The Economic Journal, **117** (*July*), 1122–1144. © The Author(s). Journal compilation © Royal Economic Society 2007. Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

WHY FIRMS FORM (OR DO NOT FORM) RJVS*

Lars-Hendrik Röller, Ralph Siebert and Mihkel M. Tombak

In this article, we examine why it is difficult to induce firms to form Research Joint Ventures (RJVs). We examine various incentives and disincentives for RJV formation by estimating an endogeneous switching model using data from the US National Cooperative Research Act. The empirical findings support hypotheses that firms of different sizes have disincentives to form RJVs and that cost-sharing is an important incentive for RJV participation.

In the early 1980s, a shift in technology policy took place in both the US and Europe. This was seemingly motivated by increased international competition, particularly from Japan, in high-technology sectors. Many scholars, policy makers and industrialists identified Japan's more cooperative business environment as a competitive advantage (Jorde and Teece, 1990). Japan's 1961 Act on the Mining and Manufacturing Industry Technology Research Association and the proactive efforts of MITI that encouraged joint ventures were identified as policy tools by which the Japanese created such a cooperative atmosphere. The response by US policy makers was to enact the 1984 National Cooperative Research Act (NCRA) and to provide government support for ventures such as SEMATECH.[1] In Europe, a block exemption for Research Joint Ventures (RJVs) was provided under EU Competition Law. Furthermore, billions of euros have been spent in framework programmes in the EU, which subsidises R&D undertakings only when different organisations work in collaboration with one another. Despite these policy changes, the number of RJVs registered under the NCRA has been quite small, ranging from 17 in 1986 to 67 in 1993 and averaging about 40 RJVs per year from 1985 to 1994. Thus the question: why is it so difficult to induce firms to form an RJV?

The goal of this article is to determine empirically what the incentives are for firms to participate or not participate together in RJVs. Our study is based on US data available through the NCRA. In order to obtain consistent evaluations of the hypotheses, we address two important empirical issues. Since much of the literature suggests a link between RJVs and R&D expenditures (see below), we must control for changes in R&D expenditures and its effect on RJV formation. As we can observe R&D changes at the time of RJV formation for only those firms that form an RJV, we have a missing data problem. To deal with this first issue, we need a consistent estimate of the expected effect on R&D expenses when RJVs are not formed. A second issue is one of simultaneity between the change in R&D expenditure and the decision to form an RJV. The decision to form an RJV may be determined by its impact on R&D expenses. R&D investments, however, are in turn also determined by RJV participation, which implies

---

* We thank Pedro P. Barros, Leonardo Felli, Dietmar Harhoff, Ken Hendricks, Morton Kamien, Ariel Pakes, F. Michael Scherer and Konrad Stahl for helpful comments and suggestions on a previous version of the article.

[1] Irwin and Klenow (1996) report that SEMATECH received hundreds of millions of dollars in federal government support in addition to facilities provided by the state of Texas.

an endogeneity between changes in R&D expenses and RJV formation. To address this concern, we estimate an endogenous switching model as proposed by Lee (1978).

The impact of the size distribution of firms on market performance has received increasing attention. In Section 1, we illustrate with a simple theoretical model the effect of firm asymmetries on RJV formation when firms are differentiated through their initial marginal costs. In this model, the effect of an RJV is to reduce marginal costs of the participating firms. It is shown that with no RJV, large firms have an incentive to spend more on R&D than small firms. Large firms thereby increase their market share and firm-size asymmetries increase. We show that firm heterogeneity has a significant impact on the incentives to form an RJV, as the market power of large firms increases in the absence of an RJV. Increased asymmetries involve not only increased profits for the large firm but also increased producers' surplus. As a result, large firms have a disincentive to becoming involved with small firms in an RJV. Rosen (1991), in his theoretical analysis, finds that firm-size differences can lead to a different emphasis on types of R&D. That is, Rosen finds that large firms tend to invest more but in less risky R&D projects. Small firms, however, invest more on the invention of radical new technologies. The implications of this analysis for RJVs is again that they tend not to be formed between differently sized firms.

The present study tests whether there is evidence that supports the firm-size and other hypotheses without imposing any of the special assumptions used in the above models. We do not estimate a structural model due to data limitations. To identify the impact of size heterogeneity on RJV formation, we test for other factors that are also hypothesised to affect RJV formation. One such factor is the effect on how much firms invest in R&D, as discussed extensively in the theoretical literature. Essentially, this literature focuses on two main rationales for RJV formation:

(1) Overcoming the free-rider problem, and
(2) Cost-sharing

The most influential papers are Katz (1986); d'Aspremont and Jacquemin (1988) and Kamien *et al.* (1992), hereafter KMZ.

A third possibility is introduced by Majewski and Williamson (2002), which is that RJVs may increase R&D expenditures due to complementarities between firms. One of the key elements examined in this literature is the free-rider effect: when R&D by one firm spills over to other firms, private incentives to conduct R&D are reduced. If firms form an all-inclusive RJV and cooperatively choose R&D investment levels, positive externalities through spillovers are internalised. This results in an increase in the effective R&D investments for all firms, and an RJV raises welfare. Contrary to the free-rider scenario, cost-sharing leads to a decrease in R&D investment for the individual firm. It is claimed that R&D cost-sharing can be quite substantial when it reduces duplication of effort.[2] Thus, the effect of an RJV on R&D spending can be either positive or negative. Whether the cost-sharing or the free-rider effect dominates in terms of their combined impact on firm-level R&D spending is ultimately an empirical question.

---

[2] This argument, however, does not consider a salient feature of R&D – that it is uncertain. Many independent trials can raise the probability of an invention occurring. In particular, Nalebuff and Stiglitz (1983) argue that the gains from competition in the form of lower risk and better incentives may more than offset the cost of duplicate research.

The implication of the above studies for the empirical testing of the hypotheses is that the probability of two firms forming an RJV is influenced in part by the expected change in R&D spending. The only scenario in which R&D spending is unaffected by RJV formation is when the free-rider, complementarities, and cost-saving effects exactly offset each other. Since this restriction is unlikely to hold in the data, R&D must be included in the analysis of RJV formation. In addition, there is the previously mentioned simultaneity of R&D spending and RJV formation. Finally, we control for the industries in which the firms are involved, accounting for the relatedness of technology as well as the degree of competition in the product market. When firms are from different industries, they are not confronted by a high degree of product market competition, and thus smaller changes in R&D spending may be a sufficient incentive for them to form an RJV. Furthermore, firms from different industries may be vertically linked, providing increased incentives to form RJVs due to complementary assets. We therefore control for industry effects in our analysis.

Our work is related to a growing empirical literature on RJV formation. Link (1996) provides an overview of the R&D joint ventures announced under the NCRA. Majewski and Williamson (2002) examine the contract details of 96 NCRA registrants. A number of studies also use European data sets on RJV formation, e.g., Cassiman and Veugelers (2002), Marin *et al.* (2000) and Kaiser (2002). Marin *et al.* (2000) study the determinants of RJV participation and find that R&D intensity, technological spillovers and firm size are significant in explaining RJV formation. Finally, Kaiser (2002) finds that in the German service sector firms engaging in cooperative research invest more than non-cooperating firms.

The remainder of the article is organised as follows. Section 1 formulates and analyses theoretical models of RJV formation. Section 2 describes the data and defines the variables used. Section 3 describes the empirical model and gives the results of the estimation. We conclude in Section 4.

1. A Theory of RJV Formation with Asymmetric Firms

In this Section, we illustrate the firm-size effect with a theoretical model. Here we show how the duopoly model of KMZ can be extended to illustrate the firm-size effect. KMZ show that symmetric firms have an incentive to form an RJV when firms engage in Cournot competition.[3] We show that asymmetries lead to a disincentive to form an RJV. The purpose of this Section is to provide a consistent theoretical argument to show how firm asymmetries affect the incentives to form an RJV.

1.1. *RJV Formation with Asymmetric Cournot Competition*

Following the approach by KMZ, we apply a three-stage game. In the first stage, firms decide on RJV participation. In the second stage, the partners determine their R&D investment (X), which reduces marginal costs by a function of the effective R&D

---

[3] We do not claim that the KMZ model is necessarily the most relevant one. Indeed, there are other approaches to modeling RJVs or spillovers. See, for example, Katsoulacos and Ulph (1998) and Beath *et al.* (1998).

investment $f(X)$. The effective R&D is the firm's own R&D investment when it is engaged in R&D competition, and it is the sum of the firms' R&D investments when they form an RJV. The third stage entails a Cournot product market game with homogenous products. We assume that the firms indexed by i and j have different *ex ante marginal costs* c_i and c_j, such that $c_i < c_j$. Given that R&D will reduce the marginal costs, we refer to $c_i - f(X)$ as the *ex post marginal costs of firm i*. We further assume that there are no fixed costs and that there is a linear demand structure given by $p = a - (q_i + q_j)$. Our assumption regarding the R&D production function and the profit functions that guarantee existence and uniqueness of the equilibrium are analogous to KMZ.

