i$. We similarly calculated 'transaction-based information flow' from the average among the four transaction-based sources and denote it by $FLOWT_i$[14].

As shown in Table 2, $FLOWT_i$ is on average larger than $FLOWS_i$, suggesting that information is more frequent from transaction-based sources than from scientific sources although in a few industries, particularly pharmaceuticals, information flow from scientific sources overwhelms.

With a larger information flow, there must be a larger technological opportunity, which tends to stimulate in-house R&D as confirmed by Cohen and Levinthal, (1990) and Goto et al. (2002). This positive effect will take place not only because the larger information flow provides more opportunities for firms to innovate but also because firms need to enhance their absorptive capacity so as to take advantages of information flow (Cohen and Levinthal, 1989). We therefore expect $FLOWS_i$ and $FLOWT_i$ to have positive coefficients in the first hurdle.

Its impact on procured R&D can be more complex. On the one hand, a larger information flow implies that more 'seeds' are available outside of the firm, prompting the firm to invest in procured R&D to internalize these seeds. For instance, it may be easier to find partners with high technological competence and the firm may be tempted to take advantages of it through commissioned or joint R&D (Cassiman and Veugelers, 2002). On the other hand, such information flow may occur through knowledge spillovers, for instance, through published papers and human contacts, without the firm paying for it. Then the firm would have a lesser need for commissioned or joint R&D and licensing. Since these two types of information flow coexist in the Goto-Nagata survey, we cannot

joint venture is more likely a result than a cause of procured R&D, the impact of consultants and other external sources is difficult to predict, and intra-firm sources are unlikely to affect procurement of R&D from outside.

[14]One may alternatively define the two information flow variables based on the maximum among the relevant set of sources on the assumption that, if any one of the sources is very useful, the firm will attempt to take advantage of this source through R&D. However, similarly to the discussion in footnote 12 above, we found that these maximum-based variables are susceptible to a few extreme values (particularly in industries with small numbers of respondents); hence, we only report the results with mean-based variables.

a priori determine which of these two effects dominate. Consequently, the coefficients of $FLOWS_i$ and $FLOWT_i$ may be positive or negative in the second hurdle.

Speed of innovation

Goto and Nagata asked the firms to evaluate on a 5-point Likert scale how fast product innovation or process innovation took place in the past ten years in the industry. Based on the industrial average of this measure of 'innovation speed', we calculated $SPEED_i$ again as a weighted average.

When technological change is rapid, competition in terms of new products and/or new process is keen and the firm is under a strong pressure to innovate. Therefore, firms competing in such markets are more likely to undertake R&D, suggesting a positive coefficient in the first hurdle. Yet, even in high-tech industries, there are also firms who are not competing on the basis of technological strength but on the basis of low cost and/or non-R&D-based knowhow. These firms are often subcontractors or low-cost suppliers to large-scale assemblers and may have opted for non-R&D-based competition in fear of escalating R&D costs. That is, there may be a divide between R&D-intensive firms and non-R&D-based firms, and an accelerated speed of innovation may actually tilt this divide towards non-R&D-performing firms. If this is the case, we may have a negative coefficient for $SPEED_i$ in the first hurdle.

The effect on procured R&D can be also complex. Again, firms in an industry with fast innovation can be under a stronger pressure to come up with new products and/or new processes and, to fulfil this purpose, they may be more willing to utilize external resources. That is, they may be inclined to commission R&D or take part in joint R&D to speed innovation up, or to acquire new technology at once through licensing. Technological inter-firm partnerships have been in fact found in such high-tech industries as computers, semi-conductors, and biotechnology (Freeman, 1991; Hagedoorn, 2002; Powell et al., 1996).

However, there may not be a sufficient number of firms in such industries that are good enough to perform required commissioned R&D works. Although BSR does not give a precise definition of 'commissioned R&D', it states, in another part of the questionnaire, that 'commissioned production of products' is to have other firms manufacture or process finished products, semi-finished products, components, accessories, or materials by instructing them the specifications and standards. By analogy, respondents may have taken ‘commissioned R&D’ as including not only the commissioning of scientific discovery or development activities as described in Section 2 but also subcontracting of routine R&D-related works, say, data input, routine experiments and computation, and the maintenance of laboratory. Such subcontracting may be more prevalent in an industry with slower technological change as many of the works are standardized and many low-cost suppliers may provide such services. By contrast, in high-tech industries, commissioning of real and advanced R&D, as opposed to subcontracting, may be needed and yet only a small number of firms may have accumulated sufficient capabilities to perform such R&D. In consequence, the firm may have no choice but to perform it themselves.

In view of these possibilities, it is extremely difficult to predict the effect of $SPEED_i$ on procured R&D in the second hurdle of the model.

Table 3 summarizes our hypotheses on the signs of the coefficients in our double-hurdle model. It also includes $PRINT_i$ as an independent variable, which we will discuss later.

7 Empirical Results

7.1 The First Hurdle

We now present the results of our empirical analyses. Table 4 presents the estimated double-hurdle model when *ICRDD* and *CRD* (suppressing hereafter the subscript *i*) are taken as the dependent variables of the first hurdle and the second hurdle, respectively. Let us begin with the first pair of the estimated model, which is in the left half of the table.

In the first hurdle, all the variables except *FLOWT* and *SPEED* have significant coefficients (except *CFS*) with expected signs. *SPEED* has a significant and negative coefficient. This result suggests that an increased speed of innovation in the industry tends to discourage marginal R&D performers from making R&D investment rather than increase R&D incentives for them. In fact, among the two-digit SIC industries, electrical machinery has the highest *SPEED* (see Appendix Table 3) but only about a half of the firms show positive in-house R&D (see Appendix Table 1) whereas, in the chemical industry, about 80 percent of the firms show positive in-house R&D but its

SPEED is lower than the entire average. However, if we look only at the firms with any R&D activity (i.e., *RDD* = 1), then the in-house R&D intensity (*RDINT*) of the electrical machinery industry is 2.4 percent, which is fourth highest among the industries and 0.7 percent point higher than the entire mean (see Appendix Table 2). That is, in this industry, R&D-intensive large firms coexist with a large number of non-R&D-performing small firms. The presence of such an industry explains why *SPEED* has a negative coefficient in the first hurdle. This result is consistent with the positive coefficient of *SPEED* on R&D intensity that is obtained in an OLS regression with R&D-performing firms only[15].

Another unexpected result is the significant and negative coefficient of *FLOWT*, as it implies that a firm in an industry where information flow from transaction-based sources is more useful is *less* likely to expend for R&D. Looking at the industrial means (see Appendix Table 3), one finds that *FLOWT* is highest in printing (including publishing). In this industry, customers are the most important information source among the four transaction-based sources. Presumably, most products are custom-made and hence close and frequent relationship with customers is required, through which the customers provide information leading to the start of new projects or the completion of existing projects, explaining the high value of *FLOWT*. This industry, however, is one of the least R&D-intensive: in fact, the proportion of firms performing in-house R&D is only 13 percent, the lowest among the industries (see Appendix Table 1), suggesting that R&D projects are carried out only by a small number of large R&D-oriented firms and/or projects need not involve R&D expenditures.

In view of this peculiar behavior of the printing industry, we added a dummy variable, *PRINT*, which is one if and only if the firm's main business is in printing, publishing, and allied industries. The estimation result with this dummy variable is shown in the right half of Table 4. As expected, *PRINT* has a very significant negative effect in the first hurdle. In addition, the coefficient of *FLOWT* turns to positive, if insignificant. This result clearly suggests that the negative coefficient of *FLOWT* in the first regression result owes to the peculiar nature of the printing industry. In other industries, there is no tendency that the propensity to perform R&D is negatively related to the level of transaction-based information flow.

[15]The result of this OLS regression is suppressed to save space but is available in Nakamura and Odagiri (2003).

7.2 The Second Hurdle

Now let us go to the estimation results of the second hurdle in Table 4. Basically, all the sign conditions shown in Table 3 are satisfied. The estimated positive coefficient of *VI* is consistent with the hypothesis that a diversified firm possesses a vertically broader capability, which makes procured R&D easier and more useful, but is inconsistent with the hypothesis that a more integrated firm is operating under an environment with higher market transaction costs. The coefficient of *CFS* is insignificant, supporting neither the hypothesis of cost-sharing and risk-sharing motivation for commissioned R&D nor the hypothesis that investment in commissioned R&D is riskier than investment in physical assets and therefore needs to be financed internally.

