Determinants for R&D Cooperation: Evidence from Spanish Manufacturing Firms

Alberto López*

Universidad Carlos III de Madrid

March 2006

Abstract

Using firm level data, this paper explores the determinants for R&D cooperation. It focuses on the impact of information flows or spillovers on R&D cooperation, but also explores the role of the traditionally considered factors (firm size, cost and risk sharing, complementarities). The estimation methods used allow for testing the endogeneity of the explanatory variables which in other papers are assumed to be endogenous a-priori. I find that the choice of an appropriate "structure" of endogeneity has important consequences for the estimates: cost risk sharing and complementarities in this case only have the expected positive effect. I also find that the overall picture of the importance of the explanatory variables depends on the estimation method. In this sense, two-step procedures overestimate the importance of spillovers. With a more efficient procedure, I find that cost risk sharing is the most important determinant for R&D cooperation in Spain. Finally, the overall results on the importance of spillovers are consistent with the existing literature, but I find that the level of legal protection in the industry has a negative effect on R&D cooperation.

Key words: R&D cooperation, Spillovers, Maximum likelihood methods

JEL Classification: O31, O32, L22, L60

*Dpto. de Economía, C/ Madrid 126, 28903 Getafe (Madrid), Spain; e-mail: alsebast@eco.uc3m.es; tel.: (34) 91 624 92 87; fax: (34) 91 624 98 75.

1. Introduction

In the ultimate interest of stimulating innovation, much attention has recently focused on the subject of cooperative firm R&D. These agreements, from the positive point of view, are likely to embody mechanisms by which firms can profitably appropriate free flows of knowledge and protect them. Hence, they are an interesting guide to normative regulation, which must try to consolidate mechanisms of incentives and at the same time avoid harming market competition.

R&D cooperation is thus becoming a major topic for policy makers. Most E.U. and national public funding for R&D is directed at stimulating cooperation between firms, and between firms and public institutions[1]. The rationale behind this policy is to generate or improve information flows or spillovers between these economic agents, as these spillovers are assumed to essentially lead to more economic growth[2] and a better performance of the national system of innovation.

Given this growing interest, literature has recently paid attention to the relationship between R&D cooperation activity and spillovers. Cassiman and Veugelers (2002), from now on CV, find that the firms' external information sources (incoming spillovers) and the flows out of the firms measured through the ability of firms to appropriate the returns from innovation (appropriability) have important and separately identifiable effects on the probability of R&D cooperation. Other works have studied the relationship between spillovers and R&D cooperation; see, for example, Belderbos et al. (2004) for evidence on this relationship from the Netherlands, and Kaiser (2002a) for evidence from the German service sector.

Besides knowledge flows, literature has identified three major classes of motives for firms getting involved in R&D cooperation: cost and risk sharing-related reasons[3], complementarities-

[1]See Acosta and Modrego (2001) for an example of public funding in Spain, and Abramovsky, Harrison and Simpson (2004) for a summary for the UK.

[2]See Griliches (1992) for a survey on the empirical evidence on the relationship between R&D spillovers and growth, and Romer (1990) for a theoretical discussion.

[3]See, among others, Belderbos et al. (2004), Hagedoorn (1993), Miotti and Sachwald (2003), Tether

related reasons or skill-sharing reasons[4], and factors related to the absorptive capacity of the firm[5]. Firstly, cooperative R&D agreements may be used by firms to set cost and risk-sharing rules in high-cost and risky settings. Hence, when cost and risk are important innovation hampering factors, firms would tend to make cooperative R&D agreements. Secondly, cooperative R&D is a vehicle for firms to learn skills and capabilities from their partners. As such, the greater the availability of technological know-how within the firm, the more likely it is to have complementarities between partners in a cooperative R&D agreement. Finally, one other determinant that is closely related to knowledge flows and complementarities is the idea of absorptive capacity. A firm's absorptive capacity is derived from its own R&D efforts and it is a measure of its ability to benefit from other firms' R&D activity. The higher the absorptive capacity of the firm, the higher the benefits from R&D cooperation.

This paper develops evidence about the determinants for R&D cooperation using a sample of Spanish manufacturing firms, focusing mainly on the importance of spillovers. The paper is based on the model introduced by CV, although it departs from these authors to explore some econometric and substantial issues.

The contribution of this paper to the empirical literature on R&D cooperation is threefold. First, I show that an adequate treatment of endogeneity matters a great deal. I find evidence supporting the existence of an important effect of spillovers on R&D cooperation, although cost and risk sharing is the most important determinant for cooperation in Spain. In obtaining these results, I apply a complete treatment for endogeneity. Two alternative estimation methods are used: Two-stage conditional maximum likelihood (2SCML) and Conditional maximum likelihood (CML). These techniques allow me both to test for the endogeneity of the explanatory variables which in other papers are assumed to be endogenous a-priori and to obtain efficient estimates. I find that the choice of an appropriate "structure" of endogeneity is crucial and has important consequences for the estimates. I

(2002), Tyler and Steensma (1995).

[4]For example, Hagedoorn (1993), Sakakibara (1997), Tyler and Steensma (1995).

[5]See, among others, Cohen and Levinthal (1989), Röller et al (2002), Sakakibara (1997), Tether (2002).

also find that, depending on the estimation method, a different picture of the importance of the explanatory variables is obtained. In this sense, two-step procedures overestimate the importance of spillovers and underestimate the impact of cost and risk sharing reasons on the probability of R&D cooperation.

Second, I obtain new insights on the topic due to the sample employed. On the one hand, I use a large sample of 2518 firms, in contrast with the 411 observations used by CV. This sample size allows me to obtain more accurate estimations and more precision applying hypothesis tests. On the other hand, compared with most European countries, the Spanish system of innovation is in an earlier stage of development[6]. As R&D cooperation is an important vehicle for improving the innovation performance of firms, this gap makes the study of those factors that may stimulate cooperation in R&D more interesting. Moreover, the structure of the Spanish manufacturing sector is characterized by a large share of small and low-technology firms, while the general finding in the literature is that firms from high-technology sectors and big firms are more likely to cooperate in R&D. It is worthy of exploring cooperation in such a context of small and low-technology firms.

Third, I extend CV's framework to the analysis of the determinants for R&D cooperation with competitors and I pay more attention to the relationship between cooperation and the effectiveness of different legal methods for protecting inventions or innovation.

The rest of the paper is organized as follows. Section 2 introduces the data and presents some descriptive analysis of the sample. Section 3 introduces the framework for the analysis. The econometric specification is shown in Section 4. Section 5 presents the results. Finally, Section 6 concludes.

[6] Compared with France, Germany and the United Kingdom, Spain presents the lowest proportion of firms with innovation expenditures and with intramural R&D. The R&D intensity (ratio of intramural R&D expenditure over total turnover) of Spanish firms performing R&D is, approximately, one third of the efforts of France, Germany and the United Kingdom. Spain also presents the lowest share of firms with R&D cooperation agreements. See Abramovsky et al. (2004) for a detailed comparison in the innovation activities and performances at the national level for France, Germany, Spain and the United Kingdom.

2. Data and descriptive analysis

The data used correspond to the Third Community Innovation Survey (CIS3; period 1998-2000), carried out in Spain by the Instituto Nacional de Estadística (INE) under the name Encuesta de Innovación Tecnológica en las Empresas. The Community Innovation Surveys take place every 4 years in European countries to investigate a firm's innovation activities. In 2001, the third wave was conducted and covered the period 1998 to 2000. The CIS3 follows the recommendations of the OSLO Manual on performing innovation surveys (see OECD and Eurostat, 1997).

The Spanish CIS3 collected data on 11778 firms[7]. The population target was firms with 10 or more employees. The participation is compulsory for firms and is based on stratified samples by size and sector. Unit non-response analysis is not carried out.

The final sample of the manufacturing sector includes 6026 firms[8], 41.8% (2518 firms) of which report having introduced innovations during the reference period. This work restricts the attention to this subsample of innovating manufacturing firms.[9]

Table 1 reports some sample statistics. It turns out that 476 firms out of our sample of 2518 innovating firms (18.9%) have at least one cooperative R&D agreement. It is helpful to further distinguish among different types of cooperative R&D agreements depending on the kind of partner: 184 firms cooperate with competitors, 316 firms cooperate with suppliers or customers (vertically-related firms), and 425 firms cooperate with research institutions.

Table 2 shows that most firms maintain cooperative R&D agreements with different partners. Sixty-one percent of firms have agreements with at least two types of partners, and 33.4% cooperate with all three types. It is important to keep this in mind when analyzing cooperation by type of partner. For example, just 144 firms which cooperate

[7] 6094 in Manufacturing (NACE 15-37), 4778 in Services (NACE from 50 to 95), and the rest in Mining and quarrying (NACE 10-14), Electricity, gas and water supply (NACE 40-44) and Building (NACE 45).

[8] In this exercise, I drop a total of 68 manufacturing firms because of partially incomplete data.

[9] Innovating firms are defined as those which report having introduced product or process innovations, having ongoing innovation activities, or having abandoned innovation activities, and, at the same time, present a positive amount spent on innovation during the period 1998-2000.

with research institutions have agreements exclusively with these institutions, while the other 281 firms also maintain agreements with at least one other type of partner.

Table 3 shows the distribution of the sample of innovating manufacturing firms across industries and size. The sample presents a larger number in small firms (fewer than 200 employees) than in big firms (200 or more workers); 1748 and 770 firms, respectively. With respect to sector distribution, the sample shows a higher share of firms in low-technology sectors (63.9% of the firms belong to low-technology sectors). These facts are consistent with the Spanish manufacturing sector characteristics shown in the introduction. Focusing on R&D cooperation activity, innovative firms in high-technology manufacturing sectors and big firms are more likely to engage in cooperative activity.