Firms' profit functions in stage three are $\pi_i = \{p - [c_i - f(X_i)]\}q_i$ (gross of R&D investment costs). The third stage Cournot quantities (q_i^*) and profits for a given X_i and X_j, can be easily solved ($q_i^* = \frac{1}{3}\{a - 2[c_i - f(X_i)] + [c_j - f(X_j)]\}$ and $\Pi_i^* = (q_i^*)^2 - X_i$ with an analogous expressions for firm j).[4]

We now examine R&D investments (stage 2) and first consider the case of R&D competition, in which firms decide on their individual R&D level (X_i) non-co-operatively. The effective level of cost-reducing R&D investment in this case is X_i. In other words, we assume that there are no spillovers.[5] Firms' objectives at this stage are to maximise their respective Cournot profit functions. The first-order condition for R&D investment derived from those profit functions for the firm of type i is

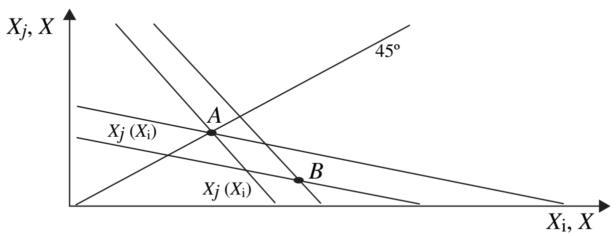
$$f'(X_i^N)q_i^* = 3/4 \tag{1}$$

with an analogous condition for firm j. Differentiating (1) with respect to X_j yields

$$\frac{\partial X_i^N}{\partial X_j} = \frac{f'(X_j)f'(X_i^N)}{3f''(X_i^N)q_i^* + 2[f'(X_i^N)]^2}$$

which is negative under the above assumptions and the second-order condition of X_i. This implies that R&D investments are strategic substitutes, which is graphically illustrated in Figure 1. For the case of symmetric *ex ante* marginal costs ($c_i = c_j$), (1) simplifies to

$$f'(X^A)[\alpha - c + f(X^A)] = 9/4 \tag{2}$$



A- Mean-preserving R&D Competition Equilibrium and RJV Equilibrium
B- Asymmetric R&D Competition Equilibrium

Fig. 1. *R&D Investments in R&D Competition and RJV*

---

[4] The derivations for these quantities and profit expressions are available upon request from the authors.

[5] This implies that the spillover parameter $\beta = 0$ in the KMZ model.

which implies that the equilibrium investments are identical, i.e., $X_i^N = X_j^N \equiv X^A$. The symmetric equilibrium is illustrated as point A in Figure 1, which is the equilibrium corresponding to the KMZ model.

We now address the issue of asymmetry and how it effects the equilibrium R&D investments. Suppose that the equilibrium is symmetric, i.e., at point A. Consider introducing a *mean-preserving* change in the *ex ante* marginal cost, such that firms' costs are $c_i + \varepsilon = c_j - \varepsilon$. In other words, a larger ε represents a greater asymmetry. To show how the asymmetry affects the reaction function of firm i, we need to implicitly differentiate (1) with respect to ε, which yields (after some manipulation)

$$\frac{\partial X_i^N(c_i, c_j)}{\partial \varepsilon} = \frac{\partial X_i^N}{\partial c_j} - \frac{\partial X_i^N}{\partial c_i} = -\frac{3}{2}\frac{\partial X_i^N}{\partial c_i} \tag{3}$$

where $\partial X_i^N/\partial c_i = 2f'(X_i^N)/\left\{3f''(X_i^*)q_i^* + 2\left[f'(X_i^*)\right]^2\right\} < 0$, which implies that firm i's reaction function shifts to the right as asymmetry increases. Similarly, firm j's reaction function shifts downwards with increased asymmetry, given that $\partial X_j^N\left(c_j, c_i\right)/\partial \varepsilon = -\partial X_i^N\left(c_i, c_j\right)/\partial \varepsilon$.

Using these features, we now can analyse the asymmetric equilibrium in R&D investments. Consider an asymmetric equilibrium depicted by point B in Figure 1. We find that the low-cost firm would invest more in R&D than the high-cost firm would, i.e., $X_i^N > X_j^N$. In addition, note that as ε increases, point B moves further to the bottom-right, which implies that the larger the *ex ante* asymmetry, the larger the asymmetry in R&D investments, i.e., $\partial X_i^N/\partial \varepsilon > 0$ and $\partial X_j^N/\partial \varepsilon < 0$. We can summarise these results in the following Lemma.

Lemma 1 *There is a positive relationship between the ex ante asymmetry in marginal costs and the asymmetry in equilibrium R&D investments, i.e.,* $\partial X_i^N/\partial \varepsilon > 0$ *and* $\partial X_j^N/\partial \varepsilon < 0$. *Endogenising R&D investments magnifies the asymmetric industry structure, i.e., the larger firm becomes even larger and the smaller firm relatively smaller.*

Given that R&D competition has such an impact on market structure, we now consider whether a similar effect exists if firms decide to form an RJV. Before examining the RJV case, it is useful to consider the mean-preserving symmetric analog to point B, which is denoted by point A in Figure 1. It is important to note that a change in ε does not affect point A, i.e., $\partial X^A/\partial \varepsilon = 0$ – see (2). In other words, a change in ε moves point B but not point A.

Now consider the R&D investment decisions when the two firms form an RJV. In this scenario, firms coordinate their R&D investments. The effective level of cost-reducing R&D investment is then $X = X_i + X_j$, which implies perfect spillovers. The industry profit function to be maximised jointly at this stage is $\Pi_i + \Pi_j$, i.e., firms coordinate. The first-order condition for R&D investment can be written as

$$f'(X^{JV})\left[a - \frac{(c_i + c_j)}{2} + f(X^{JV})\right] = 9/4. \tag{4}$$

The important aspect of the above expression is that R&D investments depend on the *average ex ante* marginal costs. Comparing the first-order conditions for the symmetric case (2) with the RJV case (4) shows that $X^A = X^{JV}$. In other words, we can depict the RJV equilibrium in Figure 1 by the same point *A*, i.e., the mean-preserving symmetric analog is identical in terms of effective R&D investments to the RJV case.[6] However, this does not imply that firms spend equal amounts, since in the RJV case firms can share their R&D expenses. Comparing points *A* and *B* yields the following Lemma:

Lemma 2 *Firms with higher marginal costs increase their effective R&D investment by participating in an RJV, while firms with lower marginal costs decrease their effective R&D investment, i.e.,* $X_i^N > X^{JV} > X_j^N$.

The above two Lemmas indicate that while a mean-preserving increase in asymmetry does not change the level of R&D investment in an RJV (i.e. X^{JV} is unaffected by ε) the R&D investments under R&D competition do change with ε. Comparing the equilibria in R&D investments, we find that the *ex ante* asymmetry in marginal costs is preserved when an RJV is formed, while the *ex ante* asymmetry is magnified when no RJV is formed. In other words, RJVs tend to keep market structure more symmetric. Since an asymmetric market structure benefits the larger (and more efficient) firm, the larger firm does not want to form an RJV whenever the rival is relatively small.

We now turn to stage 1, where firms decide whether or not to form an RJV by comparing equilibrium profits between the two cases (R&D competition and RJV). Recall from Π_i that profits are composed of product market profits minus R&D investments. Substituting the solutions for R&D investment decisions into the Cournot equilibrium profit functions, we can compare the product market incentives for firms to participate in an RJV. This leads to the following Lemma:

Lemma 3 *The RJV equilibrium yields higher product market profits for the smaller firm and lower product market profits for the larger firm.*

Proof. The difference in equilibrium pay-offs in the product market for firm *j* can be written as $\pi_j^{JV} - \pi_j^N = (q_j^{JV})^2 - (q_j^N)^2$. Thus, there is an incentive for firm *j* to participate in an RJV whenever $q_j^{JV} > q_j^N$, which implies that $f(X^{JV}) > 2f(X_j^N) - f(X_i^N)$. Given Lemma 2, the first part of this Lemma follows. Similarly, the condition for the large firm to have an incentive to join an RJV $\pi_i^{JV} - \pi_i^N > 0$ can be expressed as $f(X^{JV}) > 2f(X_i^N) - f(X_j^N)$, which does not hold by Lemma 2.

