

# Multilevel heterogeneity of R&D cooperation and innovation determinants

Sara Amoroso<sup>1</sup>

Received: 3 August 2015 / Revised: 18 December 2015 / Accepted: 29 December 2015 /  
Published online: 10 February 2016  
© European Union 2016

**Abstract** Assessing the impact of public support to innovation on R&D collaboration may require a more complex multilevel design, that describes the likely correlation present among firms characteristics within a particular sector. Using data from the 2006 edition of the Community Innovation Survey (CIS) for the Netherlands, we propose a methodology to study the effect of firm-level characteristics on the propensity to undertake a research collaborative agreement. In particular, we show that controlling for a richer variance structure yields a different picture with respect to simpler regression frameworks adopted in the literature of R&D cooperation determinants. Moreover, such a hierarchical framework can be generalized allowing for clustering at higher levels, such as sectors or geographical areas. Besides the link between public funding and R&D collaboration, our results confirm the findings of the literature: technological spillovers, risk and cost sharing rationales, firm's size, and type of innovative activity are related to the decision of engaging in different sorts of research alliances.

**Keywords** R&D collaboration · Innovation · R&D policy · Bayesian analysis

**JEL Classification** O30 · O32 · C11

---

The ideas proposed and the views expressed by the authors may not in any circumstances be regarded as stating an official position of the European Commission. The results and any possible errors are entirely the responsibility of the authors.

---

✉ Sara Amoroso  
sara.amoroso@ec.europa.eu

<sup>1</sup> Joint Research Centre, European Commission, Calle Inca Garcilaso, 3, 41092 Seville, Spain

## 1 Introduction

Given the increasing market pressures, a firm that wants to survive must not only be innovative, but also ready to face shorter business cycles,<sup>1</sup> and prompt to meet a more dynamic demand. Phenomena such as knowledge outsourcing and networking are at the core of entrepreneurial actions. In particular, firms decide to collaborate in research for various economic reasons. In fact, a research alliance could aim at strengthening the member firms' core competencies, so as to reach for new markets, or it may be a strategic decision to access complementary knowledge in order to compensate for the absence of internal competencies or to reduce the costs associated with knowledge spillovers. The theoretical literature on cooperative R&D points at the internalization of the technological spillovers as the main rationale behind the decision to cooperate (Katz 1986; d'Aspremont and Jacquemin 1988; Kamien et al. 1992). Parallel to the theory, the empirical literature confirms the relevance of such spillovers in influencing the choice of cooperating, and extends the analysis on R&D cooperation determinants by distinguishing between incoming and outgoing spillovers (Cassiman and Veugelers 2002, 2006; Lopez 2008; Kultti et al. 2015). Moreover, several authors focus on the analysis of the heterogeneity in the determinants of innovating firms' decisions to engage in R&D cooperation (Kaiser 2002; Belderbos et al. 2004a, b). These studies explore the differences in the factors affecting the firms' probability to establish different types of cooperation, namely horizontal cooperation (with competitors), vertical cooperation (with suppliers or customers), and institutional cooperation (with universities and research institutes). Furthermore, Belderbos et al. (2004a) relax the assumption on the independence among different cooperation strategies, accounting for possible correlations between the strategies that could be due to technological complementarities. However, the existing literature on the determinants of R&D cooperation with different R&D partners overlooks the role of sectoral specificities (for example, sector-specific physical assets) in influencing the expected correlation among the different cooperative strategies present among firms within a particular sector.<sup>2</sup> As a matter of fact, horizontal R&D cooperations are likely to be formed within the same sector as it will lead to *collective efficiencies* (Schmitz 1999) in the form of reduced transaction costs and accelerated innovation rates through a greater market access. These collective efficiencies are of particular interest from a policy perspective, as they represent the multiplier effect of an innovative input: the increased innovative intensity of one company or several companies multiplies the economic benefits in a given sector by helping to drive the innovative intensity of other business entities. This type of externality is demand-driven, in the sense that the private and public for

<sup>1</sup> Recent research employing spectral analysis has confirmed the presence of sinusoidal-like cycles (called Kondratiev) in the world GDP dynamics at an acceptable level of statistical significance. Korotayev and Tsirel (2010) detected shorter (on average 17 years) business cycles, approximately one third of the Kondratiev cycles.