APPRO has an expected positive coefficient, supporting the hypothesis that a stronger patent right helps the firms to reduce transaction costs. *FLOWS* has a significantly positive coefficient, suggesting that information flow from scientific sources enriches technological opportunity and the firms are motivated to commission R&D to take advantages of such opportunity. As in the first hurdle, *FLOWT* has a negative and significant coefficient but, when *PRINT* is added, the significance is lost. The coefficient of *PRINT* is negative and significant, suggesting that printing and publishing firms are inactive in commissioned R&D as well.

SPEED has negative coefficients as in the first hurdle. In an industry with fast innovation, the firm may feel a stronger need to accumulate capabilities internally and/or there may not be many qualified parties (firms, laboratories, or universities) to commission R&D to.

7.3 Determinants of the Three Modes of Procured R&D

So far we have only considered commissioned R&D as a means of procured R&D. We will now expand the analysis and estimate the double-hurdle model with each mode of procured R&D as an alternative second-hurdle variable. The results are summarized in Table 5, which shows only the estimated results of the second hurdle because we confirmed that the estimation result of the first hurdle is insensitive to the choice of the dependent variable in the second hurdle. In addition, even though we now used *RDD* as the first-hurdle variable in place of *ICRDD* for the reason discussed in Section 5, the estimated first-hurdle results are basically unchanged. In other words, the estimation results of the first-hurdle equation shown in Table 4 stand regardless of whether *RDD* or *ICRDD* is taken in the first hurdle and which of the procured R&D variable is taken in the second hurdle[16].

In view of the peculiarity of the printing industry as shown above, we only present results with *PRINT*.

The estimated *CRD* equation is slightly different from that in Table 4 (in the far right column) because the first-hurdle variable is now *RDD*; however, the signs of all the coefficients are unchanged. These signs are the same in the *CRDN* equation, that is, when in-group commissioned R&D is excluded. The only difference from the *CRD* equation is that the negative coefficient of *PC* is now significant as hypothesized in Section 6.

The same signs generally hold in the *JRD* and *TA* equations, suggesting the similarly of the determinants of the three modes of procured R&D. However, a few interesting variations appear regarding the coefficients of industrial variables. First, *FLOWS* has positive and significant coefficients in the *CRD* and *JRD* equations but the significance is lost in the *TA* equation. In an industry with abundant information flow from scientific sources, commissioned R&D and joint R&D tend to be more active but there is no significant difference regarding technology acquisitions. It is suggested that, in such an industry, firms are eager to incorporate advanced scientific knowledge through commissioning of R&D and joint R&D.

Second, the coefficient of *FLOWT*, which is insignificant but negative in the *CRD* equation, is positive and significant in the *JRD* and *TA* equations. That is, in an industry with frequent information flow from transaction-based sources, such as suppliers, customers, and competitors, joint R&D is more actively performed as well as licensing, but not commissioned R&D. It is suggested that, in an industry in which such information flow is useful, the firm finds more opportunity for joint R&D with its transaction partners or for licensing-in of technology from them. By contrast, commissioning of R&D is more likely to occur to scientific sources, such as universities, as just discussed or, possibly, to R&D specialists.

Third, *SPEED* has a significant negative coefficient in the *CRD* equation, negative but an insignificant one in the *JRD* equation, and a positive and significant one in the *TA* equation. Presumably, when innovation occurs rapidly, the firm is keen to catch up with

[16] For detailed estimation results, see Nakamura and Odagiri (2003).

innovation through licensing-in of already invented technologies, rather than through commissioning of R&D because it would take substantial time until the outcome is gained from commissioned R&D and the predictability of its outcome is low.

In contrast to these three variables, the remaining industrial variable, *APPRO*, has consistently positive and significant coefficients regardless of the modes of procured R&D, strongly supporting the hypothesis that effective protection by patents contributes to the reduction of transaction costs.

Many of these results do not hold in the *CRDI* equation. In particular, the coefficient of *PC* is positive and strongly significant as expected, making a good contrast to the negative and significant coefficient in the *CRDN* equation. Parent-controlled firms tend not to commission R&D to outside of the group but they do commission more R&D within the group. Presumably, the decision on the commissioning of R&D to outside is made by parent firms. In-group R&D capabilities are probably also concentrated to parent firms or in-group R&D companies, and the subsidiaries commission necessary R&D to these firms.

Furthermore, all the coefficients of industry variables are insignificant in the *CRDI* equation and so are the coefficients of *VI* and *DIV*, the variables showing the vertical and diversifying breadth of the firm's capabilities. We may therefore conclude that the activity of in-group commissioning of R&D is dependent on neither the industrial characteristics nor the firm's organizational form.

7.4 Estimation Results for Large Firms

In order to investigate if our use of the double-hurdle model does matter, we show, in the left half of Table 6, the determinants of procured R&D estimated as single-equation Tobit models. Comparing it with Table 5, we find that the general tendency is the same, confirming the robustness of our estimation results. Yet, z values tend to be higher in the double-hurdle model estimations (except *RDINT*) and some of the statistical significance is lost in the Tobit results (e.g., *APPRO* in the *JRD* equation and *SPEED* in the *TA* equation). This difference comes from the fact that, in the double-hurdle model, the second-hurdle equation is relevant only for R&D-performing firms (see the likelihood function in Appendix 2) whereas, in the Tobit model, all the sample firms, whether they are performing R&D or not, are treated equally. Because *RDINT* is a decisive variable to separate R&D-performing firms from others (as *RDINT* = 0 for all non-R&D-performing firms), it is estimated to have a larger-than-real explanatory power in the Tobit model. The other variables, by contrast, lose their explanatory power because of the presence of so many zero-valued dependent variables.

This difficulty with the Tobit model is mitigated when we restrict our sample to large firms, because non-R&D-performing firms become relatively unimportant among these large firms.

The Small and Medium Enterprise Basic Law of Japan defines 'small and medium enterprises' (SME) as the firms with capitalization not in excess of 300 million yen or with 300 or fewer employees. Hence, in the following, we define 'large firms' as those that do not satisfy this criterion of SME. There were 2,026 such firms in our sample, which comprise 14.4 percent of the whole sample. Among these, 1,715 (84.6 percent of 2,026) performed at least one form of R&D, whether in-house or procured, that is, *RDD* = 1. Hence, non-R&D-performing firms are minority among the large firms, in contrast to the fact that non-R&D-performing firms accounted for 52.8 percent among the entire sample.

The right-hand half of Table 6 shows the Tobit estimation results for these large firms. Comparing it with the estimation results of the entire sample in the left-hand half of the same table or with the second-hurdle estimation results in Table 4, we find that the results are reasonably similar, with only a few changes. First, the coefficients of *PC* are not significant among large firms, suggesting that large subsidiary firms are acting more like independent firms in terms of their R&D activity. This result should appear reasonable if one recalls the fact that such large Japanese firms as JVC and Hino Motors (without implying that these firms are in fact in our sample) are subsidiaries (of Matsushita Electric and Toyota, respectively).

Second, the results on industrial variables are mixed. Compared to the results with the entire sample, in the *CRD* equation, *FLOWS* and *SPEED* lose significance. When information flow from scientific sources is abundant, large firms may be able to acquire and absorb such information through papers and other spillovers without commissioning R&D. A larger incentive to speed up innovation through commissioning R&D may be stronger among large firms, offsetting the desire to accumulate capabilities through in-house R&D. In the *TA* equation, the positive coefficient of *SPEED* is strengthened in terms of both the coefficient and its statistical fit. In the *JRD* and *TA* equations, *APPRO* becomes insignificant. This result is unsatisfactory, particularly because *APPRO* is constructed from the survey of large firms only and should therefore reflect the views of large firms better. The estimation results of the double-hurdle model appear more reasonable in this regard.