Lemma 3 indicates that the larger firm's product market profits decline if it participates in an RJV. This creates a disincentive for the large firm to participate in an RJV. So far, we have only considered profits in the product market. Whether the larger firm *de facto* participates in the RJV also depends on its R&D costs. This in turn will depend on how the participating RJV members decide to split their R&D costs. In particular, given Lemma 3, the smaller firm may have an incentive to pay a larger share of the R&D

---

[6] The reason for this is that we have assumed that the spillovers in the RJV are perfect, that there are no spillovers outside the RJV, and that products are perfectly homogeneous. For instance, if products are allowed to be less than perfect substitutes, then the effective RJV investments would be larger for the RJV case.

costs. As was first shown by Bergstrom and Varian (1985) – see also Salant and Shaffer (1998, 1999) – industry profits in a Cournot product market are increased with a larger mean-preserving asymmetry. This implies that the RJV equilibrium leads to lower producer surplus (relative to the R&D competition case) whenever there is enough asymmetry. This in turn implies that there is no R&D budget sharing rule that yields a higher payoff for both firms in an RJV. We therefore find:

LEMMA 4 *If the ex ante asymmetry is relatively large, no RJV is formed.*

Since industry profits are lower under RJV formation (i.e., the large firm loses more than the small firm gains) whenever the *ex ante* asymmetry is high, no mutually beneficial RJV formation exists. The role of asymmetry in this context is worth emphasising. Without asymmetry, the RJV equilibrium must lead to higher industry profits, since collusion and cost-sharing can only increase pay-offs. The only factor leading to lower industry profits under RJV is the asymmetry, which in turn produces the no-RJV equilibrium outcome. The crucial aspect of the model is therefore that an RJV leads to less asymmetry than R&D competition does.

The model can be easily extended to firms having complementary products. In this case, RJVs are more likely to be formed. The same analysis as that above using an asymmetric differentiated Bertrand setting has been conducted with similar results.[7] Furthermore, the theory of RJVs has been expanded to n independent firms under the rubric of endogenous coalition formation (Bloch, 1995; Cassiman and Greenlee, 1999; Yi and Shin, 2000) for symmetric firms under various formation rules. In particular, Bloch (1995) found that for Cournot competitors the incentives to admit a new member decreases with the size of the RJV. Our model shows that asymmetries reduce the likelihood of two firms getting together in an RJV. Bloch's result implies that RJV formation becomes even less likely in markets with larger numbers of firms. Boone (2000) introduces the notion of competitive pressure in the analysis of incentives to conduct R&D of various types. These incentives to invest in R&D could, in turn, affect the motivation to join in an RJV. Whether and under which conditions firms join in an RJV is ultimately an empirical question which we now address.

2. Firm Size and RJV Formation

In the next Sections, we present empirical evidence regarding the hypotheses developed in the previous theoretical literature, and derived from our extension of the KMZ model. Summarising the theory developed above and elsewhere in the literature, we can derive the following conditions that tend to promote the formation of RJVs:

(1) when R&D spillovers create free-rider problems,
(2) when duplicative R&D efforts create opportunities for cost-sharing, and
(3) when firms are producing complementary products.

We also find that an RJV will tend not to be formed by differently sized firms when:

---

[7] This Bertrand analysis is available upon request from the authors. Lemmas 1,3 and 4 hold with Lemma 2 being modified to $X_i^N > X_j^N > X^{JV}$.

(4) it prevents large firms from gaining the market power from competitive R&D,
(5) the firms focus on different types of R&D, and
(6) spreading the benefits of cost-reducing new technologies to these asymmetric firms puts more downward pressure on prices.

We take these hypotheses to the data, making use of a database available through information made public under the NCRA. As we do not have data on the types of R&D nor on prices, we cannot test hypotheses (5) and (6). We have, however, data on R&D expenditures and thus can test whether R&D savings are inducements to RJV formation or whether R&D investment increases result from RJVs. We estimate a two-equation system that endogenises RJV formation and R&D investments.

We estimate the probability of two firms joining an RJV in terms of the relative difference in firm size. As we focus mainly on the incentive of RJV formation, we need not evaluate whether the filed RJVs have been successful. We primarily investigate the determinants of partner selection. To test the hypotheses, we examine a number of factors of RJV formation that have been studied in the literature. Among these are internalising spillovers (i.e., the free-rider effect) and cost-sharing, as well as industry and technological effects. To account for changes in R&D investments due to RJV formation, we need to specify the simultaneity as there is causality running both ways (i.e., changes in R&D spending affects RJV formation and *vice versa*). We also need to address the issue of missing data. That is, we observe changes in R&D due to an RJV only when firms are *actually* engaged in an RJV. To account for this endogeneity of R&D spending and the missing data problem, we estimate an endogenous switching model as proposed by Lee (1978).

Our data set consists of firm-level and RJV-type information that spans a number of industries. Given the sparse occurrence of RJVs in any one industry, it is not possible to perform a more structural industry-specific empirical analysis. In particular, we are not able to estimate a model in which all three stages of the theoretical model are endogenised, i.e., RJV participation, R&D investment and product market competition. That is, we test the comparative statics of the above theory with reduced form regressions of RJV participation and R&D investment. Also note that the formation of an RJV is a relatively rare event. In this instance, the rare event results from the fact that we have many more firm-pairs that do not form an RJV than firm-pairs that do form one. We test for a potential selectivity bias and estimate a Rare Events Logit using the procedure in Tomz *et al.* (2003).

We are able to test for other important determinants of RJV formation, such as the degree of substitutability or complementarity and the degree of competition in the final product market. In the context of the KMZ-like model, it is easily shown that RJVs between firms that are in complementary industries are more likely to occur. Unable to produce demand cross-elasticities for all the industries, we control for these demand effects by industry and industry pair dummies. Finally, there may be technological spillovers between RJV partners. For instance, there may be higher spillovers between certain types of product categories. Again, we are able to control for these through pairwise industry fixed effects.

2.1. *Data Sources: The Research Joint Ventures Act*

Our analysis requires data from a variety of sources. Using a report published by the US Department of Commerce (1993) and additional filings published in the Federal Register, we obtained the identities of the firms involved in an RJV, the date of formation of the RJV, and the general nature of the proposed research. Our basic data on RJVs runs from January 1985 through July 1994. It is worth noting that 59% of the RJVs focus on the development of new processes, whereas only 36% deal primarily with product innovations – see the RJV classification by Link (1996). Furthermore, many of the product technologies result in cost reductions. This supports the assumption in the previous Section that R&D expenditures result in marginal cost reductions. For a more detailed description of the RJV-filings, see also Link (1996). Note that our database is complementary to the database constructed by Link under sponsorship of the National Science Foundation (NSF). However, beyond the information contained in the NSF database, our dataset identifies the participating firms in the RJV. This enables us to take into account firm-level characteristics. For more descriptive statistics regarding the ownership of RJVs filed under the NCRA, see Majewski and Williamson (2002). Moreover, an early study by Berg *et al.* (1982) provide interesting descriptions of the organisation and composition of Joint Ventures registered by the Federal Trade Commission.

We use the identity of the RJV firms to crosslink the RJV database with other firm-specific data obtained from Moody's company database (1995), which contains information on 17,785 firms based on financial reports and the business press. The database includes information on public, US, and international companies, representing 95% of the non-US global market capitalisation. Since the company data we require is complete from 1989 onwards, the process of merging both databases allows us to use the data from 1990 to 1994.

2.2. *Variable Definitions and Descriptive Statistics*

In the period 1990 to 1994, we have a total of 257 RJVs in our database (see Table 1 for some trends in RJV formation during this period). This is similar to the number of RJVs

Table 1

Summary Statistics Over Time

	RJV database	
Year	No. of RJVs	Avg. size (no. of firms) per RJV
1990	45	11
1991	60	14
1992	58	10
1993	67	7
1994	27	6
Total	257 (274)*	9.6

*The numbers of RJVs mentioned in parenthesis are taken from the National Science Foundation database constructed by Link (1996).