<sup>2</sup> Depending on the model assumptions, and compatibly with the data at hand, one could allow for a richer specification of the clusters, such as the geographical district, or the relevant markets. We limit ourselves to a frugal, yet general representation of a multilevel design in the context of research cooperation determinants.

innovation can be stimulated by new or growing business enterprises, which enables their suppliers to grow as well.

Such collective interactive processes may derive from organizational proximities, and when organizational proximity arises between organizations connected by a relationship of either economic or financial dependence/interdependence, Kirat and Lung (1999) state that intra-sectoral links are liable to dominate inter-sectoral links.

Therefore, in this paper, we control for a richer variance structure that describes the expected correlation present among firms' cooperative choices within a particular sector. In particular, we analyze the correlation among different R&D alliances due to both firm- and sector-level heterogeneity, adopting a multivariate hierarchical logit model. The advantages of adopting a hierarchical structure (often referred to as multilevel, random or mixed, see Goldstein 1995; Hedeker and Gibbons 1996) are several. First, it allows us to assume and specify a more complex covariance structure. This means that we can fit a regression model to firm-level data, while accounting for unexplained variation among the sectors, aiming at capturing relevant features to explain the propensities to undertake a specific cooperative agreement. Second, there is no need to have a balanced design or equally spaced measurements, as the number of firms per sector is allowed to vary. Third, unlike the multivariate probit, the logit specification is not restricted to the normal distribution assumption for the individual effects, and its statistical fit is more accessible because of the odds-ratio interpretation of the logit coefficients. Its derivation is straightforward, and simulation of its choice probabilities is computationally simple.

Our approach departs from the one used to test for complementarities (Mohnen and Röller 2005; Belderbos et al. 2006), as our main focus is not to estimate the degree of strategic complementarities or substitutabilities among firms' cooperation choices, but rather to model and estimate both individual and aggregate forms of externalities, in which the collective actions of a reference group affect an individual's choices. As pointed out by Mohnen and Röller (2005), innovation policies may have different impacts on the distinctive phases of the innovation process. As a matter of fact, while there could be firm-level policy externalities in the decision to collaborate in research, the innovative produce might well be affected by the aforementioned demand driven innovation policy externality.

Therefore, to explicitly take into account both firm- and sector-level externalities, and the different impact of innovation policy measures on the two phases of the innovative process, we divide our study in two stage. In the first stage we study the main drivers of undertaking a collaborative agreement with a research partner. In the second stage, we investigate the effects of innovation policies and R&D cooperation on innovative intensity.

Using data from the 2006 edition of the Community Innovation Survey (CIS2006) on 1356 innovating Dutch firms operating in 15 manufacturing sectors, we analyze the firm- and sector-level heterogeneity of the determinants of R&D cooperations and of the production of innovative output determinants. To our knowledge, this paper is the first attempt to use a hierarchical heteroskedastic framework to model the firms' determinants of R&D partner's choices and to assess the impact of public financial support to innovative output.

Our results, additional to the well documented firm-level heterogeneity (Belderbos et al. 2004a), confirm the presence of sector-level heterogeneity which might affect the probabilities to cooperate and the level of innovative production. Moreover, controlling for public financial support for innovation activities at different levels of government, we show that when using the suggested multilevel approach the impact of public funding has a positive and significant sign, while, when omitting this nested framework, the policies have a poor effect on innovative turnover.

The remainder of this paper is organized as follows. In Sect. 2, we look at the existing theoretical and empirical literature on the R&D cooperation determinants to guide our own empirical analysis. Section 3 describes the model. In Sect. 4, we describe the data. Sections 5 and 6 discuss the empirical results of the mixed models to describe both the propensity to undertake the different cooperation agreements, and the impact of the public funding on the firm-level innovative output. Section 7 summarizes.

## 2 Determinants of R&D cooperation: a review of firm-level and industry-level factors

The Industrial Organization (IO) literature has pointed to technological spillovers as one of the important factors influencing the firms' incentives to engage in cooperative R&D (Katz 1986; d'Aspremont and Jacquemin 1988; Kamien et al. 1992). R&D spillovers arise when new knowledge created by one firm is also beneficial to other firms. Theoretical studies suggest that a high level of R&D spillovers can increase the firms' probability of internalizing the spillovers by participating in R&D cooperation.

