

R&D Cooperation Determinants, Evidence with Chilean Firms

José Miguel Benavente*

Universidad de Chile

Rodolfo Lauterbach**

Universidad de Chile

October 19, 2007

Abstract

In this paper we study the determinants of R&D Cooperation among innovative firms in Chile using data from the 4th Chilean Innovation Survey. We have focused our attention on several variables that influence R&D Cooperation between firms. Among them we include "incoming spillovers", "easy imitation", "firm size", "cost and risk sharing", "R&D intensity" and "legal protection". From our main results we find that incoming spillovers, cost-risk sharing and legal protection affect positively and significantly the probability of cooperating among Chilean firms, while lack of information and firm size have a negative impact on the same variable. Little differences were found when distinguishing between kind of partners. Endogeneity problems present in our model have been properly addressed using a two-stage conditional maximum likelihood test (2SCML) due to Rivers and Vuong (1988).

JEL : O3, L6, L2

Keywords : R&D Cooperation, Incoming Spillovers, Maximum Likelihood Methods.

We would like to thank Ricardo Paredes, Rómulo Chumacero, Tomás Rau and all the participants of the SECHI annual meeting that took place in September 2007, in Reñaca, Chile.

*Departamento de Economía, Facultad de Economía y Administración. Campus Andrés Bello. 8330015 Santiago. Chile. Tel.:56 29783424. e-mail: jbenaven@econ.uchile.cl

**Departamento de Economía, Facultad de Economía y Administración. Campus Andrés Bello. 8330015 Santiago. Chile. Tel.:56 994416541. e-mail: rlauterbach@fen.uchile.cl

1 Introduction

R&D activities have a very important impact on the development of new products and productive processes as shown on Crepon, Duguet and Mairesse (1995). When we study the innovative process of firms it turns out to be very important to analyze not only the incentives that take a firm to allocate financial resources on R&D but also how these innovative firms spend jointly their resources on R&D with the intention of developing together some research activities that will later traduce on a process or product innovation.

There are reasons to believe that cooperation in an innovative process is an efficient manner of developing innovative products and processes mainly for two reasons. First because it avoids the same innovative process to be developed two or more times instead of one, causing an important cost reduction for all firms participating on the cooperative process. The second one is the possibility that complementing the knowledge available in two firm may turn into the development of a new process or product that would have not been possible with the information or knowledge available only for one of the firms.

If there are some specific conditions that promote R&D cooperation among Chilean firms, it will be useful to analyze them empirically in order to have a useful tool at the time of innovative policy making.

Following Cassiman and Veugelers (2002), we develop a well suited model that addresses the possible endogeneity of some variables. The estimation procedure we used has the characteristic of firstly testing the presence of endogeneity problems on a set of variables using a Two Stages Conditional Maximum Likelihood method. After this, we have applied a Probit Conditional Maximum Likelihood Method with Instrumental Variables, in order to correct for the specification problem of the model caused by the endogenous variable detected in the first stage of

our work.

Within our results we find that Incoming Spillovers affect positively and very significantly the probability of innovating. This result has been robust in all our estimation procedures. Easy Imitation also seems to affect the dependant variable in a positive way, though the reason why this result is not significant will be further discussed on the results section. Cost and Risk sharing, and Legal Protection affect positively the probability of cooperating being significant in most of the cases. On the other side Lack of Information affects negatively the probability of cooperating with other firms for innovative purposes being this impact very significant in all our estimations. Finally we conclude that larger firms cooperate significantly less than short firms in Chile suggesting that organizational problems may be representing a cooperation restraint in large firms.

The rest of this document is structured as follows. On section 2 we briefly comment on some conceptual antecedents and previous empirical findings. Section 3 explains our model. Section 4 refers to the data of the 4th Chilean Innovation Survey. Section 5 shows our main results and section 6 concludes.

2 R&D cooperation determinants

One of the most important inputs for a firm's innovative activities is knowledge. A firm that is capable of successfully incorporating outcoming information flows is probably more capable and willing of having a cooperative behavior for its innovative activities.

On the other side, firms that can't appropriate from the usefull results of their innovative activities will tend not only to practice less cooperation but also to engage less on R&D activities.

The literature on R&D cooperation determinants focuses mainly on the importance of the information flows to stimulate the firms to engage or not to engage in cooperative R&D activities. Cassiman and Veugelers (2002) have approached a distinction between incoming information flows and outgoing information flows. They show how firms are more propensive to cooperate both, when they have enough absorptive capacity to receive and use external public knowledge, and when they can appropriate of the knowledge generated by their own innovative activities.