Table 2

Summary Statistics by Industry

	Our subsample				Moodys		
Industries 2-digit SIC-Codes	No. of firms	Avg. sales	Avg. R&D	$\Delta r\&d$*	No. offirms	Avg. sales	Avg. R&D
28 Chemicals and Allied Products	60	7858.7	603.5	-0.004E^{-03}	1048	1052	114
32 Stone, clay, glass, and concrete products	5	3439.9	149.2	-0.007E^{-04}	556	597	31
33 Primary metal industries	23	3954.1	51.6	−0.001	471	1397	31
35 Industrial Machinery/Equipment	123	17250.7	1742.3	−0.004	999	757	61
36 Electronic and other Electric Equipment	96	11642.3	690.3	−0.004	1099	939	38
37 Transportation Equipment	120	43465.1	742.1	0.002	402	4244	162
38 Instruments and Related Products	32	5792.4	491.3	0.0004	691	394	21

Note. Monetary numbers are in million USD and refer to 1992. The Standard Industrial Classifications refer to the 1987 SIC-Revision.

*The variable $\Delta r\&d$ represents the annual change in R&D intensities (R&D expenditures divided by sales) before and after participating in an RJV.

in the NSF database (274). The slight difference may be due to our time series ending in July 1994 whereas the NSF database covers the whole of 1994. Thus, we find fewer RJVs in our database in 1994 as compared with the previous years. However, the number of RJVs has increased over time. Most of the RJVs consist of 5–10 participants, with each firm participating in an average of three RJVs.

After merging the RJV and company databases and eliminating firms with missing values, we are able to retain 459 unique firms that have participated in an RJV. Table 2 shows some descriptive statistics regarding the participating firms by industry and compares them to all companies included in the Moody's Global Company Database. The average sales and R&D expenditures are higher in our subsample, suggesting that larger firms are more likely to register RJVs. Firms participating in an RJV are not required to file under the NCRA. Since smaller firms are less likely to be the subject of an anti-trust investigation, it may be that an RJV consisting entirely of small firms is less likely to file. Moreover, smaller firms are often not reported in Moody's Global Company Database or they may not report R&D expenditures. Therefore, a potential defect of our sample may be that smaller firms are not represented to the same extent as larger firms.

Hence, we have a nonrandom selection of firms and lose around 70% of observations when forming firm-pairs. This selection type enters our predictions on R&D investments (switching equations) in our empirical model and may have an impact on our cost-sharing versus free-rider hypothesis. Note, however, that this selection problem occurs on both groups, e.g., RJV-partners and non-RJV partners. Therefore, we should not have any single-sided effects between both groups that change the results significantly. Moreover, this type of selection also has an impact on estimating the participation in RJVs (selection equation), which is related to the problem that rather larger firms will register with the NCRA. This type of sampling, however, is not considered to have any serious consequences on our firm heterogeneity hypothesis as this selection problem forms a more homogenous subpopulation of firms (large firms are more

Table 3

RJV Objectives Classified by 2-digit SIC-Codes

Industries 2-digit SIC-Codes	Intra-Industry* no. of RJVs	Inter-Industry* no. of RJVs
28 Chemicals and Allied Products	14	1
32 Stone, clay, glass, and concrete products	0	3
33 Primary metal industries	3	4
35 Industrial Machinery/Equipment	8	23
36 Electronic and other Electric Equipment	7	55
37 Transportation Equipment	11	12
38 Instruments and Related Products	5	4

*We refer to intra (inter)-industry RJV, if the objective of the RJV could be classified by one (more than one) 4 digit-SIC Code.

frequently selected) than represented in the underlying population. This sample selection bias may only make our estimates more conservative (e.g., we observe that firm-size differences are important among large firms even though the majority of firms are of large size). The firm size differences may also be substantial in complementary industries that are dominated by small firms which are underrepresented in the data. Consequently, our subsample should be an appropriate sample in order to test our hypotheses.

It is interesting to note that the difference in R&D expenditures over sales before and after joining an RJV is negative in most of the industries, indicating that R&D intensities increase due to RJV formation. The 'Transportation Equipment' (SIC37) and the 'Instruments and Related Products' (SIC38) industries, however, seem to be characterised by a reduction in R&D intensities. Thus, it is important to control for industry in our estimation of R&D intensities.

In the next step, we classify the RJVs into 4-digit SIC codes according to their objectives. Whenever the objective of the RJV is classified into a unique SIC-code, we refer to an intra-industry RJV, as the technology is very narrowly defined. Whenever the RJV-objective is defined by more than one 4-digit SIC code, we refer to an inter-industry RJV, as the technology is more broadly defined. Table 3 shows the number of inter and intra-RJVs for the different SIC codes. It is interesting to note that the number of intra-industry RJVs is relatively large for 'Chemicals and Allied Products' (SIC28). Inter-industry RJVs, on the other hand, are more prevalent in 'Industrial Machinery/Equipment' (SIC35) and 'Electronic and other Electric Equipment' (SIC36).

To investigate the choice of RJV partners, we begin by matching all firms into pairs. Note that some firms are characterised by missing values for some of the variables over time. Eliminating firms with missing values gives us 31,817 firm-pairs out of the possible 105,111 firm-pairs.[8] Table 4 reports the industries in our database and the

---

[8] We lose 70% of our data when forming firm-pairs. This is due to the fact that all registered firms from the NCRA filings are matched with each other. We lose those observations of firms that are registered as RJVs but that are not included in the Moody's database. We mostly lose small firms that are not publicly traded or quite small such that they are not included in the Moody's database. Consequently, we have a truncation problem of firms with respect to their size which results in a more homogenous sample compared to the underlying population of firms. The implications of this for our analysis are discussed above.

Table 4

Sample Frequencies of Industry-pairs (%)

Industries 2-digit SIC-Codes	28 Chemicals and Allied Products	32 Stone, clay, glass, and concrete products	33 Primary metal industries	35 Industrial Machinery/Equipment	36 Electronic and Electric Equipment	37 Transport. Equipment	38 Instruments and Related Products
28 Chemicals and Allied Products	2.18						
32 Stone, clay, glass, and concrete products	0.55						
33 Primary metal industries	0.94		0.13				
35 Industrial Machinery/Equipment	14.12	1.69		20.61			
36 Electronic and other Electric Equipment				16.99	3.78		
37 Transportation Equipment	5.73	0.72	1.39	18.32		3.56	
38 Instruments and Related Products		0.36			4	3.94	1.01

sample frequencies (mean of the dummies) for each one of the industry-pairs. The Table shows six intra-industry dummies (nonzero elements on the diagonal) and twelve complementarity dummies (nonzero off-diagonal elements). It is noteworthy that in more than 50% of all RJVs in our sample, one firm is from the 'Industrial Machinery and Equipment' (SIC35) industry. Since 'Industrial Machinery and Equipment' is a necessary input for many other industries, this observation is consistent with the hypothesis that RJVs occur more often when products are complementary.

We now introduce our variables. We define variable P_{ijt} $(i \neq j)$ as a dummy variable, indicating whether the matched pair of firms i and j is participating in an RJV at time period t. To test our theory presented in Section 2 we check for a firm size difference effect. Relative firm size difference is defined as

$$DASSET_{ijt} = \frac{|ASSET_{it-1} - ASSET_{jt-1}|}{\max(ASSET_{it-1}, ASSET_{jt-1})}$$

where $ASSET_{it-1}$ is the assets of firm i taken in year $t-1$. In other words, we define *DASSET* as the absolute value of the difference in the firms' assets as a proportion of the larger firm's assets.

As the Schumpeterian hypothesis suggests an essential impact of industrial concentration on R&D expenditures, we control for concentration ratios among different industries and correct for firm-size distributions and competition between different industries by constructing an average Herfindahl index *(HERF)*. When the two firms are in different industries, we take the average Herfindahl index of the two industries.

We control for cost-sharing or free-riding and construct a measure of changes in firm-level R&D. We define $\Delta r\&d$ as the change in average firm-level R&D intensities as follows:

$$\Delta r\&d_{ijt} = \frac{1}{2}(r\&d_{i,t-1} - r\&d_{i,t} + r\&d_{j,t-1} - r\&d_{j,t})$$

where $r\&d_{it}$ is the firm-level R&D investment divided by firm-level revenue in year t. In other words, $\Delta r\&d$ measures whether the growth in R&D intensities increased or decreased. If $\Delta r\&d$ is negative, the growth in R&D intensities increases. If $\Delta r\&d$ is positive, the growth in R&D intensities declines.