3. A framework for our analysis

Based on the literature reviewed in the introduction, this paper models the probability of cooperation as depending on spillovers, as well as the traditional variables which are thought to affect R&D cooperation (cost-risk sharing, complementarities, absorptive capacity of the firm, etc.). I include three variables related to the measure of spillovers, i.e., incoming spillovers (measured by the importance of publicly available information for the firm's innovation process), appropriability (measured by the effectiveness of the different strategic protection methods of innovations, the converse of which can be thought of as the extent of outgoing spillovers) and a measure of the importance of legal methods for protecting inventions or innovation at the industry level[10]. Detailed definitions of all employed variables can be found in Appendix A.

Let me briefly comment on the expected effects of the explanatory variables.

Incoming spillovers are expected to have a positive effect on the probability of cooperation. The higher incoming spillovers are, the greater the scope for learning within cooperative R&D agreements, and hence the marginal profit to be derived from cooperation.

[10] As far as legal protection can be considered an industry variable rather than a firm-specific characteristic, only the average industry score for legal protection is employed. The industry is defined at the NACE 2-digit sector level.

The sign of the effectiveness of strategic and legal protection methods (appropriability and industry level of legal protection), however, is not so clear a priori. Literature suggests two opposite effects of this variable on the probability of cooperation. The net effect will then depend on their relative importance. On the one hand, a low level of effectiveness increases the scope for the internalization of information flows between firms through cooperation in R&D. But, on the other hand, incentives to become a free rider of other firms' investments will reduce profitability and the stability of cooperative agreements.

Cost-risk variable, given the hypothesis of cost and risk sharing, is expected to show a positive effect on cooperation. To test for complementarities, a variable which measures the availability of technological know-how within the firm is included. This variable is expected to have a positive effect on cooperation.

Benefits from R&D cooperation depend on the absorptive capacity of the firm. In this sense, the higher the firm's absorptive capacity, the higher the returns that the firm can expect from access to external resources. On the one hand, theoretical models explicitly incorporate the need for a firm to conduct its own R&D in order to realize spillovers from other firms' R&D activity (Griffith, Redding and Van Reenen, 2003; Kaiser, 2002b; and Kamien and Zang, 2000). On the other hand, empirical studies, such as Cohen and Levinthal (1989, 1990); Griffith, Redding and Van Reenen (2003, 2004), have shown that firms' absorptive capacity depends on their own R&D intensity (R&D expenditures/turnover)[11]. So, R&D intensity is included as a measure of the absorptive capacity of the firm.

Additionally, firm's size is also included as a measure of the absorptive capacity of the firm. Therefore, we should expect a positive effect of the firm's size on the probability of cooperation. Size squared is considered to allow for a nonlinear effect of firm size.

Specification also includes the level of cooperation by industries, which is assumed to pick

[11] In the empirical literature, other variables have been used in order to measure the absorptive capacity of the firm. For example, Belderbos et al. (2004); and Fritsch and Lukas (2001) measure the R&D intensity by the ratio of R&D personnel to total personnel. Jaffe (1986) uses both the ratio of R&D expenditures on capital and the level of R&D expenditures. While Miotti and Sachwald (2003) and CV use a dummy variable for permanent R&D as an indicator of the absorptive capacity.

up unobserved industry-specific attributes that contribute to the decision of engaging in a cooperative R&D agreement. Table 4 summarizes the theoretical predictions along with the empirical findings at the end of Section 5.

4. Econometric specification

The problem of endogeneity

My concern is that some of the explanatory variables introduced in the former section are, in fact, endogenous. A priori, as in CV, I will consider the possible endogeneity of incoming spillovers, appropriability and R&D intensity. Additionally, in a departure from CV's paper, the cost-risk variable will also be taken as a possible endogenous variable. Endogeneity can arise in two different ways: Omitted variables that I cannot include in the model and simultaneity in the decisions.

Firstly, the propensity to cooperate in R&D can be correlated with unobserved factors that are also systematically correlated with some of the explanatory variables. First, concerning demand side factors, we can include managing capacity and quality, the choice of governance mode of R&D activities, the extent to which the firm is open to new ideas, the permeability of the firm, reputation, outward-looking style of management and tacitness of the firm's knowledge assets[12]. Second, as for supply-side factors, we can consider the geographical proximity and the accessibility to an intensive technological area[13]. Third, of supply-demand interaction factors, we can include repeated interactions with the same partner, the length of the cooperation relationship and previous R&D cooperation agreements.

In addition to the omitted variables problem, spillovers, R&D intensity and cost-risk are

[12] For example, it is reasonable to think that the higher the manager's openness to new ideas, the higher the propensity to R&D cooperation. Additionally, the culture of openness to new ideas seems to affect, among others, the use of public sources of information (incoming spillovers) and the manager's risk aversion.

[13] These factors can affect R&D cooperation simultaneity and variables such as R&D intensity, incoming spillovers, the effectiveness of strategic protection methods (appropriability) and the accessibility to appropriate sources of finance.

also expected to be endogenous variables due to a simultaneity problem. Firstly, cooperative R&D agreements can be used to manage external knowledge flows[14], which implies that the decision to cooperate can influence incoming spillovers as well as the effectiveness of appropriation strategies. Secondly, several studies[15] have found evidence supporting the endogeneity of R&D intensity when analyzing the R&D cooperation decision because of simultaneity in the decisions. In this sense, R&D intensity may increase if R&D cooperation makes own R&D expenditures more effective. Finally, when firms use cooperative R&D agreements to share cost and risk, the effects of cooperation can influence the importance given to these variables as obstacles to innovation.

System of simultaneous equations

Due to the endogeneity of a number of variables, I consider a system of simultaneous equations (see Appendix B for details). The model is composed of a structural equation that is of primary interest (the cooperation equation) and a set of reduced form equations for the potential endogenous explanatory variables (incoming spillovers, appropriability, R&D intensity and cost-risk variable). The unobservable propensity to cooperate in R&D (y_1^*) is assumed to be a linear function of a set of observed exogenous explanatory variables ($\mathbf{z}_1$); such as the firm's size and the industry level of legal protection methods, a set of (possible) endogenous explanatory variables ($\mathbf{y}_2$) and an error term (u_1). Let y_1 equal 1 if the firm cooperates.

$$y_1^* = \mathbf{z}_1\boldsymbol{\delta}_1 + \mathbf{y}_2\boldsymbol{\alpha}_1 + u_1 \tag{1}$$

$$y_1 = 1\left[y_1^* > 0\right] \tag{2}$$

I assume that the endogenous explanatory variables are a function of the exogenous variables that determine cooperation ($\mathbf{z}_1$), a set of other exogenous variables ($\mathbf{z}_2$) and an error term ($\mathbf{v}_2$).

$$\mathbf{y}_2 = \mathbf{z}_1\boldsymbol{\Delta}_{21} + \mathbf{z}_2\boldsymbol{\Delta}_{22} + \mathbf{v}_2 = \mathbf{z}\boldsymbol{\Delta}_2 + \mathbf{v}_2 \tag{3}$$

[14] See, for example, Kamien, Müller and Zang (1992).

[15] See, among others, Becker and Dietz (2004), Colombo and Garrone (1996), and Veugelers (1997).

The arguments I presented before suggest that u_1 and $\mathbf{v}_2$ are correlated. The model described by equations (1) – (3) is applicable when $\mathbf{y}_2$ is correlated with u_1 due to omitted variables and when $\mathbf{y}_2$ is correlated with u_1 because y_2 is determined jointly with y_1 if y_1^* appears in a linear structural equation for $\mathbf{y}_2$[16].

I assume that $u_1$and $\mathbf{v}_2$ have a joint normal distribution with mean zero and finite positive covariance matrix:

$$\boldsymbol{\Omega} \equiv \begin{bmatrix} \sigma_{u_1}^2 & \boldsymbol{\Sigma}_{u_1\mathbf{v}_2} \\ \boldsymbol{\Sigma}_{\mathbf{v}_2 u_1} & \boldsymbol{\Sigma}_{\mathbf{v}_2\mathbf{v}_2} \end{bmatrix} \tag{4}$$

Under joint normality of $(u_1, \mathbf{v}_2)$, I can write.

$$u_1 = \mathbf{v}_2\boldsymbol{\theta}_1 + e_1 \tag{5}$$

where $\boldsymbol{\theta}_1 = \boldsymbol{\Sigma}_{\mathbf{v}_2\mathbf{v}_2}^{-1}\boldsymbol{\Sigma}_{\mathbf{v}_2 u_1}$

Estimation methods

Once the endogeneity of some variables is recognized, it is clear that the estimation of the model by OLS or other techniques that do not allow for the endogeneity is inappropriate and has important consequences, i.e., in applying OLS, we will not be able to consistently estimate any of the coefficients of equation (1). For instance, in the empirical literature, the importance of cost and risk as obstacles to innovation has typically been considered an exogenous determinant for R&D cooperation. However, considering this variable as exogenous, it is hard to reach any broad generalization on the relation between R&D cooperation and cost-risk sharing[17]. Therefore, a proper treatment of endogeneity is necessary to obtain

[16] In this case, $\mathbf{y}_2$ has the reduced form given by equation (3) (for a further discussion on this topic, see Maddala, 1983, Chapter 7; and Wooldridge, 2002, Chapter 15). In our case, notice that the variables used are contemporaneous, so it is plausible to think that the propensity or intention to cooperate (y_1^*), and not the actual action (y_1), should be used as an explanatory variable for $\mathbf{y}_2$.