The relationship between spillovers and R&D cooperation has also been analyzed using empirical studies. Cassiman and Veugelers (2002) empirically explore the effects of knowledge flows on R&D cooperation on Belgian firms' decisions to enter into a cooperative R&D agreement, highlighting the distinction between the effect of knowledge flows into firms (*incoming spillovers*) and that of outbound knowledge flows (*appropriability*). Cassiman and Veugelers (2002) observe a significant relation between external information flows and the decision to cooperate in R&D. Indeed, the most important finding of their two-step probit model is that the probability of firms cooperating in R&D is higher when incoming spillovers are high and outgoing spillovers are low. Furthermore, cost-sharing is found to be an important motive for cooperation in R&D, while risk-sharing is not. The empirical model of Cassiman and Veugelers (2002) has been to some extent modified and applied to Spanish firms by Lopez (2008), who focuses on the role of cost and risk sharing of innovation projects for the cooperation decision, and finds evidence supporting the importance of cost/risk sharing as a hampering factor for the innovation process. Further, the paper of Lopez (2008) pays much attention to the endogeneity of the independent variables which in other papers are assumed to be endogenous a priori. The results confirm the conclusions of Cassiman and Veugelers (2002) that spillovers and appropriability play an important part in influencing R&D

cooperation decisions, only when an adequate structure of endogeneity is chosen. Indeed, the hypothesis of exogeneity of incoming spillovers and appropriability is rejected, while the exogeneity of R&D intensity cannot be rejected.

In a study based on data of European firms, Hernán et al. (2003) follow a two-step procedure. In the first stage, the entire population that could potentially participate in a cooperative organization is considered. In this first stage, it is possible to measure the effect of the relevant firm characteristics that influence a research joint venture (RJV) formation. In the second stage, the focus is on firms that are known to have a higher probability to participate in cooperative R&D projects. Using a large database of firms from almost twenty European countries, Hernán et al. (2003) find that, contrary to what has been found by Cassiman and Veugelers (2002) and more in line with the aforementioned IO theory, patents' effectiveness and, therefore, the level of appropriability, reduces R&D cooperations. Moreover, among individual firm characteristics, firm size, and previous participation experience increase the likelihood of participating. Industry-level characteristics are also significant, especially R&D intensity. RJVs are also more likely in more concentrated industries where technological knowledge diffuses rapidly. Therefore, a minimum level of industry concentration is needed for RJVs to be formed. One possible explanation for the significance of firm size is that EU programs favor large partnerships, which may be more costly to manage. An important characteristic of their analysis is that they use a large control group that is representative for the whole population of European firms. With respect to differences between countries, Hernán et al. (2003) find that mainly firms in smaller countries participate in projects funded by the EU; according to the authors this is because firms in large countries can find partners in their own country more easily. Spillovers are measured at the industry level and are proxied by the average number of months before the diffusion of an innovation in the industry and the effectiveness of patents in the industry, both based on previous analyses. Problems of endogeneity are dealt with by lagging all time-dependent right-hand-side variables by two years.

Based on a sample of Italian manufacturing firms, Piga and Vivarelli (2004) explore the determinants of collaborative R&D project external to the firm, accounting for the selection bias that derives from the antecedent choice of engaging in internal R&D activity. Adopting a censored bivariate probit, they show the relevance of correcting for such selectivity bias. Moreover, exploiting the rich information of the dataset, they find that (partially) state-owned firms are more likely to cooperate with universities and research centres. External collaborative projects with private firms are, on the other hand, favoured by firms having objectives oriented towards process and product innovation.

Additional results dealing with the relationship between R&D cooperation, spillovers, and productivity appear in Belderbos et al. (2004a), who construct a multivariate cross-sectional probit model to explore differences in the determinants of innovating firms' decisions to participate in four distinct types of partner specific innovation strategies (cooperation with competitors, suppliers, customers, and universities and research institutes). With a large Dutch dataset (627 firms with R&D cooperation of some type), evidence of a positive impact of R&D cooperation on labor productivity growth is found, but with distinct differences depending on

(the combination of) cooperation types. Competitor and supplier cooperations seem to have the most positively significant impact on productivity growth. The results for the other variables show that incoming spillovers and R&D intensity are statistically significant in explaining R&D cooperation with firms from the same industry.