In conclusion, even though most of the general results hold even with single-equation Tobit models, we believe that double-hurdle models provide more reasonable estimates, particularly when the behavior of mostly non-R&D-performing small and medium enterprises should be also accounted.

8 Summary and Conclusions

In this paper, we argued for the importance of the issue of R&D boundaries of the firm, namely, the firm's choice between performing R&D in-house *versus* procuring it from outside. Various modes of R&D procurement are available and we classified them between commissioned R&D, joint R&D, and technology acquisitions (i.e., licensing-in). There is an important difference among these in terms of the timing of contract and the extent of definability and uncertainty at the time of contract.

Making use of a large-scale database of manufacturing firms in Japan, we have empirically analyzed the determinants of these three modes of procured R&D to test the hypotheses built around the two major theories – the transaction cost theory and the capability theory. In view of the presence of a large number of firms who failed to perform any R&D activity at all, we formulated the R&D decision process as a double-hurdle model and estimated this model with a maximum likelihood methodology.

Generally, the estimation results support the two theories. Most importantly, we found positive impacts of firm size, in-house R&D intensity, diversification, and vertical integration, which supports the hypothesis that the presence of a large and broad absorptive capacity is a contributing factor for procured R&D by making it easier for the firms to seek potential partners, evaluate them, monitor R&D alliances, and utilize the outcome for commercialization. We have also found a positive effect of the index of appropriability by patents, which supports the hypothesis that appropriability reduces transaction costs. Many of these results apply to R&D commissioned to non-group organizations, joint R&D, and licensing but not necessarily to R&D commissioned to in-group firms, suggesting that groups are quasi-internal organizations and therefore monitoring and appropriability issues do not arise.

We also found that (1) information flow from scientific sources (universities, public laboratories, and academic associations) stimulates commissioned R&D and joint R&D, (2) information flow from transaction-based sources (suppliers, customers, and competitors) stimulates joint R&D and technology acquisitions and, (3) firms in fast-innovating industries tend to rely on licensing-in to acquire completed technologies and rather refrain from commissioning R&D, presumably because it would take time before the outcome is to be gained from commissioned R&D.

Of course, there still remain many issues to be addressed. The first is the adequacy of the measure of R&D procurement activity. For instance, the number of partners in a joint R&D project need not be related to the intensity and efficiency of the project, particularly because firms may be more tempted to free ride on the partners' efforts if there are many participants. Technology acquisitions are measured by the amount the firm paid for acquired technologies. However, since most payment for licensing is composed of fixed initial payment and running royalty which is usually a fixed percentage of sales, it can violently fluctuate from year to year and the firm keeps paying for many years after the actual technology acquisition has taken place.

Second, there remains a possibility of endogeneity of some of the independent variables. In reality, the firm would first determine how much resources to be invested for its entire R&D activity and, then, would determine the best mix of in-house R&D, commissioned R&D, joint R&D, and licensing. With the adoption of a double-hurdle model, we believe we have made an important step towards analyzing this sequence of decisions. Still, we have not yet fully investigated the simultaneous decision of in-house R&D and the three modes of procured R&D and instead estimated the determinants of each means of procured R&D separately, using in-house R&D intensity as an explanatory variable. How to model such simultaneous decision of the firm and how to estimate such a model are big questions that we intend to pursue in the future. The present analysis, we hope, provides a good starting point towards such a more comprehensive analysis.

Appendix 1. Data Source

The *Basic Survey of Business Structure and Activities* (BSA) was first compiled by the then Ministry of International Trade and Industry (MITI; reorganized as the Ministry of Economy, Trade and Industry or METI in 2001) in 1992 and then every year since 1995. BSA covers all the firms in Japan that meet the following three conditions; (1) the firm has an establishment classified to either major division D (mining), F (manufacturing) or I (wholesale and retail trade, eating and drinking place) of the Japanese Standard Industrial Classification (JSIC), (2) with 50 employees or more, and (3) with capitalization (i.e., the book value of equity) of 30 million yen or more. With METI's kind permission, we use the unpublished firm-level data of this survey for 1997.

Aggregated industry data of BSA has been published in a series of official reports[17]. However, our scheme of industrial classification is different from that used in these reports. They classified each firm into one of the 2-digit JSIC industries according to the 3-digit industry with largest sales. We aggregated the firm's 3-digit sales composition to that of 2-digit and, then, classified the firm into the industry with the largest sales. For instance, suppose that the firm sells products in three 3-digit industries, say, 303 (communication equipment), 304 (computers), and 329 (miscellaneous precision instrument), which comprise, respectively, 30, 30, and 40 percent of the firm's total sales. Then, the official report classifies the firm into the 2-digit industry 32 (precision instrument) whereas we classified it into the 2-digit industry 30 (electrical equipment) because the sum of the sales in industries 303 and 304 outweighs that in industry 329. Industry statistics shown in Appendix Tables 1, 2, and 3 were calculated according to this classification scheme of ours.

The sample in our analysis consists of all the manufacturing firms in the survey[18]. However, apparent 'outliers' were eliminated. For instance, one firm reported that it had joint R&D agreements with 315 partners, which is extraordinary large compared to other firms. It is, in fact, difficult to imagine that the firm can maintain effective R&D collaboration with such many partners. We eliminated 21 similarly apparent outliers so that all the samples satisfy the following conditions (see Tables 1 and 2 for the variable

[17] See http://www.meti.go.jp/english/statistics/data/h2c1tope.html for a preliminary report in English.

[18] Similarly to our 2-digit industrial classification explained in the previous paragraph, we aggregated the firm's 3-digit sales composition to that of 1-digit and defined manufacturing firms as the firms whose largest 1-digit industry classification is F (manufacturing).

symbols): (1) $CRD < 40,000$, (2) $JRD < 150$, (3) $TA < 14,000$, (4) $RDINT < 0.35$, (5) $VI < 1$, (6) $-1 < CFS < 1$. The final sample contains 14,070 firms.

The industrial distribution of these firms is shown in the second column of Appendix Table 1, together with the number of firms that reported a positive value for each dependent variable. Appendix Table 2 shows industrial means of in-house and procured R&D intensities, calculated as the ratios (in percentage) to the firm's sales. Note that *JRD* is the number of partners in joint R&D; hence, its ratio to sales is difficult to interpret and not comparable to the R&D intensity as usually defined, such as the ratio of *IRD* to sales. Appendix Table 3 gives the industrial means of the explanatory variables.

Appendix 2. The Double-Hurdle Model

Following Flood and Grasjo (2001), we write a two-equation model as follows:

$$d_i^* = x_{1i}'\beta_1 + \nu_i \tag{1}$$

$$y_i^* = x_{2i}'\beta_2 + \epsilon_i \tag{2}$$

where, in our study, d_i^* is a latent variable representing participation in R&D and y_i^* is a latent variable representing, for instance, commissioned R&D. x_{1i} and x_{2i} are observable vectors of explanatory variables, β_1 and β_2 are the vectors of parameters, and the random errors $(\nu_i, \epsilon_i)'$ are assumed to obey i.i.d. bivariate normal distribution (BVN) with mean zero and variance-covariance matrix as follows[19]:

$$(\nu_i, \epsilon_i)' \sim BVN(0, \Sigma), \quad \Sigma = \begin{bmatrix} 1 & \sigma\rho \\ \sigma\rho & \sigma^2 \end{bmatrix} \tag{3}$$

We impose the following threshold conditions:

$$d_i = \begin{cases} 1 \ \ if \ \ d_i^* > 0 \\ 0 \ \ if \ \ d_i^* \leq 0 \end{cases} \tag{4}$$

$$y_i = \begin{cases} y_i^* \ \ if \ \ d_i = 1 \ \ and \ \ y_i^* > 0 \\ 0 \ \ otherwise \end{cases} \tag{5}$$

where d_i is an observed value, which equals $ICRDD_i$ in our study. Similarly, y_i in our study is CRD_i, etc. Using these equations, we write the likelihood function as follows:

$$\begin{aligned} L = \prod_{d_i=0}(1 - F(\nu_i > -x_{1i}'\beta_1)) \prod_{d_i=1, y_i=0} F(\nu_i > -x_{1i}'\beta_1, -x_{2i}'\beta_2 \geq \epsilon_i) \\ \times \prod_{d_i=1, y_i>0} F(\nu_i > -x_{1i}'\beta_1, \epsilon_i > -x_{2i}'\beta_2) f(\epsilon_i | \nu_i > -x_{1i}'\beta_1, \epsilon_i > -x_{2i}'\beta_2) \end{aligned} \tag{6}$$

where $f(\cdot)$ and $F(\cdot)$ denote density and cumulative distribution functions respectively. The first multiplicative term corresponds to the probability of the case in which $d_i = 0$

[19] In Cragg's original model, the error terms, ν_i and ϵ_i, were assumed independent. However, as equations (1) and (2) are both related to the R&D activity of the same firm, it is likely that unobservable common factors generate correlation between the residual errors. We thus assume equation (3). While this assumption follows that of the 'double-hurdle dependent model' of Jones (1992), our model differs from his in an important way. In Jones's model, the information on d_i is lacking and it was assumed that $d_i = 1$ if and only if $y_i > 0$. We, on the other hand, have information on d_i (namely, $ICRDD_i$) and utilize this information to formulate the likelihood function below.