[17] Miotti and Sachwald (2003) find that sharing costs and risks is not a significant determinant in the probability of R&D cooperation, while Tether (2002) find a positive and significant effect. Moreover, CV find a positive and significant effect of the importance of cost as a hampering factor for the innovation process of the firm, and, at the same time, the importance of risks has a negative and significant effect on R&D cooperation.

consistent estimates.

Moreover, the fact of considering an explanatory variable to be exogenous or endogenous can yield very different pictures of its importance[18].

Due to its importance, my choice is to apply a complete treatment for the endogeneity problem. Instead of, as in CV, assuming the endogeneity of some explanatory variables and using less efficient two-step procedures to obtain the final estimates, the methods applied in this paper allow me, with a slightly computational cost, to both test for the endogeneity of some explanatory variables of interest and obtain more efficient estimates. We use a maximum likelihood estimation in order to present the final findings, while a two-step approach is used for the initial exogeneity test of some explanatory variable.

Firstly, estimating this model, I use a two-stage conditional maximum likelihood method (2SCML). This approach is due to Rivers and Vuong (1988).[19] A convenient feature of this procedure is that it provides an estimate of $\boldsymbol{\theta}_1$that can be used to test for the endogeneity of $\mathbf{y}_2$. This method is a two-step estimation procedure. In the first step, the (assumed) endogenous variables ($\mathbf{y}_2$) are regressed on all the (assumed) exogenous variables ($\mathbf{z}$). In the second step, the residuals of the first-step regressions ($\widehat{\mathbf{v}}_2$) are used as independent variables in the cooperation equation (joint with $\mathbf{z}_1$and $\mathbf{y}_2$). The usual probit t statistic on $\widehat{\mathbf{v}}_2$ is a valid test of the null hypothesis that $\mathbf{y}_2$ is exogenous. The Rivers-Vuong approach is used for the initial test of whether $\mathbf{y}_2$ is exogenous[20].

Once the exogeneity of some explanatory variables is tested, the system of equations is estimated by conditional maximum likelihood (CML)[21]. The log likelihood function for an

[18]For example, CV considers the importance of publicly available information sources as endogenous and, using a two-step procedure, find a positive and significant effect of this variable on R&D cooperation. On the other hand, considering public incoming spillovers as exogenous, Belderbos et al. (2004) find no evidence on the effect of this variable on the probability of R&D cooperation.

[19]See Wooldridge (2002) for a recent review of this method.

[20]Note that if $\theta_1 \neq 0$, we have only estimated the coefficients up to scale.

[21] *The* CML estimator is a full-information maximum likelihood estimator. It is based on the entire system of equations, treats all equations and parameters jointly and gives direct estimates of the coefficients. System methods of estimation (CML) are preferred to and asymptotically better than limited information methods, or single-equation methods (2SCML), since the latter neglect information contained in other equations

individual in this model is (see Appendix B for details).

$$\log L = y_{i1} \log \mathbf{\Phi} \left(\frac{\mathbf{z}_{i1}\boldsymbol{\delta}_1 + \mathbf{y}_{i2}\boldsymbol{\alpha}_1 + (\mathbf{y}_{i2} - \mathbf{z}_i\mathbf{\Delta}_2)\,\boldsymbol{\theta}_1}{\left(1 - \boldsymbol{\theta}_1^{'}\mathbf{\Sigma}_{\mathbf{v}_2\mathbf{v}_2}\boldsymbol{\theta}_1\right)^{\frac{1}{2}}} \right) \tag{6}$$
$$+(1 - y_{i1}) \log \left[1 - \mathbf{\Phi} \left(\frac{\mathbf{z}_{i1}\boldsymbol{\delta}_1 + \mathbf{y}_{i2}\boldsymbol{\alpha}_1 + (\mathbf{y}_{i2} - \mathbf{z}_i\mathbf{\Delta}_2)\,\boldsymbol{\theta}_1}{\left(1 - \boldsymbol{\theta}_1^{'}\mathbf{\Sigma}_{\mathbf{v}_2\mathbf{v}_2}\boldsymbol{\theta}_1\right)^{\frac{1}{2}}} \right) \right]$$
$$-\frac{m}{2}\log(2\pi) - \frac{1}{2}\log|\mathbf{\Sigma}_{\mathbf{v}_2\mathbf{v}_2}| - \frac{1}{2}\left[(\mathbf{y}_{i2} - \mathbf{z}_i\mathbf{\Delta}_2)\,\mathbf{\Sigma}_{\mathbf{v}_2\mathbf{v}_2}^{-1}\,(\mathbf{y}_{i2} - \mathbf{z}_i\mathbf{\Delta}_2)^{'} \right]$$

Identification strategy and identification assumptions

The specification described above requires a set of variables ($\mathbf{z}_2$ in the notation I have been using) that are exogenous determinants of the endogenous explanatory variables but that are not determinants of the probability of cooperation. I have included in $\mathbf{z}_2$ the basicness of R&D, export intensity, industry level of incoming spillovers, industry level of appropriability, industry level of R&D intensity and industry level of cost-risk. In what follows, I define these variables, and I present the economic intuition behind these exclusion restrictions and the cases when they are included.

Kamien and Zang (2000) propose a model in which the benefit that firms obtain from incoming spillovers depends on their own R&D approach. Firms with a basic R&D approach are more likely to benefit from incoming spillovers. Following this argument, one can expect that the more basic the R&D is, the higher the score on incoming spillovers will be. The basicness of R&D is approximated by the importance of information from universities and research institutes for the innovation process. When incoming spillovers are considered an endogenous variable, basicness of R&D is included in z_2.

The strategic protection variable can be influenced by the competitive environment of the firm. Export intensity is used as a measure of the competitive environment of the firm. The underlying premise is that competition is higher in international markets than domestic ones, and only the most productive firms are able to make positive profits from exporting, and so there is self-selection into these markets (see Melitz, 2003). The export market is

while the former *bring* efficiency gains. Moreover, the use of full information or system methods in model estimation makes use of the cross-equation correlations of the disturbances.

one of substantial dynamism and exports are an important driver of firm performance (see Bernard and Jensen, 1999). So, the higher the export intensity, the higher the competition. When appropriability is considered an endogenous variable, export intensity is included in z_2.

Also included in z_2 as exclusion restrictions are industry-level measures (at the 2-digit NACE level) of the potentially endogenous variables[22]. These 2-digit NACE level variables are intended to capture the effect of unobserved industry-specific attributes on the corresponding potentially endogenous firm-specific variable.

The relevance and validity of these instruments are discussed in the next section.

5. Results

In this section, the basic model of cooperation is estimated, and the endogeneity of some explanatory variables and the relevance and validity of the instruments are tested. Once a "structure" of endogeneity is chosen, the importance of different motives for participating in cooperative R&D is discussed without distinguishing the type of partner. Next, separate models for cooperation with competitors, cooperation with suppliers or customers, and cooperation with research institutions are estimated. Before this, Table 5 gives descriptive statistics on the main variables. As expected, most of the mean values are higher for cooperating than for non-cooperating firms.

a) Dealing with the endogeneity.—

Table 6 shows the estimated coefficients of the independent variables for probit models. Standard errors are estimated for these coefficients. Regression a ignores endogeneity and shows the results of a one-step probit model (single-equation probit), regressions b to d show the results of 2SCML estimations, while regression e shows the results of CML estimation.

[22] The idea of using industry levels as instruments is conventional in microeconometric literature (see, for example, Pakes, 1983).

Testing the endogeneity

Three different "structures" of endogeneity are considered a priori. Firstly, and following CV, regression b shows the 2SCML estimations considering incoming spillovers, appropriability and R&D intensity to be endogenous variables. Coefficients accompanying residuals of first-step regressions for incoming spillovers and appropriability are significant. And hence, exogeneity of these two variables is rejected. Meanwhile, the exogeneity of R&D intensity cannot be rejected.

Secondly, in regression c, I also consider cost-risk an endogenous variable. Again, we reject the exogeneity of incoming spillovers and appropriability, while the exogeneity of R&D intensity cannot be rejected. The exogeneity of cost-risk is rejected.

Finally, and due to the previous results, in regression d I consider incoming spillovers, appropriability and cost-risk to be endogenous variables. Consistent with the previous findings, the exogeneity of these three variables is rejected.

Note that, when estimating the model by the CML exogeneity of incoming spillovers, appropriability and cost-risk are also rejected (see regression e).This is the preferred "structure" of endogeneity and the one used to obtain the marginal effects.

Does the "structure" of endogeneity matter?

The "structure" of endogeneity is crucial and has important consequences for the significance and sign of the estimated coefficients.

Firstly, considering cost-risk to be an endogenous variable has an important effect on its sign. When it is "correctly" considered to be an endogenous variable, I find that it has a positive and significant effect (see regressions c, d and e), while, this variable presents a negative and significant effect when it is taken as exogenous (see regression b). Additionally, the sign of the variable complementarities seems to depend on the endogeneity of cost-risk. It has a positive and significant effect if cost-risk is "correctly" considered to be an endogenous variable (see regressions c, d and e), while it shows a negative and significant effect when the endogeneity of cost-risk is not taken into account (see regressions a and b).

Secondly, the character of R&D intensity also seems to affect the results. When it is "correctly" considered to be an exogenous variable, we find that it has a positive and significant effect on the probability of cooperation (see regression d), while this variable loses its significance when it is considered to be endogenous (see regressions *b* and *c*).