Busom and Fernández-Ribas (2008) find that the formation of inter-firm alliances is likely to vary among firms and can be due to a large number of reasons, such as the nature and the scope of the R&D projects. In particular, the authors show that if a firm's aim in a cooperative agreement is to share complementary technology, it will tend to cooperate with heterogeneous partners (heterogeneous in knowledge assets, market scope, location, and product range), while, when the motivation for cooperation is based on internalizing outgoing spillovers or on increasing market power, symmetric partnerships (i.e., horizontal or vertical) are more likely.

A less developed strand of literature deals with the *innovativeness* of firms and their propensity to cooperate. As a matter of fact, some R&D alliances may differ in the degree of technological effort required. Indeed, we can identify two types of innovations that R&D partnerships may develop: radical and incremental.<sup>3</sup> A radical innovation is a product, process, or service offering “significant improvements in performance or cost that transform existing markets or create new ones” (Leifer et al. 2001). On the other hand, incremental innovations are based on minor changes or improvements in the current technology. Nonetheless, similar concepts are implicitly used in the economic literature. IO theorists, as for example Reinganum (1983), use terms such as *drastic innovations* to describe those changes in technology that determine a decrease in costs such that the new equilibrium price lies below the pre-innovation cost and consequently turn the innovator into a monopolist. On the other hand, non-drastic or *gradual* innovations only introduce costs asymmetries that do not transform the market into a monopoly. Tether (2002) observed that true or radical innovators cooperate more than those who introduce only imitative innovations.

None of these studies, however, controls for the likely multilevel structure of the data. As a matter of fact, data may occur in clusters, such as sectors in which firms are nested. One approach to modeling such type of data includes random effects for subjects (firms) or clusters (sectors) into account. This provides a mechanism of accounting for certain correlation structures among the clustered observations.

### 3 Model specification

To investigate the relationship between the factors driving the propensity to collaborate with different research partners, we assume a hierarchical structure of the model specification. In particular, we adopt a *multivariate mixed logit model*<sup>4</sup>

<sup>3</sup> The labels *radical* and *incremental* belong mostly to the managerial literature (see Dewar and Dutton 1986; Henderson 1993).

<sup>4</sup> The class of mixed logit models is a highly flexible as it can approximate any random utility model (Train 2009). The results we present can be generalized and extended to panel data.

(multi-response Generalized Linear Mixed Model—GLMM, Hedeker and Gibbons 1996). Using the terminology of multilevel analysis, let  $i$  denote the level-1 units (nested observations, i.e., firms) and let  $j$  denote the level-2 units (subjects, i.e., sectors). Assume there are  $j = 1, \dots, J$  sectors and  $i = 1, \dots, N$  firms. The total number of firms is given by  $N = \sum_{j=1}^J n_j$ , where  $n_j$  is the number of firms nested within each sector. Each firm  $i$  is faced with  $c = 1, \dots, C$  different choices of cooperation strategies.

Let us define the utility that firm  $i$  in sector  $j$  obtains from choosing cooperation  $c$  as

$$\begin{aligned} U_{ij}^c &= V_{ij}^c + \epsilon_{ij}^c \quad \text{where} \\ V_{ij}^c &= \beta_0^c + x'_{ij}\beta^c + z'_{ij}\alpha_j^c + \gamma_i^c. \end{aligned} \quad (1)$$

$\beta_0^c$  and  $\beta^c$  are the intercept and the vector of category-specific fixed effects, respectively,  $x_{ij}$  and  $z_{ij}$  are vectors of observed variables. The error term  $\epsilon_{ij}^c$  is assumed to be independently, identically extreme value distributed. We then assume that the vector of firm- and sector level random intercepts and slopes,  $\gamma_i^c$  and  $\alpha_j^c$  are normally distributed and have a firm- and sector-level cooperation strategy-specific covariance matrices to allow for heteroskedasticity deriving from different cooperation alternatives. A detailed description of the covariance structure is given in Appendix 1.1.