(and hence $y_i = 0$), and the second term, $d_i = 1$ and yet $y_i = 0$. The last term gives the probability that, given that these double hurdles are cleared, y_i^* is realized as y_i.

Combining all these equations, we have

$$L = \prod_{d_i=0} \Phi(-x'_{1i}\beta_1) \prod_{d_i=1,y_i=0} \Phi_2(x'_{1i}\beta_1, -x'_{2i}\beta_2/\sigma, \rho) \times \prod_{d_i,y_i>0} \left\{ \Phi\left(\frac{x'_{1i}\beta_1 + \frac{\rho}{\sigma}(y_i - x'_{2i}\beta_2)}{\sqrt{1-\rho^2}} \right) \frac{1}{\rho} \phi((y_i - x'_{2i}\beta_2)/\rho) \right\} \quad (7)$$

where Φ and Φ_2 are the standard normal distribution functions for, respectively, univariate and bi-variate cases. By maximizing (7), we get consistent estimates of β_1, β_2, and Σ.

References

Ahuja, Gautam and Katila, Ritta (2001) "Technological Acquisitions and the Innovation Performance of Acquiring firms: A Longitudinal Study," *Strategic Management Journal*, 22, 197-220.

Arora, Ashish and Fosfuri, Andrea (2003) "Licensing the Market for Technology," *Journal of Economic Behavior & Organization,* 52, 277-295.

Arora, Ashish and Gambardella, Alphonso (1990) "Complementarity and External Linkages: The Strategies of the Large Firms in Biotechnology," *Journal of Industrial Economics*, 38, 361-379.

Arora, Ashish and Gambardella, Alphonso (1994) "Evaluating Technological Information and Utilizing it: Scientific Knowledge, Technological Capability, and External Linkages in Biotechnology," *Journal of Economic Behavior & Organization*, 24, 91-114.

Audretsch, David B.; Menkveld, Bert; and Thurik, A. Roy (1996) "The Decision between Internal and External R&D," *Journal of Institutional and Theoretical Economics*, 152, 519-530.

Bayona, Cristina; Garcia-Marco, Teresa; and Huerta, Emilio (2001) "Firms' Motivations for Cooperative R&D: An Empirical Analysis of Spanish Firms," *Research Policy*, 30, 1279-307.

Bönte, Werner (2003) "R&D and Productivity: Internal vs. External R&D: Evidence from West German Manufacturing Industries," *Economics of Innovation and New Technology*, 12, 343-60.

Cassiman, Bruno and Veugelers, Reinhilde (2002) "R&D Cooperation and Spillovers: Some Empirical Evidence from Belgium," *American Economic Review*, 92, 1169-1184.

Chandler, Alfred. D., Jr. (1990) *Scale and Scope*. Cambridge, Mass.: Belknap Press.

Cohen, Wesley M.; Goto, Akira; Nagata, Akiya; Nelson, Richard R.; and Walsh, John P. (2002) "R&D Spillovers, Patents and the Incentive to Innovate in Japan and the United States," *Research Policy*, 31, 1349-1367.

Cohen, Wesley M. and Klepper, Steven (1996) "A Reprise of Size and R&D," *Economic Journal*, 106, 925-951.

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

Cohen, Wesley M. and Levinthal, Daniel A. (1990) "Absorptive Capacity: A New Perspective on Learning and Innovation," *Administrative Science Quarterly*, 35, 128-152.

Cohen, Wesley M.; Nelson, Richard R.; and Walsh, John P. (2000) "Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)," NBER Working Paper, No. 7552, National Bureau of Economic Research.

Cohen, Wesley M.; Nelson, Richard R.; and Walsh, John P. (2002) "Links and Impacts: The Influence of Public Research on Industrial R&D," *Management Science*, 48, 1-23.

Colombo, Massimo G. and Garrone, Paola (1996) "Technological Cooperative Agreements and Firm's R&D Intensity: A Note on Causality Relations," *Research Policy*, 25, 923-932.

Cragg, John G. (1971) "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods," *Econometrica*, 39, 829-844.

Flood, Lennart and Gråsjö, Urban (2001) "A Monte Carlo Simulation Study of Tobit Models," *Applied Economics Letters*, 8, 581-84.

Freeman, Christopher (1991) "Networks of Innovators: A Synthesis of Research Issues," *Research Policy*, 20, 499-514.

Gans, Joshua S.; Hus, David H.; and Stern, Scott (2000) "When Does Start-Up Innovation Spur the Gale of Creative Destruction?," NBER Working Paper, No. 7851, National Bureau of Economic Research.

Goto, Akira; Koga, Tadahisa; and Suzuki, Kazuyuki (2002) "Waga Kuni Seizogyo ni Okeru Kenkyu Kaihatsu Toshi no Kettei Youin" [The Determinants of R&D Investment in the Japanese Manufacturing Industries: Small Firms and Large Firms], *Keizai Kenkyu*, 53, 18-23.

Goto, Akira and Nagata, Akiya (1996) “Sabei Deta ni Yoru Inobeshon Purosesu no Kenkyu” [A Study of Innovation Process by Survey Data], unpublished report, National Institute of Science and Technology Policy.

Granstrand, Ove; Patel, Pari; and Pavitt, Keith (1997) “Multi-Technology Corporations: Why They Have ‘Distributed’ Rather than ‘Distinctive Core’ Competencies,” *California Management Review*, 39, 8-25.

Griliches, Zvi (1992) “The Search for R&D Spillovers,” *Scandinavian Journal of Economics*, 94, 29-47. Reprinted in Zvi Griliches, *R&D and Productivity*. The University of Chicago: Chicago Press, 1998, 251-268.

Hagedoorn, John (1993) “Understanding the Rationale of Strategic Technology Partnering: Interorganizational Modes of Cooperation and Sectoral Differences,” *Strategic Management Journal*, 14, 371-385.

Hagedoorn, John (2002) “Inter-Firm R&D Partnerships: An Overview of Major Trends and Patterns since 1960,” *Research Policy*, 31, 477-492.

Hernán, Roberto; Marín, Pedro L.; and Siotis, George (2003) “An Empirical Evaluation of the Determinants of Research Joint Venture Formation,” *The Journal of Industrial Economics*, 51, 75-90.

Himmelberg, Charles P. and Petersen, Bruce C. (1994) “R&D and Internal Finance: A Panel Study of Small Firms in High-Tech Industries,” *Review of Economics and Statistics*, 76, 38 - 51.

Hoskisson, Robert E.; Hitt, Michael A.; and Hill, Charles W. L. (1993) “Managerial Incentives and Investment in R&D in Large Multi-Product Firms,” *Organization Science*, 4, 325-341.

Howells, Jeremy (1999) “Research and Technology Outsourcing,” *Technology Analysis & Strategic Management*, 11, 17-29.

Huber, George P. (1991) “Organizational Learning: The Contributing Processes and the Literatures,” *Organization Science*, 2, 88-115.

Inkpen, Andrew C.; Sundaram, Anant K.; and Rockwood, Kristin (2000) "Cross Border Acquisitions of US Technology Assets," *California Management Review*, 42, 50-71.