Also, the sign and significance of the industry level of legal protection depend on considering the endogeneity problem. The single-equation probit is the only case where this variable is not significant, while the other estimates show a negative and significant effect.

Finally, and fortunately, the character of cost-risk and R&D intensity does not seem to affect the sign and significance of incoming spillovers and appropriability (compare regressions *b*, *c*, *d* and *e*).

Testing the relevance and validity of instruments

A plausible instrument must satisfy two conditions: relevance and validity. The relevance condition can be tested by examining the results of the first-step regressions. Table A1 shows the first-step regressions from which the residuals of incoming spillovers, appropriability and cost-risk for regression *d* in Table 6 have been obtained[23]. As expected, each instrument is significant in the first-step regression in question. The F tests for joint significance of the exclusion restrictions in the first-step regression for incoming spillovers, appropriability and cost risk are, respectively, 66.57, 16.40 and 13.87, which allows me to reject the null hypothesis.

In addition, Table A1 shows two different R^2 as measures of the relevancy of instruments, i.e., the partial R^2 (R_p^2) and the corrected partial R^2 ($\overline{R}_p^2$)[24]. For the appropriability and the cost-risk regressions, our estimations yield larger values for $\overline{R}_p^2$ than for R_p^2. In the case of incoming spillovers regression, R_p^2 is slightly greater than $\overline{R}_p^2$. Showing these results, I can conclude that the instruments have enough relevance to explain all the endogenous

[23]First-step regressions from which the residuals of the endogenous variables for regressions *b* and *c* in Table 6 have been obtained are available upon request.

[24]The R_p^2 (see, for example Bound, Jaeger and Baker, 1995) is the R^2 of the first-step regressions with the included instruments partialled out (note that equations (1) and (3) include common exogenous variables). The $\overline{R}_p^2$, proposed by Shea (1997) takes the correlations among the instruments into account.

regressors.[25]

In my estimation framework, testing the orthogonality condition is more problematic. The usual tests of overidentifying restrictions applied in IV or GMM estimation are not valid in a probit estimation framework. To test the orthogonality condition, a regression of the generalized residuals obtained from estimate d in Table 6 on the exclusion restrictions is shown in Table A2[26]. Only the industry level of cost is weakly significant (with an associated t-ratio equal to 1.64). This regression gives some faith in the instruments used. However, the validity of this regression for testing the orthogonality condition is not conclusive, and other tentative "experiments" have not been so optimistic[27].

I assume that it is difficult to find perfectly exogenous instruments within the CIS, where every question is closely related and, moreover, cross-section data is used. In what follows, and taking this caveat into account and having found some evidence about the orthogonality of the instruments, I will assume the validity of the instruments and the results obtained will be conditional on this assumption.

Additionally, two arguments are in my favor. Firstly, when the instruments used are not perfectly exogenous, the inconsistency of IV estimates depends on the relevance of the instruments[28]. The lower the relevance, the higher the inconsistency. And thus, the high relevance of my instruments can mitigate the inconsistency with not-perfectly exogenous instruments. Secondly, assuming the existence of invalid instruments, Hahn and Hausman (2005) find that the 2SLS has a smaller finite sample bias and MSE than the OLS under a wide range of conditions. So, in such a context of not-perfectly exogenous instruments, the 2SLS does better than the OLS in many cases[29].

[25] See Baum, Schaffer and Stillman (2003) for a comparative interpretation of R_p^2 and $\overline{R}_p^2$.

[26] I am grateful to an anonymous referee for this suggestion.

[27] Considering the case of a linear probability model, the Sargan test of overidentifying restrictions rejects the joint null hypothesis of correct model specification and the validity of the overall set of instruments.

[28] See, for example, Buse (1992); Hall, Rudebusch and Wilcox (1996); Nelson and Startz (1990a, 1990b); and Staiger and Stock (1997) for the study of the consequences of low relevance of instruments in an instrumental variables estimation context.

[29] The conditions under which the 2SLS is still preferred to the OLS are derived for a linear model with one endogenous variable, and I cannot check them in my framework.

b) Determinants for cooperation.—

Table 7 shows the impact of the explanatory variables considered throughout this study on the probability of R&D cooperation. Regression *a* pays no attention to endogeneity problems, while regressions *b* and *c* estimate the model by 2SCML and CML respectively, considering incoming spillovers, appropriability and cost-risk to be endogenous variables. The preferred outcome is estimate *c*. Estimates *a* and *b* are used for checking the importance of the estimation method on the results. I find that the overall picture of the importance of the explanatory variables depends on the estimation method. In this sense, two-step procedures overestimate the importance of spillovers.

I can conclude that incoming spillovers and appropriability have a positive and significant impact on the probability of cooperation, although the impact of the effectiveness of strategic methods is almost double. In the first place, the higher incoming spillovers are, the greater the scope for learning within cooperative R&D agreements, and hence the marginal profit to be derived from cooperation. Secondly, the more effective strategic protection is, the better firms control the outflow of commercially sensitive information, and the more likely they are to engage in cooperative agreements. Fortunately, the sign and significance of these variables do not depend on the estimation method, but the magnitudes clearly vary according to the method. Above all, 2SCML and CML yield very different pictures of the impact of incoming spillovers on R&D cooperation.

It seems that the industry level of legal protection has a negative effect on R&D cooperation. A high level of legal protection methods in an industry may hamper the internalization of information flows between firms through cooperation in R&D, and hence their negative effect on this kind of practice. Taking this together with the findings on appropriability, it may be that cooperative activity is a method of internalizing outgoing knowledge flows in industries where legal protection methods are weak, and for firms for whom more strategic methods of appropriating returns are more important.

Cost-risk sharing is the most important determinant for cooperation[30]. This variable

[30]This fact is clear when estimating by CML (see regression *c*). 2SCML estimation does not yield a clear picture about the importance of the determinants for R&D cooperation. When estimating by 2SCML,

has the greatest impact on the probability of cooperation (with a marginal effect equal to 0.687). This fact possibly stresses the lack of external private finance for innovative activity and the lack of venture capital investment, which is particularly true in Spain.

The effect of firm size is positive and significant, with evidence of a concave relation. In this case, the estimated marginal effect is similar among the estimation methods.

The hypothesis that firms with a higher availability of technological know-how are more likely to cooperate is confirmed. Finally, R&D intensity seems to lose significance when estimating by CML (the associated t-ratio is 1.42).

c) Determinants for cooperation with different types of partners.—

As shown in Section 2, most firms in the sample maintain agreements with different partners. For example, it is important to take into account that when I am analyzing the subsample of firms that cooperate with research institutions, I am considering almost the whole sample of cooperating firms.

Table 8 presents the marginal effects for CML estimations of separate models for cooperation with different types of partners. I consider incoming spillovers, appropriability and cost-risk to be endogenous variables[31].

The effectiveness of strategic protection has a significant and positive effect on cooperation with the three types of partners. The higher the control over the information flows out of the firm (through strategic protection methods), the higher the probability of cooperation with any type of partner. Moreover, apart from the level of cooperation in the industry, appropriability is the most important determinant for cooperation with competitors. Only those firms with very effective strategic protection methods will share "knowledge" with their competitors.

---

appropriability, cost-risk and industry level of cooperation have impacts around 0.6 (see regression *b*).

[31] Table A3 shows the tests for endogeneity. In some cases, I find only weak evidence for endogeneity of incoming spillovers and appropriability. However, for consistency, I still consider these variables endogenous. Note that, when analyzing the pooled cooperation decision, the exogeneity of R&D intensity is not rejected with an estimated coefficient accompanying residuals of the first-step regression for R&D intensity smaller than its estimated standard error (see regressions *b* and *c* in Table 6).

Coinciding with CV's findings, incoming spillovers seem to have an effect only on cooperation with research institutions. Firms which find publicly available information more important for their innovation process are more likely to benefit from cooperation with research institutions.

The effectiveness of the industry level of legal protection methods has only a significant and negative effect on cooperation with research institutions. It seems that a high level of these types of protection methods hampers the internalization of information flows between firms and research institutions more than with the other types of partners.

For cooperation with suppliers and customers and cooperation with research institutions, cost-risk sharing is the most important determinant for cooperation. Also, availability of technological know-how has a positive and significant effect on cooperation.

For cooperation with suppliers or customers, we find a positive and (weak) significant effect of R&D intensity. This variable loses significance for cooperation with competitors and cooperation with research institutions. For the three types of cooperative agreements, firm size is an important determinant for R&D cooperation.

The empirical results are summarized and compare to the hypothesis in Table 4.

6. Conclusions

This paper is aimed at exploring the determinants for R&D cooperation using a sample of Spanish manufacturing firms. A first step focuses on studying the endogeneity of the explanatory variables which in other papers are assumed to be endogenous a priori. We find evidence supporting endogeneity of spillovers and the importance of cost-risk as a hampering factor for the innovation process. The choice of an appropriate "structure" of endogeneity is revealed to be crucial in the significance and sign of some of the estimated effects. In this sense, cost risk sharing and complementarities have only the expected positive effect on R&D cooperation when the appropriate "structure" of endogeneity is imposed, while if this "structure" is not imposed, these variables have a negative effect.

I also find that the overall picture of the importance of the explanatory variables on the probability of R&D cooperation depends on the estimation technique. Specifically, two-step procedures overestimate the importance of spillovers and underestimate the impact of cost and risk-sharing reasons. So, in obtaining the final estimated effects, I apply a more efficient method, i.e., CML estimation.