The regression parameters are collected in the vector  $\theta = (\beta_0, \beta^{c'}, \alpha', \gamma')$ . Given the extreme value distribution assumption of the error term  $\epsilon_{ij}^c$ , the probability for firm  $i$  in sector  $j$  of choosing to cooperate with partner  $c$  follows a logistic distribution. To retrieve firm- and sector-level coefficients assuming a heteroskedastic nested cooperation choice model, we follow a Bayesian approach (see Appendix 1.2).

## 4 Data

The data used for the present study corresponds to the 2006 edition of the Community Innovation Survey<sup>5</sup> (also referred to as CIS2006), carried out by Statistics Netherlands. The Dutch CIS2006 collected data on product and process innovation, as well as organisational and marketing innovation during the three-year period 2004 to 2006. The total number of firms participating to the survey was 9940, 26.7 % (2660 firms) of which operated in the manufacturing industry. Of 2660 firms 1356 (more than 50 %) introduced an innovation—product or process—694 (26 %) participated in a cooperative activity, and all the participants to a cooperative agreement were innovators.

<sup>5</sup> The Community Innovation Surveys are designed to provide an extensive description of the general structure of innovative activities at the country and industry levels. Within the guidelines of the OSLO Manual on performing innovation surveys (OECD 1997), information about innovation activities is collected.



Table 1 reports the numbers of firms across the different technology classes, namely high-tech, medium/high-tech, medium/low-tech, low-tech,<sup>6</sup> and the percentages of innovators, product and process innovators, the percentage of firms that invest in R&D and undertake a cooperative agreement. The table shows that the high- and medium/high-tech industries exhibit a higher concentration of innovating firms, both in product and process innovations. Moreover, firms in high- and medium/high-tech industries invest more in R&D and cooperate more than those firms operating in medium/low- or low-tech sectors. Additionally, the number of firms across the different technological classes is unevenly distributed, as the majority of the companies (39.85 % of the total sample) serve the low-tech sector, while only the 5.5 % is in the high-tech industries.

Table 2 reports the averages per industry class of firms' R&D expenditure, the ratio between extramural and intramural R&D expenditure, percentages of net sales from radical and incremental innovations, net sales per employee in 2006, and the growth rate of net sales. The high-tech sectors spend in research almost six times more than the average R&D spender and have the lowest extramural/intramural R&D ratio, meaning that they do not outsource as much as, for example, low-tech industries (ratio equal to 77 %), however they do cooperate twice as much. The sectors with the highest turnover per employee are the medium/high-tech ones, while the medium/low-tech sector experience the highest net sales per employee growth rate (18.53 %). The high-tech sectors have on average a net sales of 100 million of euros (more than twice the average innovative firm), whereas the average number of employees in high-tech is nearly twice the number of employees in the average innovative company. As a matter of fact, among the high-tech sectors, the manufacturing sector of radio, television and communication equipment (an ICT sector) is by nature labor-intensive and accounts for more than 20 % of firms in high-tech sectors.

In terms of innovative output, high- and medium/high-tech industry report higher innovative sales in both drastic (radical) and gradual (incremental) innovations than the low- and medium-low counterparts.

#### 4.1 Dependent variables

The dependent variables of the model are dummy variables equal to one if the firm, during the three years 2004 to 2006, actively participated with other enterprises or non-commercial institutions on innovation activities. In particular, as in Belderbos

<sup>6</sup> Following the guidelines of the OECD Directorate for Science, Technology and Industry, the manufacturing industry can be classified into four categories according to technology intensity using the ISIC Rev. 3 breakdown of activity: high technology sectors (aircraft and spacecraft; pharmaceuticals; office, accounting and computing machinery; radio, TV and communications equipment; medical, precision and optical instruments), medium/high-technology industries (electrical machinery and apparatus; motor vehicles, trailers and semi-trailers; chemicals excluding pharmaceuticals; railroad equipment and transport equipment; machinery and equipment), medium/low-technology industries (building and repairing of ships and boats; rubber and plastics products; coke, refined petroleum products and nuclear fuel; non-metallic mineral products; basic metals and fabricated metal products), and low-technology industries (recycling; wood, pulp, paper, paper products, printing and publishing; food products, beverages and tobacco; textiles, textile products, leather and footwear).