2.1 Previous empirical findings

Cassiman and Veugelers (2002) develop a model in which they study separately the effects of Incoming Spillovers and Apropriability over the probability of participating in cooperative R&D agreements. They use data from the Belgian Community Innovation Survey. They conclude that incoming spillovers and apropiability have important and separately identifiable effects on R&D Cooperation among Belgian firms. Furthermore, they found differences in the effect of these variables over the cooperation with different types of partners.

López (2004) uses a similar approach to that developed by Cassiman and Veugelers (2002) and studies the determinants for R&D cooperation in the Span-

ish manufacturing firms. Results show a positive effect on the probability of R&D Cooperation of the importance of external information sources and the ability of the firm to appropriate the returns from innovations.

Mark and Graversen (2004) discuss the rationale for firms' choices of cooperative R&D partners which can be public research institutions or private firms. They use micro data from the Danish Centre for Studies in Research and Research Policy collected for year 2001. The determinants for R&D cooperation include R&D competence stocks, absorption capacity and R&D time horizon. They find heterogeneity in the cooperation determinants for different kinds of partners.

Atallah (2005) considers an R&D cooperation model with asymmetric firms, it focuses on cooperation among firms with different levels of efficiency. The results suggest that firms' preferences over whom to cooperate depend on spillovers and on cost differences between firms. With low (high) spillovers, a firm prefers to cooperate with the most (least) efficient among the remaining firms. As the cost differential between firms increases efficient (inefficient) firms prefer to cooperate with the most (least) efficient firm more often.

Schmidt (2005) analyzes the determinants of R&D cooperation among German manufacturing firms. Using data from the Third Community Innovation Survey, the paper focuses on the effect of information flows over the cooperation in R&D activities. It is found a positive relation between knowledge flows and cooperative activities. Moreover the paper found that firms with high intramural R&D budgets are more likely to cooperate with universities and research institutions than with suppliers and customers.

Vencatachellum and Vesaevel (2006) study the effects of spillovers on the likelihood that firms cooperate in R&D. Their results find that a firm which benefits

from higher spillovers from her rivals is more likely to cooperate horizontally in R&D. Furthermore, the impact of incoming spillovers over R&D cooperation is positive and significant only when they are above a threshold, it is to say, the value and the precision of the estimates increase with the information flows that firms report receiving from their competitors.

3 The Model

The model is based on the idea that information flows are one of the most important aspects determining R&D cooperation. Cassimann and Veugelers (2002) separate information flows between those that come from outside and come to increase firm's knowledge, and those that come from the firm and allow other agents to benefit from the firm's experience and knowledge.

It is expected that the more important are incoming knowledge flows or spillovers, more likely it is that the firm will be able to benefit from R&D agreements. When public information sources are important on the firm's knowledge improvement, it is expected that the return of cooperating will be greater because the partners are supposed to be more capable of transmitting useful information.

The effects of outgoing information flows is nevertheless uncertain and must be yet discussed since the literature does not provide a conceptual approach of this effect that is clear-cut. When the appropriability of the knowledge that has been generated into the firm is lower, this is to say, when imitating is easy, different opposite effects may take place that are not easy to disentangle.

On the first place, easy imitation generates more cooperation agreements since it is more likely that the firm is going to be able to internalize information flows that occur between firms under a cooperative agreement. We could otherwise think that these R&D agreements may be a source for setting up some rules on the use and the propriety of the knowledge that is generated within an R&D process. This way, when imitation is easy, there could be more incentives to perform R&D under cooperation agreements which are surely a way of sharing the costs of the generation of a knowledge whose benefits are anyway going to be shared by more than one firm.

But we can also expect that the effects of easy imitation or low level of appropriability affect R&D cooperation agreements in a negative way. When the possibility of becoming a free rider of other firms R&D investment exists, it is also probable that some firms will prefer not to cooperate. This way firms may rather perform only internal R&D and with no participation of other firms or institutions in order to be more able to protect the knowledge generated by their own research investment.

In this model we include a variable that represents easy imitation as an obstacle for the innovative process. This is a way of considering outgoing information flows in our estimation procedure. We expect to find no significant effects of this variable over R&D cooperation. This is because the expected effects go in both directions and we are not able to disentangle them in order to distinguish the positive and the negative effects of appropriability over cooperation.

We also include other variables in the model that are supposed to affect R&D cooperation agreements. Lack of Information as a constrain for innovating is a variable that takes value 1 if lack of information is declared by the firm as an important restraint for innovating. Lack of information being important for innovative firms could mean that they do not have yet information of what is being done by others in their research topic, so they are not in good position for cooperating with other firms or institutions. Thus, we expect this variable to affect the probability of cooperating in a negative way.