Evidence supporting the existence of important and separately identifiable effects of incoming spillovers and appropriability on R&D cooperation is obtained: the higher incoming spillovers are and the more effective the strategic appropriation methods of the returns from innovation is, the higher the probability of R&D cooperation. However, and in a departure from other empirical works, the level of legal protection in the industry has a negative effect on R&D cooperation.

In spite of the importance of spillovers, I find that cost-risk sharing is the most important determinant for R&D cooperation. This fact possibly stresses the lack of external private finance for innovative activity and the lack of venture capital investment, which is particularly true in Spain.

Results also show that firm size and the availability of technological know-how within the firm are significant and positive determinants for R&D cooperation.

The results are not so clear when analyzing cooperation with each different type of partner. Most firms have simultaneous agreements with different types of partners and this makes identification difficult. However, two principal ideas can be advanced. First of all, for cooperation with suppliers and customers and cooperation with research institutions, cost-risk sharing is the most important determinant for cooperation. Secondly, effectiveness of strategic protection methods is the most important determinant for cooperation with competitors.

Acknowledgements

I would like to thank Jordi Jaumandreu for his valuable suggestions and discussion. I also wish to acknowledge useful comments by Pedro Albarrán, Manuel Arellano, Andrés Barge, Aija Leiponen, Georg Licht, Reinhilde Veugelers, other members of the IEEF research project and the audiences at the ESSID 2003 (Corcega), and at the 32^{nd} EARIE meeting (Porto). I thank the editor, Stephen Martin, and two anonymous referees for useful comments. This research is part of the project "Innovation and Employment in European Firms: Microeconometric Evidence," supported by the European Commission under contract IHPSE-CT-2001-00047. I acknowledge support from project SEJ2004-02525/ECON and from FBBVA. Thanks is due to the INE for access to the data. Errors are mine.

Appendix A: Variable definitions

Appropriability: Sum of the scores of the following strategic methods for protecting inventions or innovations (number between 1 (high) and 4 (not used)): Secrecy; Complexity of design; Lead-time advantage on competitors. Rescaled between 0 (not used) and 1 (high).

Basicness of R&D: Sum of the scores of importance of the following information sources for innovation process (number between 1 (high) and 4 (not used)): Universities; government or private non-profit research institutes. Rescaled between 0 (not used) and 1 (high).

Complementarities: Importance of lack of information on technology as an obstacle to innovation (number between 1 (high) and 4 (not relevant). Rescaled between 0 (high) and 1 (not-relevant).

Cooperation: Variable which takes the value 1 if the firm cooperates with suppliers, customers, competitors, commercial laboratories/R&D enterprises, universities, or government or private non-profit research institutes.

Cooperation with Competitors: Variable which takes the value 1 if the firm cooperates with competitors.

Cooperation with Research Institutions: Variable which takes the value 1 if the firm cooperates with commercial laboratories/R&D enterprises, universities, or government or private non-profit research institutes.

Cooperation with Suppliers or Customers: Variable which takes the value 1 if the firm cooperates with suppliers or customers.

Cost-Risk: Sum of the scores of importance of the following obstacles to innovation process (number between 1 (high) and 4 (not relevant)): Innovation costs too high; Lack of appropriate sources of finance; Excessive perceived economic risks. Rescaled between 0 (not relevant) and 1 (high).

Export intensity: Export share in total turnover.

Incoming Spillovers: Sum of the scores of importance of the following information sources for innovation process (number between 1 (high) and 4 (not used)): Professional conferences, meetings and journals; Fairs and exhibitions. Rescaled between 0 (not used) and 1 (high).

Industry Level of Appropriability: Mean of Appropriability at the industry level. Industry is defined at 2-digit NACE.

Industry Level of Cooperation: Mean of Cooperation at the industry level. Industry is defined at 2-digit NACE.

Industry Level of Cooperation with Competitors: Mean of Cooperation with competitors at the industry level. Industry is defined at 2-digit NACE.

Industry Level of Cooperation with Research Institutions: Mean of Cooperation with research institutions at the industry level. Industry is defined at 2-digit NACE.

Industry Level of Cooperation with Suppliers or Customers: Mean of Cooperation with suppliers or customers at the industry level. Industry is defined at 2-digit NACE.

Industry Level of Cost-Risk: Mean of Cost-risk at the industry level. Industry is defined at 2-digit NACE.

Industry Level of Incoming Spillovers: Mean of Incoming Spillovers at the industry level. Industry is defined at 2-digit NACE.

Industry Level of Legal Protection: Mean of Legal Protection at the industry level. Legal Protection is the sum of the scores of the following legal methods for protecting inventions or innovations (number between 1 (high) and 4 (not-used)): Patents; Registration of design patterns; Trademarks; Copyright. Rescaled between 0 (not-used) and 1 (high). Industry is defined at 2-digit NACE.

Industry Level of R&D intensity: Mean of R&D intensity at the industry level. Industry is defined at 2-digit NACE.

R&D intensity: Ratio between intramural R&D expenditure and turnover.

Size: Log of number of employees.

Appendix B: Econometric details

Let y_1^* represent a firm's unobservable propensity to cooperate in R&D. y_1^* is assumed to be a linear function of the previously observed explanatory variables. Let y_1 equal 1 if the firm cooperates.

I assume that the (possible) endogenous explanatory variables ($\mathbf{y}_2$) are a function of the exogenous variables that determine cooperation ($\mathbf{z}_1$), a set of other exogenous variables ($\mathbf{z}_2$), and an error term ($\mathbf{v}_2$).

So, the model can be written as follows:

$$y_1^* = \mathbf{z}_1\boldsymbol{\delta}_1 + \mathbf{y}_2\boldsymbol{\alpha}_1 + u_1 \tag{1.a.}$$

$$\mathbf{y}_2 = \mathbf{z}_1\boldsymbol{\Delta}_{21} + \mathbf{z}_2\boldsymbol{\Delta}_{22} + \mathbf{v}_2 = \mathbf{z}\boldsymbol{\Delta}_2 + \mathbf{v}_2 \tag{2.a.}$$

$$y_1 = 1\left[y_1^* > 0\right] \tag{3.a.}$$

Where $\mathbf{z}_1, \mathbf{y}_2, \mathbf{z}_2$ and $\mathbf{z}$ are $1 \times p, 1 \times m, 1 \times k$ and $1 \times (p+k)$ vectors, respectively. Note that $\boldsymbol{\Delta}_{21}, \boldsymbol{\Delta}_{22}$ and $\boldsymbol{\Delta}_2$ are $p \times m, k \times m$ and $(p+k) \times m$ matrices, respectively.

$u_1$and $\mathbf{v}_2$ have a joint normal distribution with mean zero and finite positive covariance matrix:

$$\boldsymbol{\Omega} \equiv \begin{bmatrix} \sigma_{u_1}^2 & \boldsymbol{\Sigma}_{u_1\mathbf{v}_2} \\ \boldsymbol{\Sigma}_{\mathbf{v}_2u_1} & \boldsymbol{\Sigma}_{\mathbf{v}_2\mathbf{v}_2} \end{bmatrix}$$

The most convenient normalization is:

$$\sigma_{u_1}^2 = Var(u_1) = 1$$

This is the normalization imposed by Wooldridge (2002), and is different from that used by Rivers and Vuong (1988), who use the normalization

$$Var(y_1^* \mid z, y_2) = \sigma_{u_1}^2 - \boldsymbol{\theta}_1^{\prime}\boldsymbol{\Sigma}_{\mathbf{v}_2\mathbf{v}_2}\boldsymbol{\theta}_1 = 1$$

Under joint normality of $(u_1, \mathbf{v}_2)$, I can write

$$u_1 = \mathbf{v}_2\boldsymbol{\theta}_1 + e_1 \tag{4.a.}$$

where $\boldsymbol{\theta}_1 = \boldsymbol{\Sigma}_{\mathbf{v_2 v_2}}^{-1} \boldsymbol{\Sigma}_{\mathbf{v}_2 u_1}$. To obtain the joint distribution of $(y_1, \mathbf{y}_2)$, conditional on $\mathbf{z}$, recall that

$$f(y_1, \mathbf{y}_2 \mid \mathbf{z}) = f(y_1 \mid \mathbf{y}_2, \mathbf{z}) f(\mathbf{y}_2 \mid \mathbf{z}) \tag{5.a.}$$

Since $\mathbf{v}_2$ has a joint normal distribution with mean zero and covariance matrix $\Sigma_{\mathbf{v_2 v_2}}$, the joint density $f(\mathbf{y}_2 \mid \mathbf{z})$ is easy to write down:

$$f(\mathbf{y}_2 \mid \mathbf{z}) = (2\pi)^{-\frac{m}{2}} \left|\boldsymbol{\Sigma}_{\mathbf{v_2 v_2}}\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y}_2 - \mathbf{z}\boldsymbol{\Delta}_2) \boldsymbol{\Sigma}_{\mathbf{v_2 v_2}}^{-1} (\mathbf{y}_2 - \mathbf{z}\boldsymbol{\Delta}_2)'\right\} \tag{6.a.}$$

I can also derive the conditional density of y_1 given $(\mathbf{y}_2, \mathbf{z})$. Because of joint normality of $(u_1, \mathbf{v}_2)$, e_1 is also normally distributed with $E(e_1) = 0$ and $Var(e_1) = Var(u_1) - \boldsymbol{\theta}_1' \boldsymbol{\Sigma}_{\mathbf{v_2 v_2}} \boldsymbol{\theta}_1 = 1 - \boldsymbol{\theta}_1' \boldsymbol{\Sigma}_{\mathbf{v_2 v_2}} \boldsymbol{\theta}_1$