**Table 1** Innovating and cooperating firms across industry classes

Industry	% innovators	% product	% process	% R&D	% coop.	N. firms
Hi tech	64.63	55.10	41.50	46.26	42.86	147 (5.53 %)
Medium/hi tech	63.37	53.92	38.95	37.94	35.61	688 (25.86 %)
Medium/lo tech	49.67	36.21	34.90	21.57	23.40	765 (28.76 %)
Lo tech	41.98	30.85	30.09	16.42	19.53	1060 (39.85 %)
Total	50.98	39.70	34.40	25.11	26.09	2660 (100 %)

**Table 2** Summary statistics

Industry	R&D exp ( $\times 10^6$ €)	R&D <sup>in/ex</sup> ratio	% sales radical	% sales incremental	Sales/ L <sub>2006</sub>	$\Delta$ sales/L (%)
Hi tech	17.956	22.42	7.48	8.81	297.578	9.35
Medium/hi tech	4.967	41.33	5.31	6.01	321.238	16.41
Medium/lo	0.861	31.56	3.61	3.64	302.449	18.53
Lo tech	1.262	74.44	2.31	2.99	284.664	12.00
Average	3.512	47.59	3.74	4.28	299.952	14.88

et al. (2004a), we consider four R&D partnerships, namely with suppliers, clients, competitors, or research institutes and/or universities (institutional cooperation).

Cooperation networks can be further distinguished in two types of cooperative behavior. The first is based on the synergies obtained by combining complementary assets. This combination of resources enables a more complete or intense use of the different types of assets possessed by each firm. One of the most important of these complementary cooperation agreements is the *vertical* or *supply-chain* cooperation, in which the company cooperates with its customers and/or suppliers (Tether 2002).

The second rationale which characterizes the other type of cooperative agreement consists in competitive positioning (Cassiman and Veugelers 2002), i.e., seeking market power. This type of cooperation, also known as *horizontal* R&D cooperation tends to form matches between competing firms that might have similar needs in terms of product or process development, looking for resources of the same type (technological, human, and so on). A summary of the number and percentages of firms adopting the four types of R&D cooperation is presented in Table 3. Next to the number of firms adopting a cooperative agreement per sector class, and the four cooperation types, we report the number of cases and the percentage relative to the total number of firms in the different industry classes for vertical cooperation, and for cooperation with public research institutes combined with competitors. Vertical cooperation is the most frequent type of agreement, as 45.5 % of innovative firms decide to engage in either a collaboration with suppliers (41.7 %), or with customers (23.9 %). The medium- and high-tech sectors are the most collaborative in all cooperative agreement (35.3 %), however, the partners most frequently chosen for cooperating are suppliers. R&D cooperation with other competing enterprises is the

**Table 3** Types of cooperations per industry class

Industry	Non Coop	Coop	Cooperating partner			
			Competitor	Customer	Supplier	Public institute
Hi tech	32	63	22	45	55	39
	4.83 (column %)	9.08 (column %)	23.16 (row %)	47.37	57.89	41.05
Medium/hi tech	191	245	58	163	187	127
	28.85 %	35.30 %	13.30 (row %)	37.38	42.89	29.13
Medium/lo tech	201	179	48	109	146	68
	30.36 %	25.79 %	12.63 (row %)	28.68	38.42	17.89
Lo tech	238	207	55	100	177	90
	35.95 %	29.83 %	5.19 (row %)	9.43	16.70	8.49
Total	662	694	183	417	565	324
	100 %	100 %	13.49	30.75	41.67	23.89

Numbers and row percentages of firms engaging in different types of cooperation agreements: competitors, suppliers, clients, or public research institutions. Vertical cooperation is here defined as the cooperation with suppliers or customers

least frequently observed choice (13.5 % of the innovating firms sample), however high-tech firms are those with the highest relative percentage of horizontal cooperative agreements (22 firms out of 95). Collaboration with universities and public institution of research is chosen by 324 out of 1356 firms.

Table 4 reports the correlation coefficients between the different types of cooperation for the entire sample of firms and for the sectors classified according to technological intensity. Clearly, the highest correlation is found between cooperation with suppliers and cooperation with customers; however, for high-tech sectors, this is no longer the case, as the highest correlation is between cooperative agreements with suppliers and with public research institutes (0.650). The highest correlation between vertical cooperation agreements is found in medium/high-tech industries (0.743).