We also consider the possibility of sharing the risks and the costs that innovative activities may have when this are high and expensive as declared by the firms. When the costs and risks of innovative activities are considered high by the firms, it is more likely that they will tend to cooperate in order to produce process and product innovations in a less expensive way. For this work we have

constructed a cost risk sharing variable based on the following innovations restraints: Difficulty for finding sources of finance; High cost of innovating; High risk perceived; Payback period too long.

When a firm is protecting its knowledge by legal means, it is to say, when a firm has patents or know-how agreements, it may be more likely that it will cooperate. We include this variable in order to check if it has any impact over R&D cooperation. We consider, as in the previous literature, that firm size and R&D intensity may be R&D cooperation determinants. We have included size and size squared as well as R&D intensity and industry level of cooperation in order to control for unobservable industry level characteristics that may affect firms decision to cooperate with other firms or institutions.

As the dependent variable is dichotomic and takes values 1 or 0. We need our econometric specification to be defined as follows:

$$E(y_i) = P_r(y_i = 1) \rightarrow 1 \quad \text{when} \quad \alpha + \beta x_i \rightarrow \infty$$

and

$$E(y_i) = P_r(y_i = 1) \rightarrow 0 \quad \text{when} \quad \alpha + \beta x_i \rightarrow -\infty$$

This way we can define our main equation which under the assumption of no endogeneity problems may be properly estimated by a simple Probit Model as follows:

$$y_1 = F(\alpha + \beta_1 x_{Inc.Spill.} + \beta_2 x_{EasyIm.} + \beta_3 x_{LackInfo} + \beta_4 x_{CostRisk} + \beta_5 x_{Leg.Protec.} + \beta_6 x_{Ind.Coop.} + \beta_7 x_{Ind.Leg.Protec.} + \beta_8 x_{R\&DInt.} + \beta_9 x_{Size} + \beta_{10} x_{Size^2})$$

Where:

$y_1 = 1$ if the firm cooperates and 0 in other case.

β_1 : the effect of incoming spillovers over the probability of cooperating.

β_2 : the effect of Easy Imitation over the probability of cooperating.

- β_3 : the effect of Lack of Information over the probability of cooperating.
- β_4 : the effect of Cost and Risk sharing over the probability of cooperating.
- β_5 : the effect of Legal Protection over the probability of cooperating.
- β_6 : the Industry Level Cooperation effect over the probability of cooperating.
- β_7 : the Industry Level of Legal Protection effect over the probability of cooperating.
- β_8 : the effect of R&D intensity over the probability of cooperating.
- β_9 : the effect of size of firm (log of employees) over the probability of cooperating.
- β_{10} : the nonlinear effect of size of firm over the probability of cooperating.

As it has been shown in previous literature, some of the variables included in this model may be endogenous. According to Cassiman and Veugelers (2002), incoming and outgoing information flows may be endogenous variables. Since firms that cooperate may also try to control information flows with their R&D agreements, the fact of cooperating may also be a determinant of the incoming information flows or spillovers. Moreover, an R&D agreement may also be used by firms to control both the outgoing information flows to partner firms and to non-partner firms.

Other works like Veugelers (1997) and Colombo and Gerrone (1996) have found reasons to believe that R&D intensity may also be endogenous, we also consider their experience on this paper. Cost-risk has also been treated as an endogenous variable in Lopez (2004) because the effects of cooperating are also possible to affect how much these variables are considered by firms as obstacles to innovation.

In order to approach our model in a more consistent way we will test for the presence of endogeneity problems in the four variables that have been mentioned as possible endogenous variables in previous works that consider the determinants

of R&D agreements. We firstly estimate the model using a two stages conditional maximum likelihood method which has been recently reviewed by Wooldrige (2002). This model, that was firstly considered in Heckman and MaCurdy (1985) and Rivers and Voung (1988), has the characteristic of providing a simple way to test for the endogeneity of some of the variables considered in the model. In the first step of this procedure all the variables that are possibly endogenous are regressed one by one on all the variables that are assumed to be exogenous. The test is performed on the second step of this procedure where we consider the resids of each of the first step equations as independent variables of our main equation which has R&D cooperation as dependant variable.

We can now define a model that considers the presence of endogenous explanatory variables. We asume that the endogenous explanatory variables (y_2) are a function of the other explanatory variables described in our previous equation (X) (exogenous) and that they also depend on a set of instruments (Z).

So we can write the model as follows:

$$y_1 = F(\alpha_0 + \beta_1 x_i + \beta_2 y_2)$$

$$y_2 = F(\alpha_1 + \beta_3 x_i + \beta_4 z_i)$$

Where x_i , z_i are $k \times n$ and $p \times n$ matrices respectively while y_2 is a $n \times j$ matrix. Then β_1 and β_3 are both $1 \times k$, β_4 is $1 \times p$ and β_2 is a $1 \times j$ vector.