Since (1.a.) and (4.a.), I can write

$$y_1^* = \mathbf{z}_1 \boldsymbol{\delta}_1 + \mathbf{y}_2 \boldsymbol{\alpha}_1 + \mathbf{v}_2 \boldsymbol{\theta}_1 + e_1 \tag{7.a.}$$

$$e_1 \mid \mathbf{z}, \mathbf{y}_2, \mathbf{v}_2 \sim N\left(0, 1 - \boldsymbol{\theta}_1' \boldsymbol{\Sigma}_{\mathbf{v_2 v_2}} \boldsymbol{\theta}_1\right) \tag{8.a.}$$

Since $\mathbf{v}_2 = \mathbf{y}_2 - \mathbf{z}\ \boldsymbol{\Delta}_2$ and $y_1 = 1\left[y_1^* > 0\right]$

$$P(y_1 = 1 \mid \mathbf{y}_2, \mathbf{z}) = \boldsymbol{\Phi}\left(\frac{\mathbf{z}_1 \boldsymbol{\delta}_1 + \mathbf{y}_2 \boldsymbol{\alpha}_1 + (\mathbf{y}_2 - \mathbf{z}\boldsymbol{\Delta}_2) \boldsymbol{\theta}_1}{\left(1 - \boldsymbol{\theta}_1' \boldsymbol{\Sigma}_{\mathbf{v_2 v_2}} \boldsymbol{\theta}_1\right)^{\frac{1}{2}}}\right) \tag{9.a.}$$

Let w denote the term inside $\boldsymbol{\Phi}(\cdot)$ in equation (9.a.). Then I derive

$$f(y_1 \mid \mathbf{y}_2, \mathbf{z}) = \{\boldsymbol{\Phi}(\cdot)\}^{y_1} \{1 - \boldsymbol{\Phi}(\cdot)\}^{1-y_1} \tag{10.a.}$$

Substituting (6.a.) and (10.a.) in (5.a.), I can write

$$\begin{aligned} f(y_1, \mathbf{y}_2 \mid \mathbf{z}) &= \{\boldsymbol{\Phi}(\cdot)\}^{y_1} \{1 - \boldsymbol{\Phi}(\cdot)\}^{1-y_1} (2\pi)^{-\frac{m}{2}} \left|\boldsymbol{\Sigma}_{\mathbf{v_2 v_2}}\right|^{-\frac{1}{2}} \\ &\quad \exp\left\{-\frac{1}{2}(\mathbf{y}_2 - \mathbf{z}\boldsymbol{\Delta}_2) \boldsymbol{\Sigma}_{\mathbf{v_2 v_2}}^{-1} (\mathbf{y}_2 - \mathbf{z}\boldsymbol{\Delta}_2)'\right\} \end{aligned} \tag{11.a.}$$

and so the log likelihood for observation i is

$$\begin{aligned} & y_{i1} \log \boldsymbol{\Phi}(w_i) + (1 - y_{i1}) \log\left[1 - \boldsymbol{\Phi}(w_i)\right] \\ & -\frac{m}{2}\log(2\pi) - \frac{1}{2}\log\left|\boldsymbol{\Sigma}_{\mathbf{v_2 v_2}}\right| - \frac{1}{2}\left[(\mathbf{y}_{i2} - \mathbf{z}_i \boldsymbol{\Delta}_2) \boldsymbol{\Sigma}_{\mathbf{v_2 v_2}}^{-1} (\mathbf{y}_{i2} - \mathbf{z}_i \boldsymbol{\Delta}_2)'\right] \end{aligned} \tag{12.a.}$$

where I understand that w_i depends on the parameters $(\boldsymbol{\delta}_1, \boldsymbol{\alpha}_1, \boldsymbol{\Delta}_2, \boldsymbol{\theta}_1)$:

$$w_i \equiv \frac{\mathbf{z}_{i1}\boldsymbol{\delta}_1 + \mathbf{y}_{i2}\boldsymbol{\alpha}_1 + (\mathbf{y}_{i2} - \mathbf{z}_i\boldsymbol{\Delta}_2)\,\boldsymbol{\theta}_1}{\left(1 - \boldsymbol{\theta}_1^{'}\boldsymbol{\Sigma}_{\mathbf{v}_2\mathbf{v}_2}\boldsymbol{\theta}_1\right)^{\frac{1}{2}}}$$

Summing expression (12.a.) across all i and maximizing with respect to all parameters gives the MLEs of $\boldsymbol{\delta}_1, \boldsymbol{\alpha}_1, \boldsymbol{\Delta}_2, \boldsymbol{\Sigma}_{\mathbf{v}_2\mathbf{v}_2}, \boldsymbol{\theta}_1$. The estimate of $\boldsymbol{\theta}_1$ can be used to test for endogeneity of $\mathbf{y}_2$.

Notice that, if u_1 and v_2 are uncorrelated and thus $\theta_1 = 0$, the log likelihood function in equation (12.a.) can be broken into two terms. The first line would be the log likelihood function for a single equation probit associated with y_1, and the second line would be the log likelihood function for the normal linear least-squares model associated with y_2. Thus, if $\theta_1 = 0$, there is no gain in considering the simultaneous equation model. If $\theta_1 \neq 0$, however, the single-equation model and the simultaneous equation model can yield very different coefficient estimates.

References

Abramowsky, L., Jaumandreu, J., Kremp, E. and Peters, B., (2004). "National Differences in Innovation Behaviour: Facts and Explanations." mimeo.

Abramovsky, L, Harrison, R. and Simpson, H., (2004). "Increasing Innovative Activity in the UK? Where Now for Government Support for Innovation and Technology Transfer." IFS Briefing Note No. 53.

Acosta, J. and Modrego, A. (2001), "Public Financing of Cooperative R&D Projects in Spain: the Concerted Projects under the National R&D Plan." *Research Policy*, 30, pp. 625-641.

Baum, C.F., Schaffer, M.E. and Stillman. S., (2003). "Instrumental Variables and GMM: Estimation and Testing." Boston College Working Paper 545.

Becker, W. and Dietz, J., (2004). "R&D Cooperation and Innovation Activities of Firms: Evidence for the German Manufacturing Industry." *Research Policy*, 33, pp. 209-223..

Belderbos, R., Carree, M., Diederen, B., Lokshin, B. and Veugelers., R., (2004). "Heterogeneity in R&D Cooperation Strategies." *International Journal of Industrial Organization*, 22, pp. 1237-1263.

Bernard, A.B. and Jensen, J.B., (1999). "Exceptional Exporter Performance: Cause, Effect, or Both?" *Journal of International Economics*, 47, pp. 1-25.

Bound, J., Jaeger, D. and Baker, R., (1995). "Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak." *Journal of the American Statistical Association,* 90, pp. 443-450.

Buse, A., (1992). "The Bias of Instrumental Variable Estimators." *Econometrica,* 60(1), pp. 173-180.

Cassiman, B. and Veugelers, R., (2002). “R&D Cooperation and Spillovers: Some Empirical Evidence from Belgium.” *American Economic Review,* 92(4), pp. 1169-1184.

Cohen, W.M. and Levinthal, D.A., (1989). “Innovation and Learning: The Two Faces of R&D.” *Economic Journal,* 99, pp. 569-596.

———— (1990). “Absorptive Capacity: a New Perspective on Learning and Innovation.” *Administrative Science Quarterly*, 35(1), Special Issue: Technology, Organizations, and Innovation, pp. 128–152.

Colombo, M. G. and Gerrone, P., (1996). “Technological Cooperative Agreements and Firm*s'* R&D intensity. A note on causality relations.” *Research Policy,* 25(6), pp. 923-932.

Fritsch, M. and Lukas, R., (2001). “Who Cooperates on R&D?” *Research Policy*, 30, pp. 297-312.

Griffith, R., Redding, S. and Van Reenen, J., (2003). “R&D and Absorptive Capacity: from Theory to Data.” *The Scandanavian Journal of Economics*, 105(1), pp. 99-118.

———— (2004). “Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries.” *Review of Economics and Statistics*, 86(4), pp. 883-895.

Griliches, Z., (1992). “The Search for R&D Spillovers.” *The Scandanavian Journal of Economics*, 94, pp. 29-47.

Hagedoorn, J., (1993). “Understanding the Rationale of Strategic Technology Partnering: Interorganizational Modes of Cooperation and Sectoral Differences.” *Strategic Management Journal,* 14(5), pp. 371-385.

Hahn, J. and Hausman, J., (2005). “Instrumental Variable Estimation with Valid and Invalid Instruments.” *Annales d'Economie et Statistique,* forthcoming.

Hall, A., Rudebusch, G. and Wilcox, D., (1996). “Judging Instrument Relevance in Instrumental Variables Estimation.” *International Economic Review,* 37, pp. 283-298.

Jaffe, A., (1986). "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value." *American Economic Review*, 75(6), pp. 984-1002.

Kaiser, U., (2002a). "An Empirical Test of Models Explaining Research Expenditures and Research Cooperation: Evidence for the German Service Sector." *International Journal of Industrial Organization*, 20, pp. 747-774.

——— (2002b). "R&D with Spillovers and Endogenous Absorptive Capacity." *Journal of Institutional and Theoretical Economics*, 158, pp. 286-303.

Kamien, M. I., Müller, E. and Zang, I., (1992). "Research Joint Ventures and R&D Cartels." *American Economic Review*, 82(5), pp. 1293–1306.

Kamien, M. I. and Zang, I., (2000). "Meet Me Halfway: Research Joint Ventures and Absorptive Capacity." *International Journal of Industrial Organization,* 18(7)*,* pp. 995-1012.