Jones, Andrew M. (1992) "A Note on Computation of the Double-Hurdle Model with Dependence with an Application to Tobacco Expenditure," *Bulletin of Economic Research*, 44, 67-74.

Jovanovic, Boyan (1993) "The Diversification of Production," *Brookings Papers on Economic Activity. Microeconomics*, 1993, 197-247.

Kleinknecht, Alfred and Reijnen, Jeroen O. N. (1992) "Why Do Firms Cooperate on R&D? An Empirical Study," *Research Policy*, 21, 347-360.

Levin, Richard C.; Klevorick, Alvin K.; Nelson, Richard R.; and Winter, Sidney G. (1987) "Appropriating the Returns from Industrial Research and Development," *Brookings Papers on Economic Activity*, 3, 783-832.

Nakamura, Kenta and Odagiri, Hiroyuki (2003) "Determinants of R&D Boundaries of the Firm: An Empirical Study of Commissioned R&D, Joint R&D, and Licensing with Japanese Company Data," Discussion Paper No. 32, National Institute of Science and Technology Policy.

Nelson, Richard R. (1959) "The Simple Economics of Basic Scientific Research," *Journal of Political Economy*, 67, 297-306.

Nelson, Richard R. and Winter, Sidney G. (1982) *An Evolutionary Theory of Economic Change*. Cambridge, Mass.: Belknap Press.

Odagiri, Hiroyuki (1992) *Growth through Competition, Competition through Growth: Strategic Management and the Economy in Japan*. Oxford: Clarendon Press.

Odagiri, Hiroyuki (2003) "Transaction Costs and Capabilities as Determinants of the R&D Boundaries of the Firm: A Case Study of the Ten Largest Pharmaceutical Firms in Japan," *Managerial and Decision Economics*, 24, 187-211.

Odagiri, Hiroyuki and Goto, Akira (1996) *Technology and Industrial Development in Japan: Building Capabilities by Learning, Innovation, and Public Policy*. Oxford: Clarendon Press.

Odagiri, Hiroyuki; Koga, Tadahisa; Kenta, Nakamura (2002) “Baio Tekunoroji Kenkyu Kaihatsu to Kigyo no Kyokai: Kenkyu Teikei, Gijyutsu Donyu, Autosoushingu, Kaigai Kenkyu Kaihatsu ni Kansuru Chosa Houkoku,” [Biotechnology R&D and the Boundaries of the Firm: Results from a Survey Study on R&D Alliances, Thechnology Acquisition, Outsourcing, and Overseas R&D,”] Chosa Shiryo [Research Material] No. 90, National Institute of Science and Technology Policy.

Patel, Pari and Pavitt, Keith (1997) “The Technological Competencies of the World’s Largest Firms: Complex and Path-Dependent, but Not Much Variety,” *Research Policy*, 26, 141-156.

Penrose, Edith T. (1959) *The Theory of the Growth of the Firm*. Oxford: Basil Blackwell.

Pisano, Gary P. (1990) “The R&D Boundaries of the Firm: An Empirical Analysis,” *Administrative Science Quarterly*, 35, 153-176.

Powell, Walter W.; Koput, Kenneth W.; and Smith-Doerr, Laurel (1996) “Interorganisational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology,” *Administrative Science Quarterly*, 41, 116-145.

Prahalad, C. K. and Hamel, Gary (1990) “The Core Competence of the Corporation,” *Harvard Business Review*, May-June, 68, 79-91.

Rocha, Frederico (1999) “Inter-Firm Technological Cooperation: Effects of Absorptive Capacity, Firm Size and Specialization,” *Economics of Innovation and New Technology*, 8, 253-271.

Schumpeter, Joseph A. (1942) *Capitalism, Socialism, and Democracy*. New York: Harper and Row.

Teece, David J. (1986) “Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy,” *Research Policy*, 15, 285-305.

Teece, David J.; Pisano, Gary; and Shuen, Amy (1997) “Dynamic Capabilities and Strategic Management,” *Strategic Management Journal*, 18, 509-533.

Tobin, James (1958) "Estimation of Relationships for Limited Dependent Variables," *Econometrica*, 26, 24-36.

Ulset, Svein (1996) "R&D Outsourcing and Contractual Governance: An Empirical Study of Commercial R&D Projects," *Journal of Economic Behavior & Organization*, 30, 63-82.

Veugelers, Reinhilde (1997) "Internal R&D Expenditures and External Technology Sourcing," *Research Policy*, 26, 303-316.

Veugelers, Reinhilde and Cassiman, Bruno (1999) "Make and Buy in Innovation Strategies: Evidence from Belgian Manufacturing Firms," *Research Policy*, 28, 63-80.

Wakasugi, Ryuhei (1999) "Mikurodeta ni Motoduku Kigyoukatsudou no Takakuka, Kokusaika, Sohutoka ni Kansuru Teiryoubunseki," [A Quantitative Study of Diversification, Internationalization, and Softenization of Corporate Activities with Micro Data], unpublished report.

Wernerfelt, Birger (1984) "A Resource-Based View of the Firm," *Strategic Management Journal*, 5, 171-180.

Williamson, Oliver E. (1975) *Markets and Hierarchy*. New York: Free Press.

Williamson, Oliver E. (1985) *The Economic Institutions of Capitalism*. New York: Free Press.

Table 1. R&D Variables: Descriptive Statistics

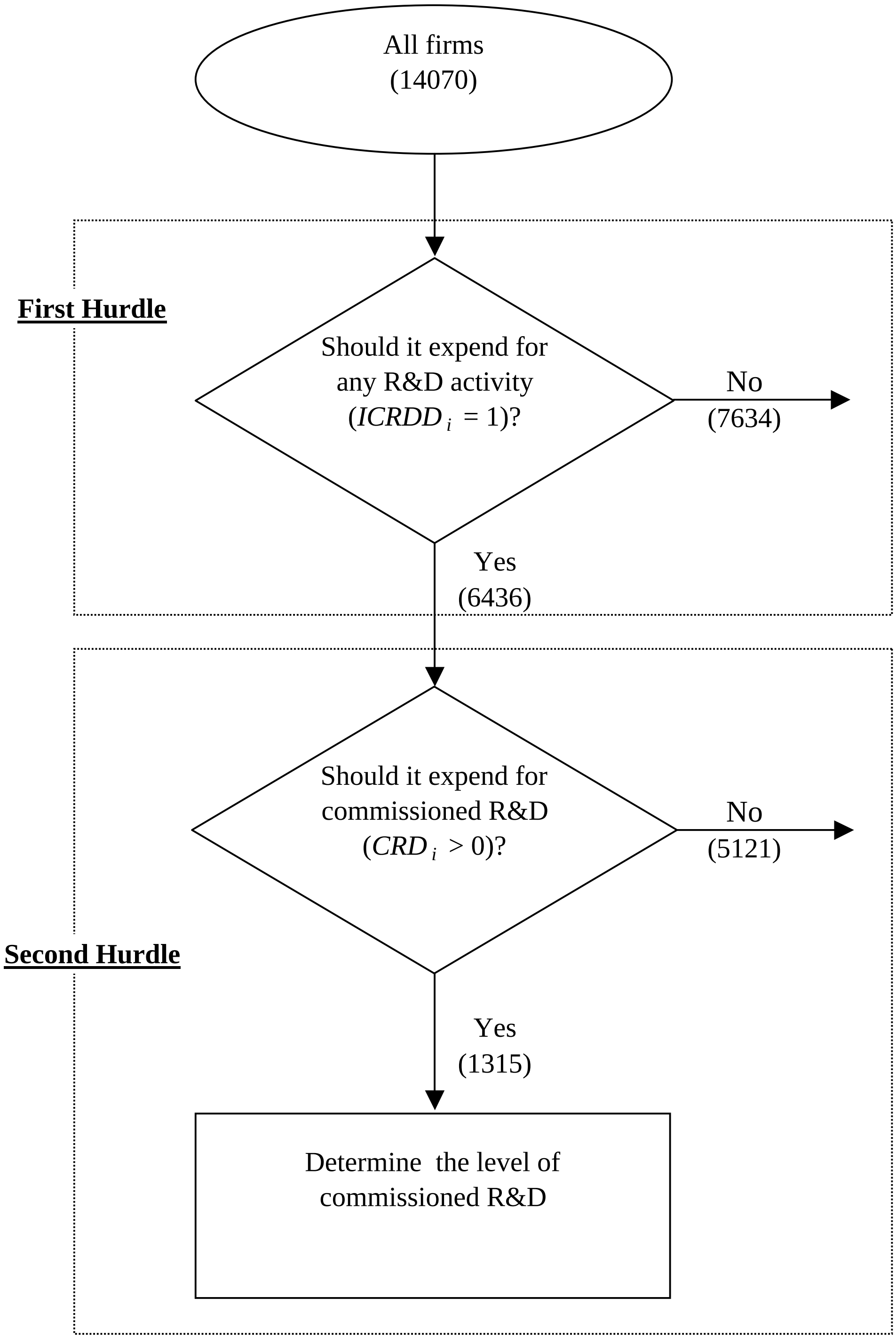
(in million yen, except *n* and *JRD*)