4 Description of the Data

The data used in this paper has been obtained from the Fourth Chilean Innovation Survey, collected by the National Institute of Statistics during the year 2005.

Table 1 shows some basic description of the database that has been used for this work. The sample is composed by 3,122 firms of which 1,494 firms (48%) has declared to have carried out innovations in the 2 year period before the survey. The survey has considered firms from the manufacturing, mining, energy and services sectors. The most innovative enterprises in Chile are those from the mining sector with a 58% of innovative firms. The sector in which the smallest proportion of firms carries out some type of innovation is the Services sector with a 43% of innovative firms.

Of all the innovative firms included in the sample only a 12.4% of them has cooperated with other innovative firms in order to perform R&D activities. Firms that cooperate with Universities are 106 representing a 7.1%, while firms that cooperate with Competitors are 96 and represent a 6.2% of the total innovative firms. Firms that cooperate with Consulting professionals are 93 and represent a 6.2% of the innovative firms. Finally, the 166 firms that cooperate with Customers represent an 11% of the innovating firms.

Table 2 presents the mean values of the most important variables of our model. Incoming Spillovers has a sample mean of 0.655 and is significantly higher for firms that cooperate with Consulting professionals. Easy Imitation has a sample mean of 0.727 and presents no significant mean difference between cooperating and non-cooperating firms. Lack of Information has a sample mean of 0.283 and has a significantly lower mean value for firms that cooperate with any kind of partner. Cost-Risk has a sample mean of 0.223. This variable has a mean that is not significantly different between cooperating and non-cooperating firms. Legal

Protection has a sample mean of 0.198 and has a significantly higher mean values for firms that cooperate with any kind of partner. Employees has a mean value of 461.5 and is significantly higher for firms that cooperate with any kind of partner. R&D Intensity has a mean of 0.012 for innovative firms and its significantly higher for firms that cooperate with at least one kind of partner and for firms that cooperate with universities.

5 Results

5.1 Preliminary Probit Regressions

Results exposed on tables 3 and 4 suggest that there is evidence to believe that Incoming Information flows determine a greater likelihood of performing R&D cooperation. This result is robust on all regressions no matter which kind of partner is being considered. The highest impact of Incoming Spillovers is over the cooperation agreements with consultants and competitors. Cooperation with suppliers and customers and cooperation with universities seem to be affected in a smaller magnitude by the existence of Incoming Spillovers.

Even though Easy Imitation has a positive sign in all regressions of tables 3 and 4, it is clear that this value is persistently not significant. This result has a very simple intuition. Since in our theoretic model we have specified this variable to impact R&D cooperation agreements in both directions, then the result is that the mean impact of outgoing spillovers is positive but is not significant. This means that we are not able to be sure that the impact captured by this coefficient is being caused by this variable. More likely this effect may be caused by both, Easy Imitation that influences the dependant variable in both directions, and other unobservable effects that are being captured at this point.

Lack of Information affects negatively and significantly the probability of having R&D cooperation agreements inside of the Chilean firms. This variable takes a higher absolute value when the dependant variable is cooperation with universities and cooperation with consultants. This result may be because firms that declare to have Lack of Information do not have yet information of what is being done by others in their research topic. So they are not prepared for cooperating with other firms or institutions.

Cost-Risk variable seems to affect cooperation agreements in a positive way. This result is consistent in all our specific-partner estimations. Cost-Risk shows a more intensive impact over cooperation with competitors and cooperation with consultants.

Legal Protection affects positively and significantly the probability of engaging in R&D agreements. This is consistent with our theoretic model previously specified. We have found that, for Chilean firms, the greater impact of this variable is on the cooperation with universities.

This model also controls by Industry Level of Cooperation, Industry Level of Legal Protection, R&D Intensity during the period, Size and Size Squared. All controls are always significant except by R&D Intensity which is significant for cooperation with all partners and for cooperating with universities.

5.2 Two-Step Endogeneity Test Results

We have used for our estimations a two-step endogeneity test that can be found in Heckman and MaCurdy (1985) and Rivers and Voung (1988). This allows us to state the direction of the causality relations for all the variables that we believe may be endogenously determined. In this case we have tested the endogeneity of four of the independent variables included in our model. These variables are; Incoming Spillovers, Appropriability, R&D Intensity and Cost-Risk.

In the first step of our test we regress each of the allegedly endogenous regressors with all the assumed independent variables of the model. Then we use the predicted value of the latent variable in the second step of the estimation procedure. The second equation has R&D cooperation as dependent variable. It includes the residuals of each of the first step equations as regressors. We check the

significance of these regressors in order to see if there are endogeneity problems.

We find that only Incoming Spillovers is endogenous. So we can proceed now to our final estimation procedure which will use instruments for Incoming Spillovers.