Our variable for the average change in firm-level R&D intensities can be expressed in terms of the variables from our theoretical model as follows:

$$\Delta r\&d_{ijt} = \frac{1}{2}\left[\frac{X_i^N}{p^N q_i^N} - \frac{aX^{JV}}{p^{JV} q_i^{JV}} + \frac{X_j^N}{p^N q_j^N} - \frac{(1-a)X^{JV}}{p^{JV} q_j^{JV}}\right]$$

where, a is the proportion of R&D investment firm i contributes to the RJV. How firm size differences can affect the above variable can be analysed by taking the derivative of the above with respect to our exogenous variable the initial marginal cost differences (ε)

$$\frac{\partial(\Delta r\&d_{ijt})}{\partial\varepsilon}=\frac{1}{2}\left\{\frac{\frac{\partial X_i^N}{\partial\varepsilon}}{p^N q_i^N}-\frac{X_i^N\frac{\partial(p^N q_i^N)}{\partial\varepsilon}}{(p^N q_i^N)^2}+\frac{\frac{\partial X_j^N}{\partial\varepsilon}}{p^N q_j^N}-\frac{X_j^N\frac{\partial(p^N q_j^N)}{\partial\varepsilon}}{(p^N q_j^N)^2}\right.$$
$$\left.-X^{JV}\left[\frac{a\frac{\partial(p^N q_i^N)}{\partial\varepsilon}}{(p^N q_i^N)^2}+\frac{(1-a)\frac{\partial(p^N q_j^N)}{\partial\varepsilon}}{(p^N q_j^N)^2}\right]\right\}.$$

As shown in Section 1.1, $\partial X_i^N/\partial\varepsilon > 0$, and $\partial X_i^N/\partial\varepsilon = \partial X_j^N/\partial\varepsilon$. We also know that $p^N q_i^N > p^N q_j^N$. Consequently, the direct effect of marginal cost differences $\left(\frac{\partial X_i^N/\partial\varepsilon}{p^N q_i^N}+\frac{\partial X_j^N/\partial\varepsilon}{p^N q_j^N}\right)$ would be negative. The indirect effect of marginal cost differences, however, depend on the sharing proportion of the R&D investment and on the specific parameters and type of our theoretical model. For example, with the linear demand function and Cournot model of Section 1.1, it follows that $p^{JV}q_i^{JV} = \frac{1}{9}(a+c)(a-c+3\varepsilon)$, $p^{JV}q_j^{JV} = \frac{1}{9}(a+c)(a-c-\varepsilon)$, $\partial(p^N q_i^N)/\partial\varepsilon = \frac{1}{3}(a+c)$, and $\partial(p^N q_j^N)/\partial\varepsilon = -\frac{1}{9}(a+c)$ where c is the average marginal cost. Applying Lemma 2 that $X_i^N > X^{JV} > X_j^N$, this implies that the indirect effect is also negative if $(a-c-\varepsilon)/(a-c+3\varepsilon) > 1/\sqrt{3}$ and $(a-c-\varepsilon)/(a-c+3\varepsilon) > \sqrt{(1-a)/3a}$.

In other words, so long as the cost difference is sufficiently small and the proportion that firm i contributes to the budget is sufficiently large. Thus, our theoretical model suggests that there would be an effect of *DASSET* on our change in R&D intensity $\Delta r\&d_{ijt}$. The sign of that effect, however, depends (among other things) on variables such as the R&D cost-sharing rule that is used by the RJV which is unobserved. In addition, we control for the size of the RJV as larger RJVs may allow for greater heterogeneity between firms. It seems more likely that two firms would be together in an RJV if that RJV were more inclusive. Hence, we add the variable *MEMBERS*, which is defined as the number of members in the RJV. In general, one might expect an effect on the variable $\Delta r\&d$, the sign of that effect depending on whether cost-sharing or free-rider effects dominate. If cost-sharing dominates as the rationale for the RJV then more *MEMBERS* means that each firm in the RJV can reduce its R&D intensity and *MEMBERS* would have a positive effect on $\Delta r\&d$. If, however, free-rider effects dominate then a larger membership implies less free-riders and a greater internalisation of the spillovers leading to greater R&D investment. To illustrate these arguments we introduce *MEMBERS* into our variable for the average change in firm-level R&D intensities to obtain:

$$\Delta r\&d_{ijt}=\frac{1}{2}\left(\frac{X_i^N}{p^N q_i^N}-\frac{X^{JV}/MEMBERS}{p^{JV}q_i^{JV}}+\frac{X_j^N}{p^N q_j^N}-\frac{X^{JV}/MEMBERS}{p^{JV}q_j^{JV}}\right)$$

if we assume that the costs of R&D are shared equally among the RJV members. Setting $MEMBERS = M$ and $p^{JV}q_k^{JV} = r_k^{JV}$ for $k = i$ or j the derivative of our change in the $\Delta r\&d$ intensity with respect to *MEMBERS* is

$$\frac{\partial(\Delta r\&d_{ijt})}{\partial M} = \frac{-1}{2M}\left\{\left(\frac{\partial X^{JV}}{\partial M} - \frac{X^{JV}}{M}\right)\left(\frac{1}{r_i^{JV}} + \frac{1}{r_j^{JV}}\right) - X^{JV}\left[\frac{\partial r_i^{JV}/\partial M}{(r_i^{JV})^2} + \frac{\partial r_j^{JV}/\partial M}{(r_j^{JV})^2}\right]\right\}$$

The above shows that an effect of *MEMBERS* exists and that it is positive if $\partial X^{JV}/\partial M < X^{JV} + (X^{JV}M/r_k^{JV})(\partial r_k^{JV}/\partial M)$ for both $k = i$ and j, i.e., it depends on the effect of *MEMBERS* on the amount invested by the RJV in R&D and on its revenue effects. Given that the variable $\Delta r\&d$ is not observed for firm-pairs that do not engage in an RJV, we proxy *MEMBERS* by taking the average size of all other RJVs in which the firms are engaged.

Another control variable is the total RJV activities by the firm-pair under consideration. The variable *FRJVS* equals the number of other RJVs in which the firms are engaged

$$FRJVS_{ijt} = \frac{1}{2}[(FRJVS_{it} - 1) + (FRJVS_{jt} - 1)].$$

We also use a set of dummy variables to control for intra and inter-industry effects. We define industry dummies (*SICs*) as having a value of one if the two firms under consideration are in the same major *SIC* industry group; otherwise, the value is zero. In addition, we define inter-industry dummies (*COMPs*) as indicating firms that are from different industries. They take on a value of one if the two firms under consideration are from different major industry groups; otherwise, the value is zero. In the empirical analysis below, we will interpret the *COMP* dummy as an indicator of how closely related the products are on the technology dimension. Note that SIC classifications are often based on cost-side considerations, i.e., they are technology oriented and not demand-side-oriented.[9] Finally, we introduce time dummies in order to account for the panel data structure in our dataset.

The statistics of the variables we use in the estimation are given in Table 5. This Table shows descriptive statistics for firm-pairs in the groups for cooperating and non-cooperating firms. There are a total of 496 cases in which the firms in a firm-pair are cooperating together in an RJV, and there are 31,321 firm-pairs in which the firms do not cooperate. We use those firm-pairs as a control group. The advantage of using a control group of firms that have engaged in an RJV but not with the firm under current consideration is that we select firms conditional on having joined an RJV instead of having never been engaged in any RJV. We thus avoid selection biases between two very different groups of firms, one group which may, as a matter of policy, not participate in any RJV. Overall, the variable *DASSET* has a mean of 0.73, indicating that there is a large amount of heterogeneity in firm-size between paired firms, which guarantees a fair amount of variance in firm-size in our sample of all firms. Table 5 also reveals several interesting features in our data. First, the non-RJV pair group has a lower mean on *DASSET* than the RJV pair group. This finding seems to reject our conjecture that

---

[9] As usual, there may be relevant variables for the formation of RJVs that have been excluded from the empirical analysis due to a lack of measures or data. In addition to financial risk and organisational variables already mentioned, there may be other factors. KMZ, for example, have identified the organisation of the RJV as an important variable. Geographic location of the partners may be another variable affecting RJV formation. These variables may be correlated with some of the variables that have been included (e.g., the organisation of the RJV may be correlated with the number of members).