Symbol	Description	Whole sample				Sample with positive values			
		n	Mean	Std. Dev.	Max	*n*	Mean	Median	Std. Dev.
IRD	In-house R&D expenditures	14070	466.46	6632.40	427800	6281	1044.91	54.00	9896.60
CRD	Total commissioned R&D expenditures	14070	21.20	294.42	14907	1315	226.79	15.00	938.87
CRDN	Non-group commissioned R&D expenditures	14070	16.85	271.93	14907	1150	206.15	10.69	930.81
CRDI	In-group commissioned R&D expenditures	14070	4.35	80.07	4336	296	312.85	29.91	910.61
JRD	Number of joint R&D partners	14070	0.23	1.98	75	950	3.47	2.00	6.83
TA	Payment for technology acquisitions	14070	11.37	178.53	8215	834	191.76	16.00	709.72

Notes: 1. n = number of observations (i.e., number of firms).
2. For any variable, the median for the 'whole sample' equals 0 and the maximum value for the 'sample with positive values' equals that for the 'whole sample'.
3. Subscript *i* is suppressed.

Data Source: BSA

Figure 1. R&D Decision Flow Chart: The Double-Hurdle Model



Notes: 1. $ICRDD_i = 1$ if $IRD_i + CRD_i > 0$
$= 0$ if $IRD_i + CRD_i = 0$.
2. In parentheses are the number of firms.

Table 2. List of Independent Variables and Descriptive Statistics

Symbol	Name	Description	Whole sample (*n* =14070)					Sample with *RDD* =1 (*n* = 6648)		
(subscript *i* suppressed)			Mean	Median	Std. Dev.	Min	Max	Mean	Median	Std. Dev.
RDINT	In-house R&D intensity	In-house R&D expenditure / sales	0.008	0	0.018	0	0.332	0.017	0.009	0.024
LSALE	Size	Sales in natural logarithm	8.407	8.191	1.287	4.454	15.866	8.907	8.675	1.375
VI	Vertical integration	Value-added / sales	0.293	0.277	0.131	0.001	0.986	0.283	0.272	0.112
DIV	Diversification	Index of product diversification (1-$H^{1/2}$, where H = Herfindahl index)	0.143	0.092	0.151	0	0.656	0.167	0.133	0.156
CFS	Cash flow ratio	Cash flow / sales	0.044	0.039	0.059	-0.931	0.964	0.047	0.042	0.058
PC	Parent-controlled	A dummy variable that equals 1 if and only if the firm has a parent company	0.280	0	0.449	0	1	0.262	0	0.440
APPRO	Appropriability	Appropriability by patents*	0.322	0.314	0.070	0.142	0.615	0.332	0.331	0.075
FLOWS	Scientific information flow	The average of information flow from three scientific sources (universities, public laboratories, and academic associations)*	0.385	0.369	0.081	0.247	0.825	0.394	0.369	0.101
FLOWT	Transaction-based information flow	The average of information flow from four transaction-based sources (suppliers with or without share ownership relationship, customers, and competitors)*	0.463	0.472	0.080	0.341	0.625	0.458	0.462	0.074
SPEED	Innovation speed	Speed of innovation change*	3.064	3.090	0.281	2.038	3.786	3.064	3.063	0.281
PRINT	Printing and allied industry dummy	A dummy variable that equales 1 if and only if the firm is in printing, publishing, and allied industries	0.056	0	0.230	0	1	0.017	0	0.130

Note: *n* = number of observations (i.e., number of firms). *RDDi=1* if and only if min (*IRDi*, *CRDi*, *JRDi*, *TAi*)>0.
Data Source: BSA, except * by Goto and Nagata (1996)

Table 3. Hypothesized Signs of the Coefficients

	Dependent Variables	
	First Hurdle	Second Hurdle
Independent Variables	*RDD* or *ICRDD*	*CRD* (or *CRDN*, *CRDI*), *JRD*, or *TA*
RDINT		+
LSALE	+	+
VI	+	+/−
DIV	+	+
CFS	+	+/− (likely - for *TA*)
PC	−	− (likely + for *CRDI*)
APPRO	+	+
FLOWS	+	+/−
FLOWT	+	+/−
SPEED	+/−	+/−
PRINT	−	−

Note: Subscript i is suppressed.

Table 4. Estimation Results of the Double Hurdle Model

	First hurdle	Second hurdle	First hurdle	Second hurdle
Dependent var.	*ICRDD*	*CRD*	*ICRDD*	*CRD*
RDINT		5,194.386		5,167.408
		(5.97)***		(5.91)***
LSALE	0.476	340.369	0.484	341.480
	(38.50)***	(7.30)***	(38.75)***	(7.30)***
VI	1.174	625.825	1.271	656.230
	(10.94)***	(3.31)***	(11.70)***	(3.44)***
DIV	0.793	314.247	0.707	278.098
	(10.40)***	(2.93)***	(9.20)***	(2.61)***
CFS	0.102	30.981	-0.033	-5.982
	(0.45)	(0.08)	(-0.14)	(-0.02)
PC	-0.200	-6.386	-0.221	-13.353
	(-7.87)***	(-0.20)	(-8.66)***	(-0.42)
APPRO	2.292	2,124.185	1.518	1,898.605
	(12.56)***	(5.82)***	(8.09)***	(5.50)***
FLOWS	1.015	715.663	1.308	827.528
	(6.24)***	(3.46)***	(8.16)***	(3.86)***
FLOWT	-0.957	-525.078	0.201	-134.098
	(-6.70)***	(-2.52)**	(1.24)	(-0.64)
SPEED	-0.170	-139.879	-0.122	-123.609
	(-3.96)***	(-2.48)**	(-2.83)***	(-2.22)**
PRINT			-1.040	-492.434
			(-14.67)***	(-3.54)***
Constant	-4.683	-4,971.448	-5.253	-5,163.867
	(-25.45)***	(-7.16)***	(-27.94)***	(-7.17)***
SIGMA	938.492		938.291	
	(7.21)***		(7.20)***	
RHO	0.953		0.955	
	(66.29)***		(55.57)***	
Log likelihood	-20474.98		-20347.38	

Notes: 1. No. of observations = 14070.
2. Robust z statistics in parentheses.
3. * significant at 10%; ** at 5%; *** at 1%.