5.3 CML with Instrumental Variables Results

Tables 6 and 7 contain the Conditional Maximum Likelihood results with the use of instrumental variables for Incoming Spillovers. The instruments that we used are Industry Level of Incoming Spillovers¹, and Basicness of R&D².

Results show that Incoming Spillovers determine a greater probability of R&D cooperation. This result is seen on all regressions independently of which kind of partner is being considered. The highest impact of Incoming Spillovers is over the cooperation agreements with universities and consultants. Cooperation with suppliers and customers and cooperation with competitors seem to be affected in a smaller magnitude by the existence of Incoming Spillovers.

Easy Imitation has a positive signed parameter in all regressions of tables 6 and 7. Nevertheless it is clear that this value is persistently not significant. Since in our theoretic model we have specified this variable to impact R&D cooperation agreements in both directions, the result is that the mean impact of outgoing spillovers is positive but is not significant. This effect may be caused by both, Easy Imitation that influences the dependant variable in both directions, and other unobservable effects that are being captured.

¹The correlation between Incoming Spillovers and Industry Level of Incoming Spillovers is 0.15

²The correlation between Incoming Spillovers and Basicness of R&D is 0.58

Lack of Information impacts negatively and significantly the probability of having R&D cooperation agreements for the Chilean firms. This variable takes a higher absolute value when the dependant variable is cooperation with universities and cooperation with consultors. Lacking information of what is being done by others, may be the cause of not cooperating.

Cost-Risk variable seems to affect cooperation agreements in a positive way. This result is significant in all cases with the exception of the regression where the dependant variable is cooperation with all kinds of partner. Cost-Risk shows a more intensive impact over cooperation with competitors and cooperation with consultors.

Legal Protection affects positively and significantly the probability of engaging in R&D agreements. This is consistent with our theoretic model previously specified. We have found that, for Chilean firms, the greater impact of this variable is on the cooperation with universities.

This model also controls by Industry Level of Cooperation, Industry Level of Legal Protection, R&D Intensity during the period, Size and Size Squared. All controls are always significant except by R&D Intensity which is significant for cooperation with all partners and for cooperating with universities.

6 Concluding Remarks

This paper contributes to the literature by studying the R&D cooperation determinants with data obtained from the fourth Chilean Innovation Survey. We have used Instrumental Variables and Conditional Maximum Likelihood Methods in order to estimate properly our model due to the existence of endogeneity problems in our estimations. Moreover, following Rivers and Young (1988), we have tested for the presence of endogeneity in several of the variables that our model includes as regressors.

Following previous literature on R&D cooperation determinants, Cassiman and Veugelers (2002) and also Colombo and Gerrone (1996) and Lopez (2004), we have tested for the presence of endogeneity on Incoming and Outgoing information flows variables and also on R&D Intensity and Cost-Risk variables. Within our two-step test results we only find evidence to suggest that Incoming Spillovers must be treated as an endogenous variable.

Results show that Incoming Spillovers determine a greater probability of R&D cooperation. This result is seen on all regressions independently of which kind of partner is being considered.

Easy Imitation has a positive signed parameter in all regressions, nevertheless it is clear that this value is persistently not significant. Since in the theoretic model we have specified that this variable impacts R&D cooperation agreements in both directions, the expected result is that the mean impact of outgoing spillovers may be positive or negative but it is not significant.

Lack of Information impacts negatively and significantly the probability of having R&D cooperation agreements for the Chilean firms.

Cost-Risk variable seems to affect cooperation agreements in a positive way, this result is significant in most of our regressions. Legal Protection also affects positively and significantly the probability of engaging in R&D agreements.

In this paper the results of evaluating R&D cooperation determinants in Chilean firms are exposed. Our results found differences and similarities with previous studies that have been done by Cassiman and Veugelers (2002) with data from Belgium and Lopez (2004) with data from Spain.

Among the main similarities of our results with previous studies we find that Incoming Spillovers affect positively and significantly the R&D cooperation agreements of firms.

The main difference that we found is that Outgoing Information flows or Easy Imitation variable is not at all significant in our results, even though it is significant in previous works. According to the theoretic model we used for this work, this variable could be affecting R&D agreements in both directions and our results are consistent with this model.