Table 5

Statistics for Firm-pairs in Both Groups

		Group of non RJV-pairs				Group of RJV-pairs			
Variables	Description	*N*	Mean	Min.	Max.	*N*	Mean	Min.	Max.
MEMBERS	Number of members in an RJV (see text for precise definition).	31,321	29.43	2	138	496	53.39	2	138
HERF	Herfindahl Index of the firm's primary industry classification.	31,321	132.60	20.5	1545	496	59.39	26	501
DASSET	Measure of firms' difference in assets prior to forming an RJV.	31,321	0.73	0	0.99	496	0.75	0	0.99
FRJVS	Number of further RJVs undertaken by firms.	31,321	3.66	0	23.5	496	2.87	0	17
$\Delta r\&d$	The change in firm-level R&D intensities by forming an RJV.	31,321	−0.003	−0.39	0.04	496	−0.007	−0.38	0.03

The monetary data are measured in million USD in current prices and are deflated by the producer price index taken from the Main Economic Indicators (OECD).

cooperation tends to occur among similarly sized firms. The Herfindahl index is higher for the noncooperating than for the RJV pair group. Consequently, we control for heterogeneity in industry concentration across industries by adding the Herfindahl index to our empirical analysis. The variable $\Delta r\&d$ measures the change of R&D intensities over time for a firm-pair. This variable captures the evolution (growth) of R&D intensities. $\Delta r\&d$ is negative when the free-rider and complementarity effects dominate and positive when the cost-sharing effect is larger. It is interesting to note that the average change in firm-level R&D intensities before and after forming an RJV is negative but is three times higher for the non-RJV firm-pairs than for RJV firm-pairs. This suggests that the free-rider and complementarity effects are more significant for cooperating firms. They seem to overcome free-rider problems as their growth in R&D intensities is higher such that they increase their R&D investments.

3. Econometric Specification and Estimation

In order to investigate our hypotheses, we estimate a probit equation that determines whether a firm-pair forms an RJV.[10] As the theory suggests effects of firm-size, R&D investment, industry concentration, and industry effects (whether products are substitutable or complements), we control for these variables. Moreover, as we endogenise R&D investments in our theoretic models, we do so also in our econometric specification below. Thus the latent variable

---

[10] The decision process by which firms choose their RJV partners may be more complicated than a simple probit model suggests. The fact that RJVs are composed of many firms suggests a more sophisticated model where the decision to participate in an RJV depends on which and how many other firms are willing to join (Bloch, 1995; Cassiman and Greenlee, 1999; Yi and Shin, 2000).

$$\begin{aligned}P_{ijt}^{*} &= \delta_1 DASSET_{ijt} + \delta_2 HERF_{ij} + \delta_3 DRDJV_{ijt} + \delta_4 MEMBERS_{ijt} \\ &\quad + \delta_5 FRJVS_{ijt} + \sum_{k=1}^{6} \delta_6^k SIC_{ij}^k + \sum_{l=1}^{12} \delta_7^l COMP_{ij}^l + \sum_{m=1}^{5} \delta_8^m TIME_{ij}^m + \varpi_{ijt}.\end{aligned} \tag{5}$$

indicates whether a firm-pair forms an RJV or not. P_{ijt} is the dummy variable introduced in the previous Section where $P_{ijt} = 1$ iff $P_{ijt}^{*} > 0$ and $P_{ijt} = 0$ iff $P_{ijt}^{*} \leq 0$. As suggested by much of the literature on RJVs, the incentive to form an RJV should depend on the expected effect that the RJV formation will have on R&D expenditures. We therefore include the variable $DRDJV_{ijt} = \Delta r\&d_{ijt}(\text{when } P_{ijt} = 1) - \Delta r\&d_{ijt}(\text{when } P_{ijt} = 0)$ in (5), which measures the difference in R&D intensity growth due to an RJV. If *DRDJV* is positive (negative), the cost-sharing (free-rider) effect dominates the other. As discussed above, the firm-size hypothesis implies that the variable *DASSET* has a negative impact on the probability of forming an RJV ($\delta_1 < 0$). The theory of Bloch (1995) suggests that the effect of the size of the RJV is negative ($\delta_4 < 0$). In addition, we control for concentration ratios among different industries *(HERF)* and total RJV activities by the firm-pair *(FRJVS)*. Finally, the dummies SIC_{ij}^k and $COMP_{ij}^l$ control for intra and inter-industry effects. In the empirical analysis below, we interpret the *COMP* dummy as an indicator of how related the products are on the technology side. The *TIME* dummies capture fixed effects in the propensity to engage in RJVs over time. The error term ϖ is supposed to be normally, identically and independently distributed.

We are not able to obtain consistent estimates unless we address two important empirical issues. First, we observe $\Delta r\&d_{ijt}(\text{when } P_{ijt} = 1)$ only when firms are *actually* engaged in an RJV. In this case, we can construct a measure of the R&D effect from our observed changes in R&D. Whenever a firm-pair does not form an RJV, we have no such measure of $\Delta r\&d_{ijt}(\text{when } P_{ijt} = 1)$. This also applies to some observations of $\Delta r\&d_{ijt}(\text{when } P_{ijt} = 0)$. We consequently have a missing data problem and need a consistent estimate of the expected effect on R&D expenditures for the unobserved events.

A second issue is one of simultaneity between changes in R&D expenditures and the decision to form an RJV. As is specified in (5), the decision to create an RJV is determined by its impact on R&D (i.e., free-rider and cost-sharing effects). However, R&D investments are in turn also determined by RJV formation, which implies that *DRDJV* is endogenous in (5). Not accounting for endogeneity in (5) leads to its inconsistent estimation.

Given these two concerns, we estimate our model using a switching model originally suggested by Lee (1978). The endogenous switching model can be written as (5) and

$$\begin{aligned}\Delta r\&d_{ijt} &= \alpha_1 DASSET_{ijt} + \alpha_2 HERF_{ij} + \alpha_3 MEMBERS_{ijt} + \alpha_4 FRJVS_{ijt} \\ &\quad + \sum_{k=1}^{6} \alpha_5^k SIC_{ij}^k + \sum_{l=1}^{12} \alpha_6^l COMP_{ij}^l + \sum_{m=1}^{5} \alpha_7^m TIME_{ij}^m + v_{ijt} \quad \text{if } P_{ijt} = 1,\end{aligned} \tag{6}$$

$$\begin{aligned}\Delta r\&d_{ijt} &= \beta_1 DASSET_{ijt} + \beta_2 HERF_{ij} + \beta_3 MEMBERS_{ijt} + \beta_4 FRJVS_{ijt} \\ &\quad + \sum_{k=1}^{6} \beta_5^k SIC_{ij}^k + \sum_{l=1}^{12} \beta_6^l COMP_{ij}^l + \sum_{m=1}^{5} \beta_7^m TIME_{ij}^m + \varepsilon_{ijt} \quad \text{if } P_{ijt} = 0.\end{aligned} \tag{7}$$

Table 6

Sources and Complementarities in RJV Formation and R&D Intensities

Variables	Selection Equation*		Equation (6)**		Equation (7)**		Counterfactual Probit	
	Dep. Var: P_{ij}		Dep. Var: $\Delta r\&d1$		Dep. Var: $\Delta r\&d0$		Equation (5)	
	Param.	Std. Err.	Param.	Std. Err.	Param.	Std. Err.	Param.	Std. Err.
MEMBERS	0.002	0.001	0.003	0.0004	−0.03	0.017	−0.026	0.002
DASSET	−1.17	0.08	−0.76	0.04	−5.09	2.09	−6.57	0.31
HERF	−0.03	0.0004	0.001	0.0002	−0.14	0.05	−0.18	0.01
FRJVS	0.02	0.01	0.47	0.007	0.54	0.12	0.52	0.03
DRDJV							1.21	0.06
D90	0.65	0.11	−1.63	0.03	1.19	2.49	6.54	0.37
D91	0.94	0.10	0.38	0.07	2.14	1.71	4.84	0.29
D92	1.52	0.11	0.49	0.06	2.88	2.37	6.42	0.35
D93	1.13	0.11	0.81	0.14	3.79	1.95	6.78	0.38
SIC28	−0.89	0.15	−1.18	0.04	−1.62	2.43	−0.37	0.14
SIC33	1.13	0.33	0.61	0.82	7.44	20.03	10.96	0.63
SIC35	−0.59	0.08	−0.31	0.06	−2.95	1.30	−2.61	0.15
SIC36	−0.76	0.13	0.17	0.12	−1.29	2.06	−1.72	0.15
SIC37	13.02	0.46	−0.08	0.27	58.07	22.92	73.18	3.49
SIC38	0.54	0.40	0.07	0.46	3.71	11.12	4.20	0.41
COMP2832	−0.76	0.42	0.08	0.59	−3.17	11.57	−3.89	0.36
COMP2833	−0.34	0.36	−0.79	0.05	1.43	11.93	2.79	0.28
COMP2835	−2.15	0.37	−0.77	0.03	−8.66	7.27	−11.44	0.53
COMP3235	−1.26	0.63	−0.14	0.25	−5.54	8.30	−7.26	0.39
COMP3237	5.89	0.31	0.02	0.43	26.66	12.19	33.35	1.63
COMP3337	6.62	0.22	0.08	0.30	32.54	12.78	40.35	1.94
COMP3536	−0.72	0.08	−0.12	0.06	−2.64	1.31	−3.04	0.16
COMP3537	5.46	0.12	−0.35	0.09	25.33	10.40	31.63	1.54
COMP3638	−0.09	0.18	0.15	0.16	1.68	1.81	1.46	0.20
COMP3738	6.23	0.29	−0.03	0.24	29.72	12.95	36.61	1.79
SIGMA(1)			5.86	1.51				
RHO(1)			0.84	0.10				
SIGMA(0)					2.13	0.002		
RHO(0)					−0.87	0.03		