Table 5. The Determinants of Procured R&D: The Second-Hurdle Estimation Results

Mode of Procured R&D	*CRD*	*CRDN*	*CRDI*	*JRD*	*TA*
RDINT	5,535.647	5,518.671	3,238.026	25.166	2,316.237
	(6.04)***	(5.55)***	(3.72)***	(3.44)***	(4.03)***
LSALE	338.874	316.566	188.746	2.164	325.721
	(7.30)***	(6.56)***	(5.72)***	(8.45)***	(7.01)***
VI	643.295	505.411	150.253	5.907	830.595
	(3.41)***	(2.76)***	(0.63)	(3.17)***	(4.25)***
DIV	278.249	336.786	103.825	7.398	254.792
	(2.62)***	(3.04)***	(0.87)	(6.08)***	(2.48)**
CFS	-15.035	169.404	-30.193	-2.932	-429.075
	(-0.04)	(0.54)	(-0.06)	(-0.98)	(-1.20)
PC	-12.785	-114.789	229.138	-0.654	3.585
	(-0.40)	(-2.91)***	(5.09)***	(-1.84)*	(0.11)
APPRO	1,875.869	2,036.336	213.622	6.746	1,122.837
	(5.47)***	(5.48)***	(0.70)	(2.68)***	(4.18)***
FLOWS	824.089	904.348	-158.163	9.519	271.224
	(3.85)***	(3.88)***	(-0.67)	(5.10)***	(1.57)
FLOWT	-133.020	-203.695	186.441	9.438	347.438
	(-0.63)	(-0.95)	(0.75)	(3.96)***	(1.66)*
SPEED	-125.439	-147.891	-24.047	-0.724	101.083
	(-2.26)**	(-2.51)**	(-0.37)	(-1.22)	(1.68)*
PRINT	-491.648	-488.668	-259.343	-7.243	-461.003
	(-3.54)***	(-3.16)***	(-1.86)*	(-5.40)***	(-3.33)***
Constant	-5,126.654	-4,922.450	-3,326.851	-44.426	-5,387.194
	(-7.17)***	(-6.45)***	(-5.89)***	(-9.49)***	(-6.96)***
SIGMA	937.280	922.209	694.359	9.653	744.535
	(7.20)***	(6.48)***	(7.22)***	(12.13)***	(7.19)***
RHO	0.945	0.958	0.893	0.977	0.946
	(60.29)***	(37.75)***	(33.57)***	(35.03)***	(58.98)***
Log likelihood	-20379.05	-18930.08	-11245.52	-13125.74	-15979.52

See Note to Table 4.

Table 6. Determinants of Procured R&D in All Firms vs. Large Firms: Tobit Estimation Results

	All firms					Large firms				
	CRD	*CRDN*	*CRDI*	*JRD*	*TA*	*CRD*	*CRDN*	*CRDI*	*JRD*	*TA*
RDINT	11,685.849	11,120.790	6,421.634	86.856	5,981.097	14,432.390	11,502.228	10,350.146	82.638	7,756.675
	(6.99)***	(6.30)***	(5.35)***	(8.31)***	(6.67)***	(6.02)***	(5.04)***	(4.84)***	(4.50)***	(5.77)***
LSALE	288.714	271.490	159.560	1.747	294.525	524.083	510.865	202.525	2.587	383.065
	(7.21)***	(6.48)***	(5.49)***	(7.54)***	(6.89)***	(6.44)***	(5.94)***	(3.56)***	(4.18)***	(6.20)***
VI	448.304	340.099	67.637	2.989	669.808	1,264.879	1,359.488	-181.010	6.450	511.918
	(2.86)***	(2.19)**	(0.34)	(1.90)*	(4.10)***	(2.36)**	(2.60)***	(-0.32)	(1.22)	(1.26)
DIV	260.406	310.778	113.985	7.010	268.207	461.769	535.954	358.062	11.430	572.740
	(2.50)**	(2.87)***	(0.98)	(5.89)***	(2.60)***	(1.68)*	(1.88)*	(1.35)	(3.68)***	(2.62)***
CFS	-50.276	105.780	-62.109	-2.892	-403.908	-822.224	-242.578	-1,023.508	-2.231	-961.008
	(-0.14)	(0.32)	(-0.12)	(-0.94)	(-1.15)	(-0.85)	(-0.28)	(-1.00)	(-0.32)	(-1.33)
PC	-3.200	-105.933	233.253	-0.604	5.629	143.547	-5.216	353.024	0.318	93.184
	(-0.10)	(-2.71)***	(5.16)***	(-1.70)*	(0.17)	(1.63)	(-0.06)	(3.47)***	(0.34)	(1.18)
APPRO	1,492.612	1,672.470	50.103	3.614	861.509	2,702.370	3,402.256	353.302	7.926	603.164
	(4.89)***	(5.09)***	(0.17)	(1.46)	(3.42)***	(3.64)***	(4.37)***	(0.46)	(1.08)	(1.15)
FLOWS	559.864	675.827	-401.943	6.881	104.623	635.035	912.865	-1,598.804	8.792	491.411
	(2.73)***	(3.09)***	(-1.57)	(3.65)***	(0.58)	(1.17)	(1.68)*	(-2.37)**	(1.72)*	(1.29)
FLOWT	-107.736	-159.893	131.507	9.449	399.853	-259.445	-77.902	-432.060	11.736	1,175.550
	(-0.51)	(-0.74)	(0.53)	(3.92)***	(1.88)*	(-0.44)	(-0.13)	(-0.66)	(1.71)*	(2.38)**
SPEED	-144.672	-163.827	-42.843	-0.833	81.588	-212.037	-281.625	-81.557	-2.778	372.744
	(-2.64)***	(-2.80)***	(-0.68)	(-1.42)	(1.38)	(-1.53)	(-1.91)*	(-0.54)	(-1.77)*	(2.70)***
PRINT	-442.983	-449.797	-218.378	-6.501	-459.106	-858.567	-1,079.357	-114.036	-6.803	-966.434
	(-3.28)***	(-2.99)***	(-1.59)	(-4.92)***	(-3.31)***	(-2.05)**	(-2.25)**	(-0.26)	(-1.74)*	(-2.79)***
Constant	-4,388.492	-4,269.847	-2,824.888	-38.027	-4,912.897	-7,869.947	-8,037.920	-3,117.333	-50.708	-7,489.769
	(-7.05)***	(-6.35)***	(-5.65)***	(-8.88)***	(-6.85)***	(-6.25)***	(-5.82)***	(-3.43)***	(-4.59)***	(-6.25)***
Observations	14070	14070	14070	14070	14070	2026	2026	2026	2026	2026
Log likelihood	-12720.49	-11178.98	-3171.37	-5419.72	-8057.65	-5098.42	-4507.08	-1403.96	-1736.56	-4254.83

Note: *z* statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Appendix Table 1. Industrial Distribution of Firms

Industry	No. of all sample firms	No. of firms with positive in-house or procured R&D											
		IRD		*CRD*		*CRDN*		*CRDA*		*JRD*		*TA*	
Food	1285	535	(41.6)	53	(9.9)	45	(8.4)	10	(1.9)	36	(6.7)	13	(2.4)
Beverages, tobacco and feed	281	148	(52.7)	25	(16.9)	23	(15.5)	5	(3.4)	10	(6.8)	6	(4.1)
Textile mill products	372	143	(38.4)	28	(19.6)	21	(14.7)	10	(7.0)	17	(11.9)	6	(4.2)
Apparel and other finished products	520	92	(17.7)	20	(21.7)	19	(20.7)	1	(1.1)	9	(9.8)	3	(3.3)
Lumber and wood products	151	28	(18.5)	3	(10.7)	3	(10.7)	0	(0.0)	7	(25.0)	1	(3.6)
Furniture and fixtures	176	78	(44.3)	10	(12.8)	10	(12.8)	1	(1.3)	5	(6.4)	3	(3.8)
Pulp, paper and paper products	379	78	(20.6)	6	(7.7)	4	(5.1)	2	(2.6)	12	(15.4)	7	(9.0)
Printing and allied industries	786	103	(13.1)	12	(11.7)	9	(8.7)	4	(3.9)	11	(10.7)	7	(6.8)
Chemical and allied products	868	699	(80.5)	264	(37.8)	243	(34.8)	47	(6.7)	129	(18.5)	115	(16.5)
Petroleum and coal products	60	35	(58.3)	13	(37.1)	13	(37.1)	3	(8.6)	7	(20.0)	6	(17.1)
Plastic products	651	278	(42.7)	53	(19.1)	47	(16.9)	12	(4.3)	49	(17.6)	36	(12.9)
Rubber products	157	93	(59.2)	15	(16.1)	11	(11.8)	4	(4.3)	15	(16.1)	9	(9.7)
Leather tanning, leather products and fur skins	47	22	(46.8)	8	(36.4)	8	(36.4)	0	(0.0)	2	(9.1)	0	(0.0)
Ceramic, stone and clay products	620	260	(41.9)	51	(19.6)	42	(16.2)	15	(5.8)	62	(23.8)	42	(16.2)
Iron and steel	339	99	(29.2)	19	(19.2)	18	(18.2)	1	(1.0)	17	(17.2)	16	(16.2)
Non-ferrous metals and products	287	123	(42.9)	28	(22.8)	22	(17.9)	9	(7.3)	18	(14.6)	18	(14.6)
Fabricated metal products	931	376	(40.4)	69	(18.4)	59	(15.7)	11	(2.9)	64	(17.0)	44	(11.7)
General machinery	1542	801	(51.9)	127	(15.9)	109	(13.6)	31	(3.9)	117	(14.6)	130	(16.2)
Electrical machinery, equipment and supplies	2099	1065	(50.7)	224	(21.0)	184	(17.3)	66	(6.2)	173	(16.2)	158	(14.8)
Transportation equipment	1419	626	(44.1)	129	(20.6)	111	(17.7)	31	(5.0)	87	(13.9)	104	(16.6)
Precision instruments and machinery	584	338	(57.9)	96	(28.4)	91	(26.9)	15	(4.4)	74	(21.9)	64	(18.9)
Ordnance and accessories	17	13	(76.5)	5	(38.5)	5	(38.5)	1	(7.7)	4	(30.8)	6	(46.2)
Miscellaneous	499	248	(49.7)	57	(23.0)	53	(21.4)	17	(6.9)	25	(10.1)	40	(16.1)
Total	14070	6281	(44.6)	1315	(20.9)	1150	(18.3)	296	(4.7)	950	(15.1)	834	(13.3)