Maddala, G.S., (1983). Limited Dependent and Qualitative Variables in Econometrics. Cambridge University Press.

Melitz, M.J., (2003). "The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 71(6), pp. 1695-1725.

Miotti, L. and Sachwald, F., (2003). "Cooperative R&D: Why and with Whom? An Integrated Framework of Analysis." *Research Policy*, 32, pp. 1481-1499.

Nelson, C. and Startz, R., (1990a). "Some Further Results on the Exact Small Sample Properties of the Instrumental Variables Estimator." *Econometrica,* 58, pp. 967-976.

——— (1990b). "The Distribution of the Instrumental Variables Estimator and its t-Ratio When the Instrument Is a Poor One." *Journal of Business,* 58, pp. 125-140.

OECD and Eurostat (1997). Oslo-Manual, Proposed Guidelines for Collecting and Interpreting Technological Innovation Data, Paris.

Pakes, A., (1983). "On Group Effects and Errors in Variables Aggregation." *Review of Economics and Statistics*, 65(1), pp. 168-173.

Rivers, D. and Vuong, Q. H., (1988). "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." *Journal of Econometrics,* 39, pp. 347-366.

Röller, L.H., Siebert, R. and Tombak, M., (2002). "RJVs and Firm Heterogeneities." *CEPR Discussion Paper.*

Romer, P.M., (1990). "Endogenous Technological Change." *Journal of Political Economy,* 98(5), pp. 71-102.

Sakakibara, M., (1997). "Heterogeneity of Firms' Capabilities and Cooperative Research and Development: An Empirical Examination of Motives." *Strategic Management Journal,* 18, Special Issue, pp. 134-164.

Shea, J., (1997). "Instrument Relevance in Multivariate Linear Models: A Simple Measure." *Review of Economics and Statistics,* 79(2), pp. 348-352.

Staiger, D. and Stock, J., (1997). "Instrumental Variables Regression with Weak Instruments." *Econometrica,* 65(3), pp. 557-586.

Tether, B.S., (2002). "Who Cooperates for Innovation, and Why. An empirical Analysis." *Research Policy,* 31, pp. 947-967.

Tyler, B.B. and Steensma, H. K., (1995). "Evaluating Technological Collaborative Opportunities: A Cognitive Modeling Perspective." *Strategic Management Journal,* 16, Special Issue, pp. 43-70.

Veugelers, R., (1997). "Internal R&D Expenditures and External Technology Sourcing." *Research Policy,* 26(3), pp. 303-316.

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

Table 1. Sample Statistics
(Number and percentage of firms)

Manufacturing Firms	6026
Innovating Firms	2518 (41.8%)[1]
Non-cooperating Firms	2042 (81.1%)[2]
Cooperating Firms (at least one cooperative R&D agreement)	476 (18.9%)[2]
Firms Cooperating with Competitors	184 (7.3%)[2]
Firms Cooperating with Suppliers or Customers	316 (12.5%)[2]
Firms Cooperating with Research Institutions	425 (16.9%)[2]

[1]percentage with respect to manufacturing firms
[2]percentage with respect to innovating firms

Table 2. Cooperative R&D Agreement Combinations
(Number and percentage[1] of firms)

Cooperating firms with three types of agreements	159 (33.4%)

Cooperating firms with one or two types of agreements

	Competitors	Suppliers or Customers	Research Institutions
Competitors	7 (1.5%)	9 (1.9%)	9 (1.9%)
Suppliers or Customers	–	35 (7.3%)	113 (23.8%)
Research Institutions	–	–	144 (30.2%)

[1] percentage with respect to Cooperating Firms (firms with at least one cooperative R&D agreement)

Table 3. Number of innovating manufacturing firms by size and sector[1,2]

	Less than 200 employees	200 or more employees
Transport equipment	85 (14)	96 (38)
Chemicals	150 (49)	109 (52)
Machinery	139 (17)	46 (13)
Electrical	196 (36)	88 (39)
High-technology sectors	570 (116)	339 (142)
Food, beverages and tobacco	166 (15)	126 (33)
Textile and leather	197 (9)	45 (8)
Wood and paper	228 (10)	50 (11)
Rubber and plastic	81 (6)	37 (9)
Non-metallic mineral products	106 (11)	50 (22)
Metallic products	219 (24)	93 (39)
NEC and recycling	181 (16)	30 (5)
Low-technology sectors	1178 (91)	431 (127)
Manufacturing firms	1748 (207)	770 (269)

[1]number of innovating manufacturing firms with at least one cooperative R&D agreement between brackets

[2]Transport equipment (NACE 34-35); Chemicals (NACE 23-24); Machinery (NACE 29); Electrical (NACE 30-33); Food, beverages and tobacco (NACE 15-16); Textile and leather (NACE 17-19); Wood and paper (NACE 20-22); Rubber and plastic (NACE 25); Non-metallic mineral products (NACE 26); Metallic products (NACE 27-28); NEC and recycling (NACE 36-37)

Table 4. Summary of hypothesis and empirical results for Spain

Variable	Hypothesis	Finding
Incoming spillovers	The importance of publicly available information has a positive effect on the probability of cooperation	True
Appropriability	The effectiveness of the strategic protection methods does not have a clear effect on the probability of cooperation	Positive effect
Industry level of legal protection	The effectiveness of the legal protection methods does not have a clear effect on the probability of cooperation	Negative effect
Cost-risk	The higher the importance of cost and risk as hampering factors for innovation, the higher the probability of cooperation	True
Complementarities	The availability of technological know-how within the firm has a positive effect on the probability of cooperation	True
R&D intensity	Benefits from R&D cooperation depend on the absorptive capacity of the firm; therefore, we expect a positive effect of R&D intensity on the probability of cooperation	True for cooperation with suppliers or customers
Size	Benefits from R&D cooperation depend on the absorptive capacity of the firm; therefore, we expect a positive effect of the firm's size on the probability of cooperation	True

Table 5. Descriptive Statistics[1]

	Sample Mean (N=2518)	Mean non-cooperating Firms (N=2042)	Mean Cooperating Firms (N=476)	Mean Cooperation with Competitors (N=184)	Mean Cooperation with Suppliers or Customers (N=316)	Mean Cooperation with Research Institutions (N=425)
Incoming Spillovers	0.438	0.427	0.485***	0.515***	0.508***	0.490***
	(0.310)	(0.310)	(0.307)	(0.299)	(0.304)	(0.305)
Appropriability	0.214	0.186	0.335***	0.363***	0.374***	0.349***
	(0.316)	(0.298)	(0.362)	(0.366)	(0.366)	(0.363)
Industry Level Legal Protection	0.131	0.129	0.137***	0.139***	0.137***	0.137***
	(0.040)	(0.040)	(0.039)	(0.039)	(0.039)	(0.039)
Size	1.930	1.845	2.298***	2.385***	2.341***	2.329***
	(0.577)	(0.547)	(0.556)	(0.600)	(0.573)	(0.543)
Cost-Risk	0.477	0.468	0.515***	0.553***	0.538***	0.517***
	(0.317)	(0.322)	(0.291)	(0.278)	(0.282)	(0.288)
Complementarities	0.689	0.694	0.666	0.663	0.674	0.665
	(0.322)	(0.327)	(0.298)	(0.295)	(0.289)	(0.297)
R&D intensity	0.016	0.012	0.035***	0.055***	0.042***	0.036***
	(0.118)	(0.038)	(0.259)	(0.414)	(0.317)	(0.274)

***difference in means between cooperating and non-cooperating firms significant at 1%, **significant at 5%, *significant at 10%

We test the null hypothesis of equality of two means. Under the null hypothesis, the quantity $T = \frac{\sqrt{\frac{n_1 n_2}{n_1+n_2}}(\overline{X}_1 - \overline{X}_2)}{\sqrt{\sigma_1^2 \frac{n_1}{n_1+n_2-2} + \sigma_2^2 \frac{n_2}{n_1+n_2-2}}}$ has the distribution t with $n_1 + n_2$ degrees of freedom, where n_i, $\overline{X}_i$, and σ_i^2 (i=1, 2) are the number of observations from the sample of population i; the mean of the variable from the sample of population i; and the variance from the sample of population i, respectively

[1]standard deviations in parenthesis

Table 6. Results of Regressions for Cooperation. Testing the Endogeneity[1]