Variable Definitions

- Easy Imitation: Variable which takes the value 0 if possible imitation is not an important restraint for innovating, and takes values between 1 (low) and 4 (high) when imitation is declared to be an obstacle for innovating. Rescaled between 1 (high) and 0 (not relevant).
- Incoming Spillovers: Variable which takes the value 0 if innovation ideas are not originated by Professional conferences, exhibitions, meetings and journals, and takes values between 1 (low) and 4 (high) when this sources are declared to be sources for innovating ideas. Rescaled between 1 (high) and 0 (not relevant).
- Basicness of R&D: Variable which takes the value 0 if innovation ideas are not originated by institutional sources, and takes values between 1 (low) and 4 (high) when institutional sources are declared to be sources for innovating ideas. Rescaled between 1 (high) and 0 (not relevant).
- Lack of Information: Variable which take the value 0 if lack of information is not an obstacle, as declared by the firm, for the innovative process, and takes vales between 1 (low importance) and 4 (high importance) when the lack of information is an obstacle for innovating. Rescaled between 0 (not relevant) and 1 (high).
- Legal Protection: Variable that takes value 0 if the firm has no patents or know-how agreements and 1 in other case.
- Cooperation: Variable which takes value 0 if the firm does not declare to have cooperated for innovative activities with any other agent (suppliers, clients, competitors, consultors, universities, research institutes) and takes the value 1 if the firm has declared to cooperate with at least one of them.
- Cooperation with competitors: Variable which takes the value 1 when the

firm has declared to cooperate for innovating with its competitors, and takes value 0 in other case.

- Cooperation with Consultors: Variable which takes the value 1 when the firm has declared to cooperate for innovating with private consultors or R&D laboratories, and takes value 0 in other case.
- Cooperation with Suppliers or Customers: Variable which takes the value 1 when the firm has declared to cooperate for innovating with its suppliers or customers, and takes value 0 in other case.
- Cooperation with Universities: Variable which takes the value 1 when the firm has declared to cooperate for innovating with universities or public research institutes, and takes value 0 in other case.
- Cost-Risk: Sum of the scores of importance of the following declared obstacles for innovating (values between 0 (not relevant) and 4 (very important obstacle): Difficulty for finding sources of finance; High cost of innovating; High risk perceived; Payback period too long. Rescaled between 0 (not relevant) and 1 (high).
- Industry Level Imitation: Mean of Imitation at each industry. The industry has been defined with a 2-digit classification.
- Industry Level of Incoming Spillovers: Mean of Incoming Spillovers at each industry. The industry has been defined with a 2-digit classification.
- Industry Level of Legal Protection: Mean of Legal Protection at each industry. The industry has been defined with a 2-digit classification.
- Industry Level of Cooperation: Mean of Cooperation at each industry. The industry has been defined with a 2-digit classification.

- Industry Level of Cooperation with Competitors: Mean of Cooperation with Competitors at each industry. The industry has been defined with a 2-digit classification.
- Industry Level of Cooperation with Consultants: Mean of Cooperation with Consultants at each industry. The industry has been defined with a 2-digit classification.
- Industry Level of Cooperation with Suppliers or Customers: Mean of Cooperation with Suppliers or Customers at each industry. The industry has been defined with a 2-digit classification.
- Industry Level of Cooperation with Universities: Mean of Cooperation with Universities at each industry. The industry has been defined with a 2-digit classification.
- R&D Intensity 2004: Ratio between intramural R&D expenditures and turnover.
- Size: Log of number of employees of the firm.
- Size squared: square of log of number of employees.

References

- [1] Atallah, G. (2005). "Partner Selection in R&D Cooperation." CIRANO Scientific Series.
- [2] Cassiman, B. and Veugelers, R. (2002). "R&D Cooperation and Spillovers: Some Empirical Evidence from Belgium." *American Economic Review*, 92(4), pp. 1169-1184.
- [3] Cohen, W.M. and Levinthal, D.A., (1989). "Innovation and Learning: The Two Faces of R&D." *Economic Journal*, 99(397).
- [4] Colombo, M. and Gerrone, P. (1996). "Technological Cooperative Agreements and Firm's R&D intensity. A note on causality relations." *Research Policy*, September 1996, 25(6), pp. 923-32.
- [5] Geroski, P. A. (1995). "Do Spillovers Undermine the Incentives to Innovate" in Steven Dowrick, ed., *Economic approaches to innovation*. Aldershot, U.K.: Elgar, 1995, pp. 76-97.
- [6] Griliches, Z. (1992) "The Search for R&D Spillovers". *Scandinavian Journal of Economics*, Supp. 1992, 94, pp. 29-47.
- [7] Heckman, J. and MaCurdy, T. (1985). "A Simultaneous Equations Linear Probability Model." *Canadian Journal of Economics*, February 1985, 18(1), pp. 28-37.
- [8] López, A. (2004) "Determinants for R&D Cooperation: Evidence From Spanish Manufacturing Firms". Universidad Carlos III de Madrid, part of the research project: "Innovation and Employment in European Firms: Microeconomic Evidence"
- [9] Mark, M. and Graversen, E. "Determinants of Danish Firms' choice of R&D-cooperation partners." The Danish Centre for Studies in Research and Research Policy, University of Aarhus, Working Paper 2004/6