Note: In parentheses are, for *IRD* , the percentage to the whole sample; and, for *CRD, CRDN, CRDA, JRD* , and *TA* , the percentages to the number of firms with positive *IRD* .

Appendix Table 2. Mean R&D Intensity (In-house and Procured) by Industry

	All Firms (n = 14070)						Firms with *RDD* = 1 (n = 6648)						
	IRD	*CRD*	*CRDN*	*CRDI*	*JRD*	*TA*	*IRD*	*CRD*	*CRDN*	*CRDI*	*JRD*	*TA*	*n*
Food	0.267	0.008	0.005	0.003	0.002	0.002	0.619	0.019	0.011	0.008	0.004	0.005	555
Beverages, tobacco and feed	0.351	0.013	0.006	0.007	0.001	0.013	0.653	0.024	0.012	0.013	0.001	0.025	151
Textile mill products	0.412	0.084	0.035	0.048	0.004	0.001	0.996	0.202	0.086	0.117	0.011	0.003	154
Apparel and other finished products	0.119	0.014	0.012	0.002	0.001	0.000	0.614	0.070	0.060	0.010	0.008	0.002	101
Lumber and wood products	0.063	0.003	0.003	0.000	0.001	0.000	0.296	0.015	0.015	0.000	0.006	0.002	32
Furniture and fixtures	0.253	0.012	0.012	0.001	0.001	0.010	0.564	0.027	0.026	0.002	0.003	0.022	79
Pulp, paper and paper products	0.168	0.005	0.001	0.004	0.002	0.017	0.748	0.021	0.005	0.016	0.008	0.077	85
Printing and allied industries	0.067	0.004	0.002	0.002	0.001	0.005	0.461	0.029	0.016	0.013	0.004	0.037	114
Chemical and allied products	2.995	0.263	0.221	0.041	0.007	0.096	3.640	0.319	0.269	0.050	0.008	0.116	714
Petroleum and coal products	0.873	0.028	0.025	0.003	0.003	0.043	1.218	0.039	0.035	0.004	0.004	0.061	43
Plastic products	0.571	0.030	0.018	0.013	0.003	0.010	1.256	0.067	0.039	0.028	0.007	0.023	296
Rubber products	1.209	0.050	0.023	0.027	0.008	0.007	1.957	0.081	0.037	0.044	0.013	0.012	97
Leather tanning, leather products and fur skins	0.703	0.042	0.042	0.000	0.003	0.000	1.501	0.089	0.089	0.000	0.007	0.000	22
Ceramic, stone and clay products	0.561	0.035	0.024	0.012	0.008	0.017	1.195	0.075	0.050	0.025	0.017	0.036	291
Iron and steel	0.219	0.003	0.002	0.001	0.002	0.004	0.694	0.008	0.006	0.002	0.006	0.012	107
Non-ferrous metals and products	0.490	0.044	0.034	0.010	0.003	0.031	1.050	0.095	0.073	0.022	0.006	0.065	134
Fabricated metal products	0.428	0.029	0.013	0.016	0.005	0.008	0.988	0.066	0.030	0.037	0.012	0.019	403
General machinery	0.911	0.039	0.019	0.020	0.005	0.034	1.654	0.070	0.035	0.036	0.008	0.061	849
Electrical machinery, equipment and supplies	1.275	0.073	0.047	0.026	0.004	0.030	2.394	0.138	0.089	0.049	0.008	0.055	1118
Transportation equipment	0.721	0.034	0.021	0.012	0.002	0.024	1.560	0.073	0.046	0.027	0.003	0.052	656
Precision instruments and machinery	1.713	0.073	0.051	0.022	0.008	0.048	2.748	0.117	0.082	0.035	0.013	0.076	364
Ordnance and accessories	1.969	0.041	0.027	0.014	0.001	0.108	2.575	0.054	0.035	0.019	0.002	0.142	13
Miscellaneous	0.819	0.057	0.042	0.016	0.003	0.040	1.513	0.106	0.077	0.029	0.006	0.075	270
Total	0.823	0.051	0.035	0.016	0.004	0.024	1.742	0.108	0.074	0.034	0.008	0.051	6648

Note: The figures show the percentages to sales of respective R&D variables, except n which is the number of firms with *RDD* >0 for each industry.

Appendix Table 3. Summary Statistics by Industry

Industry	*RDINT*	*LSALE*	*VI*	*DIV*	*CFS*	*PC*	*APP*
Food	0.003	8.266	0.253	0.131	0.030	0.231	0.2
Beverages, tobacco and feed	0.004	9.250	0.198	0.192	0.042	0.238	0.2
Textile mill products	0.004	7.774	0.334	0.109	0.043	0.237	0.2
Apparel and other finished products	0.001	7.581	0.311	0.097	0.018	0.244	0.4
Lumber and wood products	0.001	8.201	0.203	0.140	0.022	0.311	0.4
Furniture and fixtures	0.003	7.984	0.260	0.130	0.029	0.165	0.4
Pulp, paper and paper products	0.002	8.293	0.260	0.128	0.048	0.282	0.2
Printing and allied industries	0.001	8.174	0.351	0.079	0.049	0.167	0.2
Chemical and allied products	0.030	8.943	0.290	0.177	0.059	0.305	0.4
Petroleum and coal products	0.009	10.152	0.191	0.180	0.050	0.383	0.3
Plastic products	0.006	8.372	0.264	0.133	0.049	0.338	0.3
Rubber products	0.012	8.347	0.309	0.162	0.042	0.306	0.3
Leather tanning, leather products and fur skins	0.007	7.707	0.265	0.095	0.012	0.234	0.4
Ceramic, stone and clay products	0.006	8.178	0.301	0.171	0.050	0.276	0.2
Iron and steel	0.002	8.758	0.232	0.087	0.037	0.289	0.2
Non-ferrous metals and products	0.005	8.514	0.266	0.122	0.055	0.380	0.2
Fabricated metal products	0.004	8.174	0.301	0.118	0.044	0.215	0.3
General machinery	0.009	8.300	0.314	0.135	0.045	0.224	0.3
Electrical machinery, equipment and supplies	0.013	8.569	0.307	0.141	0.046	0.406	0.3
Transportation equipment	0.007	8.670	0.297	0.159	0.048	0.280	0.3
Precision instruments and machinery	0.017	8.414	0.327	0.228	0.045	0.301	0.3
Ordnance and accessories	0.020	10.289	0.319	0.324	0.050	0.176	0.3
Miscellaneous	0.008	8.519	0.277	0.216	0.037	0.242	0.3
Total	0.008	8.407	0.293	0.143	0.044	0.280	0.3

Notes: No. of observations = 14070.