*Note that (6) and (7) were substituted into (5).
**Note that the equation is corrected by inverse Mill's ratio.
NOBS = 31,817

In other words, if $P_{ijt} = 1$, R&D expenditures are given by (6), while R&D expenditures are determined through (7) whenever $P_{ijt} = 0$. Note that OLS estimates of (6) and (7) yield inconsistent estimates since $\mathrm{E}(v_{ijt}/P^*_{ijt} > 0) \neq 0$ and $\mathrm{E}(\varepsilon_{ijt}/P^*_{ijt} \leq 0) \neq 0$. Following Lee, we apply a two-stage probit estimation where we substitute (6) and (7) into (5), yielding a reduced-form probit model. The reduced-form probit model can be consistently estimated by standard probit methods. We estimate the selection equation using a probit, which is a common procedure for switching models as it makes it easier to calculate the hazard rates in the second stage. Using the predicted probabilities $\widehat{P}_{ijt}$ from the reduced-form probit and correcting (6) and (7) by using inverse Mill's ratio, we get consistent estimates of the R&D equations by least squares: where $\mathrm{E}(u_{ij}\varepsilon_{ij}) = \sigma_1^2$, $\mathrm{Corr}(u_{ij}\,\varepsilon_{ij}) = \rho_1$, $\mathrm{E}(u_{ij}v_{ij}) = \sigma_0^2$, and $\mathrm{Corr}(u_{ij}v_{ij}) = \rho_0$, and u_{ij} is the first-stage probit error, which follows a normal distribution with $\mathrm{N}(0, \sigma^2)$. If $\rho_1 = \rho_0 = 0$, the model represents an exogenous switching model and the R&D equations can be estimated by

OLS. Otherwise, the model incorporates an endogenous switching process. To obtain asymptotically efficient estimates, we have computed the FIML estimates of the above model.

With regard to the identification of our model we show in Section 2.2, based on our theoretical model, that neither *DASSET* nor *MEMBERS* can be excluded from our R&D growth switching equations. We therefore include the variables *DASSET* and *MEMBERS* in both R&D growth switching equations.[11] We identify our system by using the nonlinearity from Mill's ratio.[12] The results are shown in the article, Table 6. We also estimated another specification in which the variables *DASSET* and *MEMBERS* enter only the R&D switching equation for firms forming an RJV. We also apply further robustness checks with regard to excluding the variables *DASSET*and *MEMBERS* from the R&D growth switching equations and test whether the results are sensitive with regard to the specification of our R&D growth switching equations and to test the identification restrictions derived from our theoretical model.[13] In further robustness checks we also tested many other different specifications in which the variables *DASSET* and *MEMBERS* entered the switching equations in different combinations with the remaining variables. The signs and magnitudes of the estimated parameters are very similar to the results reported in our article.

3.1. *Results and Interpretation*

Before we turn to the probit equation, we briefly discuss the R&D equations. The results of the R&D equations (6) and (7) corrected by Mill's ratio are presented in Table 6. Before interpreting our results, it is important to check whether the truncation terms ρ and σ are significant. We find a significant estimate for the correction. This indicates that the selectivity through the endogenous dummy variable is indeed an important issue and illustrates the necessity of estimating an endogenous switching model.

We now can report a number of empirical findings regarding the R&D spending of RJVs. As shown in Table 6, the Herfindahl index (*HERF*) is highly significant and positive in the R&D growth equation for firms forming an RJV together, indicating that firms in less concentrated industries increased their R&D intensities (recall the definition of $\Delta r\&d$, which is the change in firm-level R&D intensities). This implies that the free-rider and complementarity effects become more important relative to cost-sharing incentives as the concentration of the industry increases. The positive and significant coefficient of *FRJVS* indicates that the participation in more RJVs tends to decrease firm-level R&D intensities. Therefore, we find that cost-sharing becomes more important as firms participate in more RJVs. The estimation of the corrected R&D

---

[11] We also would like to thank one of the referees for the clarification on the identification issue.

[12] Cameron and Trivedi (2005, page 511) explicitly mention that a sample selection model is theoretically identified without any restrictions on the regressors. Wooldridge (2002, page 569–70) even mentions that imposing exclusion restrictions on a reduced form equation can be detrimental and unnecessary.

[13] The identification through the nonlinearity of inverse Mill's ratio may cause multicollinearity problems, as the switching equations are functions of the explanatory variables in the selection equation, see also Wooldridge (2002, pp. 565–9) and Maddala (1983, p.252). A few of the dummy variables (e.g., *COMP3238*) become statistically insignificant in some specifications, as they have only few observations in which those dummy variables take on the value of one. We have estimated the models with and without these dummies. Our results are not significantly affected by dropping the dummies. We report the results when some dummies have been dropped.

equation (6) shows a considerable amount of heterogeneity among the intra-industry dummies. Industries in which firm-level R&D intensity growth in an RJV is relatively large include the 'Chemicals and Allied Products' industry *(SIC28)*, and the 'Industrial Machinery and Equipment' industry *(SIC35)*. Apparently, those industries are characterised by free-rider problems. Turning to complementary industry effects, we find that firm-pairs from 'Chemicals and Allied Products' and 'Primary Metal Industries' (*COMP2833*), 'Chemicals and Allied Products' and 'Industrial Machinery and Equipment' industries (*COMP2835*), and 'Industrial Machinery and Equipment' and 'Transportation Equipment' industries (*COMP3537*) seem to increase their R&D intensities. Estimates of the corrected R&D (7) for firm-pairs that do not participate in the same RJV are presented in the third column of Table 6.

We now turn to our main objective. Table 6 also presents in the last column the structural probit estimates based on (5).[14] The variable *DASSET* has a negative impact on the probability of forming an RJV with a point estimate of −6.57, implying that RJVs tend to be formed among firms of similar size. In addition, the impact is statistically significant with a standard error of 0.31. We therefore find evidence for the firm-size hypothesis in our data. Note that what seemed to reject our hypothesis by simply comparing sample means turns out to be the reverse once we estimate an endogenous switching regression model and introduce our controls.

We also find that firms from industries with higher concentration ratios (measured by *HERF*) have a negative and statistically significant effect. This is consistent with our theory that market power considerations affect whether large firms joint venture with small firms. The negative and significant impact of *MEMBERS* indicates that the presence of larger RJVs makes it less likely that two firms will join in an RJV together, a result consistent with the theory of Bloch (1995). The positive and significant effect of *FRJVS* suggests that the more RJVs in which a particular firm is involved, the more likely it will be to join additional RJVs, i.e., the returns to RJVs are increasing. This may be suggestive of some organisational infrastructure or organisational learning required for RJVs (i.e., some upfront costs). Once some of the initial hurdles are past then participating in additional RJVs may become easier.

Turning to the industry dummies, it is interesting to compare the intra-industry dummies (*SICs*) with the inter-industry dummies (*COMPs*). As shown in Table 6, the point estimate for the 'Transportation Equipment' industry (*SIC37*) is 73.18, which is the largest significant estimate for an industry, implying that firms in *SIC37* have the highest intra-industry probability to form an RJV. As expected, the complementarity dummies vary substantially according to the industry pairs considered. However, in many cases, the *COMP* dummies are larger than the *SIC* dummies, indicating that complementarities are an important factor in the formation of RJVs.

We do find some large, statistically significant complementarities between some industry groups. The 'Transportation Equipment' industry in particular exhibits complementarities with other industries (see e.g., *COMP3237*, *COMP3337*, *COMP3537*, and *COMP3738*). The 'Primary Metal Industries' and 'Transportation Equipment' (*COMP3337*) as well as the 'Transportation Equipment' and 'Instruments and Related Products' (*COMP3738*) display the highest likelihood of forming RJVs with each other.

---

[14] Counterfactuals have bee formed by using the results from (6) and (7) and substituting those into (5).

These industries appear to have vertical relationships. Given those vertical relationships, one would expect that firms in these industries produce complementary products and that the incentive to form an RJV is high. Moreover, the time dummies are positive and mostly increasing over time, which indicates that firms are increasingly getting involved in more RJVs over time. This finding is consistent with the descriptive statistics provided in Table 1, which indicates that more RJVs have been registered over time. For robustness, we also estimate a counterfactual or structural logit equation.