	(*a*) Cooperation (Single-Equation Probit)	(*b*) Cooperation (2SCML)	(*c*) Cooperation (2SCML)	(*d*) Cooperation (2SCML)	(*e*) Cooperation (CML)
Constant	-4.524^{***}	-4.621^{***}	-7.366^{***}	-7.609^{***}	-3.849^{***}
	(0.406)	(0.505)	(1.162)	(1.197)	(0.473)
Incoming Spillovers	0.217^{**}	2.662^{***}	1.822^{***}	2.124^{***}	0.717^{*}
	(0.109)	(0.518)	(0.576)	(0.504)	(0.386)
Appropriability	0.432^{***}	3.081^{***}	2.671^{***}	2.875^{***}	1.319^{***}
	(0.096)	(0.746)	(0.770)	(0.731)	(0.473)
Industry Level Legal Protection	0.075	-4.016^{***}	-3.323^{***}	-2.977^{***}	-1.281^{**}
	(0.860)	(1.038)	(1.075)	(1.039)	(0.650)
R&D Intensity	3.136^{***}	6.558	6.407	2.039^{***}	0.873
	(0.814)	(4.343)	(4.178)	(0.720)	(0.614)
Size	1.989^{***}	1.510^{***}	1.974^{***}	1.886^{***}	0.888^{***}
	(0.326)	(0.378)	(0.412)	(0.408)	(0.284)
Size squared	-0.287^{***}	-0.219^{***}	-0.289^{***}	-0.279^{***}	-0.132^{**}
	(0.072)	(0.077)	(0.081)	(0.082)	(0.062)
Cost-Risk	0.293^{**}	-0.572^{***}	2.749^{**}	2.903^{**}	1.953^{***}
	(0.116)	(0.139)	(1.266)	(1.318)	(0.553)
Complementarities	-0.220^{**}	-0.242^{*}	1.094^{**}	1.220^{**}	0.817^{***}
	(0.111)	(0.132)	(0.521)	(0.539)	(0.224)
Industry Level Cooperation	2.887^{***}	2.599^{***}	2.520^{***}	2.850^{***}	1.209^{***}
	(0.321)	(0.492)	(0.470)	(0.363)	(0.283)
$\theta_{\text{incoming spillovers}}$	–	-2.958^{***}	-2.117^{***}	-2.418^{***}	-0.848^{**}
		(0.535)	(0.591)	(0.522)	(0.405)
$\theta_{\text{appropriability}}$	–	-2.774^{***}	-2.358^{***}	-2.561^{***}	-1.184^{**}
		(0.751)	(0.775)	(0.736)	(0.417)
$\theta_{\text{R\&D intensity}}$	–	-4.280	-4.127	–	–
		(4.375)	(4.209)		
$\theta_{\text{cost-risk}}$	–	–	-2.578^{**}	-2.736^{**}	-1.883^{***}
			(1.269)	(1.321)	(0.564)
LL	-998.411	-895.287	-891.401	-891.867	-2217.254
χ^2	332.41^{***}	449.63^{***}	459.81^{***}	451.53^{***}	–
N	2518	2518	2518	2518	2518

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

[1]standard errors between brackets.

Table 7. Results of Regressions for Cooperation. Marginal Effects[1]

	(*a*) Cooperation (Single-Equation Probit)	(*b*) Cooperation (2SCML)	(*c*) Cooperation (CML)
Incoming Spillovers	0.049**	0.440***	0.252*
	(0.024)	(0.106)	(0.130)
Appropriability	0.097***	0.596***	0.464***
	(0.021)	(0.150)	(0.164)
Industry Level Legal Protection	0.016	−0.617***	−0.451**
	(0.194)	(0.216)	(0.226)
R&D Intensity	0.709***	0.423***	0.307
	(0.186)	(0.150)	(0.213)
Size	0.450***	0.391***	0.312***
	(0.071)	(0.082)	(0.098)
Size squared	−0.064***	−0.057***	−0.046**
	(0.015)	(0.016)	(0.022)
Cost-Risk	0.066**	0.602**	0.687***
	(0.026)	(0.273)	(0.213)
Complementarities	−0.049**	0.253**	0.288***
	(0.025)	(0.111)	(0.086)
Industry Level Cooperation	0.653***	0.591***	0.425***
	(0.071)	(0.109)	(0.090)
LL	−998.411	−891.867	−2217.254
χ^2	332.41***	451.53***	–
N	2518	2518	2518

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

[1]standard errors between brackets. The coefficients are the marginal effect of the independent variable on the probability of cooperation, ceteris paribus.

Table 8. Results of Regressions for Cooperation with different types of partners. Marginal Effects[1]

	(*a*)	(*b*)	(*c*)
	Cooperation with Competitors (CML)	Cooperation with Suppliers or Customers (CML)	Cooperation with Research Institutions (CML)
Incoming Spillovers	0.145	0.152	0.219*
	(0.096)	(0.117)	(0.131)
Appropriability	0.247*	0.445***	0.485***
	(0.140)	(0.150)	(0.161)
Industry Level Legal Protection	−0.186	−0.347	−0.437*
	(0.211)	(0.220)	(0.226)
R&D Intensity	0.213	0.327*	0.180
	(0.200)	(0.188)	(0.177)
Size	0.163*	0.239***	0.331***
	(0.087)	(0.092)	(0.097)
Size squared	0.017	−0.031	−0.052**
	(0.018)	(0.020)	(0.021)
Cost-Risk	0.414	0.591**	0.740***
	(0.278)	(0.241)	(0.200)
Complementarities	0.154	0.249**	0.315***
	(0.116)	(0.099)	(0.081)
Industry Level of Cooperation with Competitors	0.587*** (0.120)	–	–
Industry Level of Cooperation with Suppliers or Customers	–	0.438*** (0.111)	–
Industry Level of Cooperation with Research Institutions	–	–	0.398*** (0.099)
LL	−1848.246	−2062.711	−2114.616
N	2518	2518	2518

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

[1]standard errors between brackets. The coefficients are the marginal effect of the independent variable on the probability of cooperation, ceteris paribus.

Table A1. Results of first-step regressions used for constructing the reduced form residuals of Incoming Spillovers, Appropriability and Cost-Risk of Table 6, regression (d)[1]

	(a) Incoming Spillovers	(b) Appropriability	(c) Cost-Risk
Size	-0.070	0.082	-0.120^{**}
	(0.056)	(0.061)	(0.057)
Size squared	0.008	-0.009	0.015
	(0.012)	(0.015)	(0.013)
Industry Level Legal Protection	0.044	0.068	-0.087
	(0.226)	(0.247)	(0.217)
R&D Intensity	0.001	0.047	0.072^{**}
	(0.021)	(0.082)	(0.033)
Complementarities	-0.053^{***}	-0.041^{**}	-0.414^{***}
	(0.018)	(0.018)	(0.018)
Basicness of R&D	0.371^{***}	0.171^{***}	0.135^{***}
	(0.023)	(0.027)	(0.021)
Export intensity	-0.016	0.126^{***}	-0.006
	(0.023)	(0.027)	(0.022)
Industry Level of Cooperation	-0.104	-0.102	0.032
	(0.077)	(0.078)	(0.078)
Industry Level of Incoming Spillovers	0.901^{***}	0.077	0.003
	(0.126)	(0.128)	(0.129)
Industry Level of Appropriability	-0.073	0.768^{***}	0.084
	(0.198)	(0.214)	(0.195)
Industry Level of Cost-Risk	0.030	0.071	0.850^{***}
	(0.179)	(0.193)	(0.167)
Constant	0.124	-0.163	0.484^{***}
	(0.123)	(0.130)	(0.116)
R_p^2	0.104	0.037	0.018
$\overline{R}_p^2$	0.081	0.061	0.180
F(11, 2506)	34.90	20.41	66.96
F(5, 2506)[2]	66.57	16.40	13.87
N	2518	2518	2518

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

[1]Estimation method: OLS. Robust standard errors between brackets.

[2]F test for joint significance of the exclusion restrictions: basicness of R&D, export intensity, industry level of incoming spillovers, industry level of appropriability and industry level of cost-risk.

Table A2. Results of the regression of the generalized residuals of estimate d in Table 6 on the exclusion restrictions[1]

	Generalized Residuals
Basicness of R&D	0.037
	(0.028)
Export intensity	0.025
	(0.029)
Industry Level of Incoming Spillovers	−0.169
	(0.122)
Industry Level of Appropriability	−0.129
	(0.125)
Industry Level of Cost-Risk	0.183*
	(0.111)
R^2	0.002
F(5, 2513)	1.35
N	2518

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

[1]Estimation method: OLS. Robust standard errors between brackets.

Table A3. Results of Regressions for Cooperation with different types of partners. Testing the Endogeneity[1]

	(a)	(b)	(c)
	Cooperation with Competitors (CML)	Cooperation with Suppliers or Customers (CML)	Cooperation with Research Institutions (CML)
Constant	-4.161^{***}	-3.932^{***}	-3.958^{***}
	(0.645)	(0.560)	(0.464)
Incoming Spillovers	0.648	0.502	0.624^{*}
	(0.529)	(0.418)	(0.389)
Appropriability	1.105^{*}	1.467^{***}	1.381^{***}
	(0.661)	(0.517)	(0.469)
Industry Level Legal Protection	-0.832	-1.146	-1.243^{*}
	(0.971)	(0.746)	(0.649)
R&D Intensity	0.953	1.078^{*}	0.513
	(0.957)	(0.654)	(0.509)
Size	0.730^{**}	0.789^{***}	0.943^{***}
	(0.364)	(0.307)	(0.286)
Size squared	-0.075	-0.102	-0.147^{**}
	(0.077)	(0.065)	(0.062)
Cost-Risk	1.852^{**}	1.950^{***}	2.107^{***}
	(0.886)	(0.662)	(0.510)
Complementarities	0.690^{*}	0.821^{***}	0.897^{***}
	(0.384)	(0.270)	(0.204)
Industry Level of Cooperation with Competitors	2.630^{***}	–	–
	(0.690)		
Industry Level of Cooperation with Suppliers or Customers	–	1.447^{***}	–
		(0.434)	
Industry Level of Cooperation with Research Institutions	–	–	1.132^{***}
			(0.307)
$\theta_{\text{incoming spillovers}}$	-0.716	-0.541	-0.767^{*}
	(0.540)	(0.428)	(0.411)
$\theta_{\text{appropriability}}$	-0.936	-1.242^{**}	-1.240^{***}
	(0.656)	(0.512)	(0.466)
$\theta_{\text{cost-risk}}$	-1.597^{*}	-1.759^{**}	-2.049^{***}
	(0.938)	(0.686)	(0.519)
LL	-1848.246	-2062.711	-2114.616
N	2518	2518	2518

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

[1]standard errors between brackets.