- [10] Rivers, D. and Vuong, Q. H., (1988). "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models
- [11] Schmidt, T. (2005). "Knowledge Flows and R&D Cooperation: Firm-level Evidence from Germany." Zentrum für Europäische Wirtschaftsforschung (ZEW), Discussion Paper 05-22.
- [12] Vencatachellum, D. and Versaevel, B. (2006). "Horizontal R&D Cooperation and Spillovers: Evidence from France." Centre National de la Recherche Scientifique, Working Paper 06-12.
- [13] Veugelers, R.(1997). "Internal R&D Expenditures and External Technology Sourcing." *Research Policy*, 26(3), pp. 303-316.
- [14] Wooldridge, J. M., (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Table 1: Sample Statistics

All Sample Firms	3,122
Innovating Firms	1,494
Percentage of All Firms	47.85%
Manufacturing Firms	1,269
Innovating Manufacturing Firms	640
Innovating Manufacturing Firms	50.43 %
Energy Firms	150
Innovating Energy Firms	72
Innovating Energy Firms	48%
Mining Firms	64
Innovating Mining Firms	37
Innovating Mining Firms	57.81%
Services Firms	1,412
Innovating Services Firms	613
Innovating Services Firms	43.41%
Non-cooperating Firms	1,309
Percentage of Innovating Firms	87.62%
Cooperating Firms	185
Percentage of Innovating Firms	12.38%
Firms Cooperating with Universities	106
Percentage of Innovating Firms	7.1%
Firms Cooperating with Competitors	93
Percentage of Innovating Firms	6.22%
Firms Cooperating with Consultors	93
(percentage of innovating firms)	6.22%
Firms Cooperating with Customers	166
(percentage of innovating firms)	11.11%

Table 2: Descriptive Statistics

Sample	Mean non-cooperating Firms	Mean Cooperating Firms	Mean Cooperation with Universities	Mean Cooperation with Competitors	Mean Cooperation with Consultants	Mean Cooperation with Customers
	(N=1,309)	(N=185)	(N=106)	(N=93)	(N=93)	(N=166)
Incoming Spillovers	0.653	0.686	0.707	0.741	0.752**	0.698
Easy Imitation	0.725	0.726	0.783	0.806	0.795	0.771
Lack of Info	0.289	0.183**	0.160*	0.182**	0.150**	0.174**
Cost-Risk	0.224	0.2	0.235	0.236	0.225	0.198
Legal Protection	0.189	0.340**	0.405**	0.333**	0.408**	0.349**
Employees	461.5	3704.2*	507.8*	627.9*	7028.2*	4099.3*
R&D Intensity	0.012	0.021**	0.032**	0.020	0.020	0.158

** difference in means between cooperating and non-cooperating firms significant at 5%.

Table 3: Results of Probit Preliminary Regressions

	(a)	(b)
Single Equation Probit Estimates	Cooperation with any organization	Cooperation with Universities
Incoming Spillovers	0.238** (0.094)	0.258** (0.117)
Easy Imitation	0.679 (0.110)	0.152 (0.137)
Lack of Info	-0.308*** (0.114)	-0.406*** (0.148)
Cost-Risk	0.175 (0.118)	0.294** (0.137)
Legal Protection	0.409*** (0.105)	0.518*** (0.123)
Industry Cooperation	7.875*** (1.472)	
Industry Cooperation Univ.		11.569*** (2.210)
Ind. Legal Protection	-1.409** (0.714)	-1.479* (0.823)
R&D Intensity	1.267* (0.728)	2.175*** (0.739)
Size	-0.241** (0.103)	-0.158 (0.116)
Size squared	0.027*** (0.010)	0.017 (0.011)
Constant	-1.396*** (0.310)	-1.909*** (0.323)
LL	-489.014	-314.526
χ^2	85.35	90.07
N	1463	1463

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

Table 4: Results of Probit Preliminary Regressions

Single Equation Probit Estimates	(c) Cooperation with Competitors	(d) Cooperation with Consultors	(e) Cooperation with Suppliers or Customers
Incoming Spillovers	0.376*** (0.124)	0.408*** (0.127)	0.321*** (0.105)
Easy Imitation	0.175 (0.144)	0.219* (0.148)	0.044 (0.118)
Lack of Info	-0.277* (0.147)	-0.452*** (0.158)	-0.292** (0.126)
Cost-Risk	0.351** (0.141)	0.328** (0.143)	0.212* (0.127)
Legal Protection	0.335** (0.132)	0.452*** (0.129)	0.457*** (0.112)
Industry Cooperation Comp	13.530*** (2.386)		
Industry Cooperation R&D		13.197*** (2.395)	
Industry Cooperation Supp			10.494*** (2.141)
Ind. Legal Protection	-1.904** (0.925)	-1.501* (0.816)	-1.269* (0.731)
R&D Intensity	1.114 (0.944)	1.171 (0.908)	1.225 (0.771)
Size	-0.238** (0.115)	-0.316*** (0.115)	-0.213** (0.105)
Size squared	0.028** (0.011)	0.037*** (0.011)	0.023** (0.010)
Constant	-1.951*** (0.329)	-1.889*** (0.330)	-1.678*** (0.297)
LL	-290.654	-288.947	-398.258
χ^2	72.96	87.42	71.61
N	1463	1463	1463