Using the predicted values for $\Delta r\&d_{ij}$ (when $P_{ijt} = 1$) and $\Delta r\&d_{ij}$ (when $P_{ijt} = 0$) based on the corrected equations (6) and (7), we take the counterfactual of each group by calculating the predicted values of what the firm-pairs would have spent on R&D had they been a member of the other group and then take the difference (i.e., $DRDJV_{ijt} = \widehat{\Delta r\&d}_{ijt}(\text{when } P_{ijt} = 1) - \widehat{\Delta r\&d}_{ijt}(\text{when } P_{ijt} = 0)$ for all observations. This difference can then be used for estimating the resulting logit-estimates, which are consistent as shown by Lee (1979). The *DRDJV* variable has a positive and statistically significant effect. This is consistent with the effect of *MEMBERS* being positive in our estimation of (6). Hence, the cost-sharing effect (net of the free-rider and complementarity effects) is an important determinant of RJV formation. The other estimates of *DASSET* and *HERF* have the same sign as in the FIML estimation from above and confirm the previous results.

In a final step, we also test for the Rare Events problem. In this instance, the rare event results from the fact that we have many more firm-pairs that do not form an RJV than firm-pairs that do form an RJV. With rare event data, logistic regression can underestimate the probability of events. King and Zeng (2001) discuss methods of computing probability estimates that correct problems due to rare events and propose an estimator that gives a lower mean square error. The Rare Events Logit (Relogit) estimates the same Logit model but with an estimator that gives a lower mean square error in the presence of rare events data for the coefficient. The basic idea is to weight the log-likelihood function by the inverse of the *ex ante* probability of the observation in the sample. We conducted a rare events logit for robustness, and can confirm no statistically different results from the regular logit. The results of the standard Logit and the Relogit estimations (using the procedure in Tomz *et al.*, 2003) provide very similar estimates of the parameters and the standard errors.[15]

4. Conclusion

The goal of this article is to test RJV formation hypotheses on US data that became available through the NCRA. We estimate a two-equation system that endogenises RJV formation and R&D investments through an endogenous switching model. We report on a number of empirical findings regarding R&D spending and RJV formation. In particular, we find that R&D cost-sharing is an important incentive for RJV formation.

---

[15] The reason that King and Zeng used the logistic distribution is given by the fact that in the logit case, prior correction method (correcting the estimates based on prior information about the fraction of ones in the population) is consistent, fully efficient and easy to apply (King and Zeng, 2001). We apply another robustness check and run the same regression using the probit estimation procedure. The results are very similar to the logit estimates.

Our main findings are that firm-size differences, number of members in the RJV, the industries of the firms, and the impact on R&D investments are significant factors in determining whether two firms join together in an RJV. The econometric estimates imply that firms are more likely to form an RJV the more similar they are in their size. We also find that a higher participation rate in other RJVs makes it more likely to engage in further RJVs, i.e., the returns to RJVs are increasing. Firms in the 'Transportation Equipment' manufacturing industry or firms in industries related to that industry are more likely to participate in RJVs. Finally, we show that the tendency to form RJVs increased over our study period from 1990 to 1994.

European School of Management and Technology
Purdue University
University of Toronto at Mississauga

Submitted: 15 July 2004
Accepted: 13 October 2006

References

Beath, J., Poyago-Theotoky, J. and Ulph, D. (1998). 'Organization design and information-sharing in a research joint venture with spillovers', *Bulletin of Economic Research*, 50 (1), pp. 47–59.

Beecy, L., Link, A., William, F. and Teece, D. (1994). 'Quantifying private rates of return to collaborative research: the case of SEMATECH', mimeo.

Berg, S.V., Duncan, J. and Friedman, P. (1982). *Joint Venture Strategies and Corporate Innovation*, Cambridge: Oelgeschlager Gunn & Hain.

Bergstrom, T. and Varian, H. (1985). 'Two remarks on Cournot equilibria', *Economics Letters*, vol. 19, pp. 5–8.

Bloch, F. (1995). 'Endogenous structures of association in oligopolies', *Rand Journal of Economics*, vol. 26(3), pp. 537–56.

Boone, J. (2000). 'Competitive pressure: the effects on investments in product and process innovation', *Rand Journal of Economics*, vol. 31(3), pp. 549–69.

Cameron, A.C. and Trivedi, P.K. (2005). *Microeconometrics*, Cambridge: Cambridge University Press.

Cassiman, B. and Greenlee, P. (1999). 'Firm objectives and the formation of research joint ventures', *Managerial Decision and Economics*, vol. 20(3), pp. 115–30.

Cassiman, B. and Veugelers, R. (2002). 'Spillovers and R&D cooperation: some evidence from Belgium', *American Economic Review*, vol. 92 (4) pp. 1169–84.

d'Aspremont, C. and Jacquemin, A. (1988). 'Cooperative and noncooperative R&D in duopoly with spillovers', *American Economic Review*, vol. 78 (December), pp. 1133–7.

Irwin, D. and Klenow, P. (1996). 'High tech R&D subsidies: estimating the effects of SEMATECH', *Journal of International Economics*, vol. 40 (3–4), pp. 323–44.

Jorde, T. and Teece, D. (1990). 'Innovation and cooperation: implications for competition and antitrust', *Journal of Economic Perspectives*, vol. 4(3) (Summer), pp. 75–96.

Kaiser, U. (2002). 'An empirical test of models explaining research expenditures and research cooperation: evidence for the German service sector', *International Journal of Industrial Organization*, vol. 20, pp. 747–74.

Kamien, M., Muller, E. and Zang, I. (1992). 'Research joint ventures and R&D cartels', *American Economic Review*, vol. 82(5), pp. 1293–306.

Katsoulacos Y. and Ulph, D. (1998). 'Endogenous spillovers and the performance of research joint ventures', *Journal of Industrial Economics*, vol. 46, pp. 333–57.

Katz, M. (1986). 'An analysis of cooperative research and development', *Rand Journal of Economics*, vol. 17(4), pp. 527–43.

King, G., and Zeng, L. (2001). 'Logistic regression in rare events data', *Political Analysis*, vol. 9(2), pp. 137–63.

Lee, L.F. (1978). 'Unionism and wage rates: a simultaneous equations model with qualitative and limited dependent variables', *International Economic Review*, vol. 19(2), pp. 415–33.

Lee, L.F. (1979). 'Identification and estimation in binary choice models with limited (censored) dependent variables', *Econometrica*, vol. 47, pp. 977–96.

Link, A. (1996). 'Research joint ventures: patterns from federal register filings', *Review of Industrial Organization*, vol. 11, pp. 617–28.

Maddala, G.S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*, New York: Cambridge University Press.

Majewski, S. and Williamson, D. (2002). 'Endogenous spillovers, strategic blocking, and the design of contracts in collaborative R&D: evidence from the NCRA filings of R&D joint ventures', Antitrust Division, US Department of Justice, Economic Analysis Group working paper.

Marin, P.L., Siotis, G. and Hernan, R. (2003). 'An empirical evaluation of the determinants of research joint venture formation', *Journal of Industrial Economics*, vol. 51 (1), pp. 75–89.

Moody's Investors Service. (1995). *Moody's Global Company Data*, New York: Moody's Investors Service.

Nalebuff, B. and Stiglitz, J. (1983). 'Information, competition and markets', *American Economic Review Papers and Proceedings*, vol. 72 (May), pp. 278–84.

Rosen, R. (1991). 'Research and development with asymmetric firm sizes', *Rand Journal of Economics*, vol. 22(3), pp. 411–29.

Salant, S. W. and Shaffer, G. (1998). 'Optimal asymmetric strategies in research joint ventures', *International Journal of Industrial Organization*, vol. 16 (2), pp. 195–208.

Salant, S. W. and Shaffer, G. (1999). 'Unequal treatment of identical agents in Cournot equilibrium', *American Economic Review*, vol. 89 (3), pp. 585–604.

Tomz, M., King, G. and Zeng, L. (2003). 'RELOGIT: rare events logistic regression', *Journal of Statistical Software*, vol. 8 (2).

US Department of Commerce. (1993). 'Research and development consortia registered under the national cooperative research act of 1984'. The Technology Administration, Office of Technology Commercialization, Washington, DC.

Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*, Cambridge MA: MIT Press.

Yi, S.-S. and Shin, H. (2000). 'Endogenous formation of research coalitions with spillovers', *International Journal of Industrial Organization*, vol. 18, pp. 229–56.