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

Table 5: Testing for Endogeneity, Results of 2SCML

Two Stages	(a)	(b)	(c)
Conditional ML	Cooperation	Cooperation	Cooperation
Incoming Spillovers	0.163** (0.098)	0.163* (0.098)	0.168* (0.098)
Easy Imitation	0.063 (0.152)	0.059 (0.117)	0.106 (0.146)
Lack of Info	-0.447** (0.176)	-0.444*** (0.157)	-0.399*** (0.169)
Cost-Risk	-0.023 (0.194)	-0.023 (0.194)	0.059 (0.177)
Legal Protection	0.304** (0.122)	0.305** (0.120)	0.266** (0.177)
Industry Cooperation	7.623*** (1.579)	7.627*** (1.576)	6.916 (1.409)
R&D Intensity	-4.362 (5.277)	-4.370 (5.276)	1.118 (0.732)
Size	-0.293*** (0.113)	-0.293*** (0.112)	-0.243** (0.102)
Size squared	0.030*** (0.010)	0.030*** (0.010)	0.027*** (0.010)
$\theta_{IncomingSpillovers}$	0.410** (0.192)	0.410** (0.192)	0.402** (0.192)
$\theta_{Appropriability}$	-0.995 (0.807)	-0.997 (0.806)	-0.590 (0.711)
$\theta_{R\&DIntensity}$	5.666 (5.403)	5.673 (5.402)	
$\theta_{Cost-Risk}$	0.026 (0.625)		0.042 (0.623)
Constant	-0.571 (0.846)	-0.563 (0.825)	-1.125* (0.668)
LL	-488.252	-488.253	-488.831
χ^2	86.87	86.87	85.71
N	1463	1463	1463

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

Table 6: Results of CML with Instrumental Variables

	(a)	(b)
Instrumental Variables	Cooperation with	Cooperation with
Probit Estimates	any organization	Universities
Incoming Spillovers	0.602** (0.273)	1.000*** (0.260)
Easy Imitation	0.078 (0.107)	0.164 (0.128)
Lack of Info	-0.302*** (0.114)	-0.366*** (0.142)
Cost-Risk	0.166 (0.116)	0.261** (0.131)
Legal Protection	0.306*** (1.340)	0.338*** (0.123)
Industry Cooperation	6.448*** (1.340)	
Industry Cooperation Univ.		8.638*** (1.964)
R&D Intensity	1.111 (0.720)	1.799** (0.722)
Size	-0.248** (0.100)	-0.183* (0.110)
Size squared	0.027*** (0.010)	0.017 (0.011)
Constant	-1.590*** (0.279)	-2.152*** (0.301)
LL	-1452.901	-1275.436
χ^2	81.51	108.30
N	1463	1463

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

Table 7: Results of CML with Instrumental Variables

Instrumental Variables	(c) Cooperation with Competitors	(d) Cooperation with Consultors	(e) Cooperation with Suppliers or Customers
Incoming Spillovers	0.813** (0.318)	1.056*** (0.282)	0.735*** (0.285)
Easy Imitation	0.176 (0.140)	0.230* (0.141)	0.055 (0.115)
Lack of Info	-0.272* (0.145)	-0.417*** (0.154)	-0.286** (0.125)
Cost-Risk	0.328** (0.139)	0.299** (0.138)	0.201 (0.125)
Legal Protection	0.197 (0.133)	0.292** (0.128)	0.338*** (0.115)
Industry Cooperation Comp	11.315*** (2.240)		
Industry Cooperation R&D		10.753*** (2.276)	
Industry Cooperation Supp			8.656*** (2.021)
R&D Intensity	1.061 (0.897)	0.802 (0.884)	1.042 (0.766)
Size	-0.250** (0.112)	-0.332*** (0.111)	-0.227** (0.104)
Size squared	0.028*** (0.011)	0.037*** (0.011)	0.024** (0.010)
Constant	-2.218*** (0.317)	-2.158*** (0.310)	-1.886*** (0.293)
LL	-1254.715	-1251.228	-1361.570
χ^2	71.50	95.85	70.80
N	1463	1463	1463

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