## 4.2 Rationales for cooperating

Following the existing theoretical and empirical work, we propose four sets of explanatory variables related to firms' characteristics, obstacles to innovation that the firm should overcome, the existence of public funding to encourage R&D, and sectoral characteristics within which the firm operates. Table 5 reports summary statistics of the regressors used to estimate the variances  $\mathbf{V}_1$  and  $\mathbf{V}_2$  and the set of coefficients  $\theta$  of the mixed logit model (see Appendices 1.1 and 1.2, respectively).

**Table 4** Correlation between cooperative agreement

	Institutional	Competitors	Customers	Suppliers
All industries				
Suppliers	0.588	0.454	0.699	1
Clients or customers	0.551	0.467	1	
Competitors	0.466	1		
Institutional	1			
Hi-tech industries				
Suppliers	0.650	0.503	0.645	1
Clients or customers	0.604	0.466	1	
Competitors	0.439	1		
Institutional	1			
Med/hi-tech industries				
Suppliers	0.585	0.391	0.743	1
Clients or customers	0.537	0.434	1	
Competitors	0.408	1		
Institutional	1			
Med/lo-tech industries				
Suppliers	0.538	0.478	0.716	1
Clients or customers	0.451	0.496	1	
Competitors	0.449	1		
Institutional	1			
Lo-tech industries				
Suppliers	0.590	0.465	0.634	1
Clients or customers	0.596	0.463	1	
Competitors	0.539	1		
Institutional	1			

### 4.3 Firm-level characteristics

With regard to firms' characteristics, firm-level knowledge inflows, also defined as incoming spillovers (Cassiman and Veugelers 2002), are derived from the scores of importance of publicly available information. We denote the spillover variable by  $spill_{ij}$ . Unfortunately, the 2006 edition of the CIS does not collect any information on the level of appropriability, namely, the degree of strategic protection the firms adopt for their innovations.

The variable  $spill_{ij}$  measures the degree of importance of publicly available source of information. The variable was originally coded from 0 (not used) to 3 (highly important), and we recoded in  $\{0, 1/3, 2/3, 1\}$ .

Variable	Description
<i>Co<sub>comp</sub></i>	Dummy variable which takes the value 1 when the firm has actively participated with its competitors on innovation activities
<i>Co<sub>cust</sub></i>	Dummy variable which takes the value 1 when the firm has actively participated with its clients or customers on innovation activities
<i>Co<sub>supp</sub></i>	Dummy variable which takes the value 1 when the firm has actively participated with its suppliers of equipment, materials, components, or software on innovation activities
<i>Co<sub>inst</sub></i>	Dummy variable which takes the value 1 when the firm has actively participated with universities or other higher education institutions, or government or public research institutes on innovation activities
<i>spill</i>	Variable which takes the value 0 if innovation ideas are not originated by Professional conferences, exhibitions, meetings and journals, professional and industry associations. The variable was originally coded from 0 (not used) to 3 (highly important), and we recoded in 0, 1/3, 2/3, 1
<i>lp</i>	Dummy variable that takes value 1 if the firm applied for a patent, registered an industrial design, or a trademark, or claimed a copyright
<i>size</i>	Log of number of employees of the firm
<i>cost</i>	Variable measuring the importance of the costs of innovation, or the lack of funds, or access to finance in hampering the firm's innovation activities or projects or influencing the decision not to innovate. The original variable takes values between 1 (high) and 4 (not relevant). Rescaled between 0 (not relevant) and 1 (high)
<i>risk</i>	Variable measuring the importance of the uncertainty of the demand for innovative goods or services in hampering the firm's innovation activities or projects or influencing the decision not to innovate. The original variable takes values between 1 (high) and 4 (not relevant). Rescaled between 0 (not relevant) and 1 (high)
<i>rad</i>	Dummy variable that takes value 1 if the enterprise introduced a new or significantly improved good or service onto its reference market before the competitors
<i>inc</i>	Dummy variable that takes value 1 if the enterprise introduced a new or significantly improved good or service that was already available from the competitors in its reference market
<i>rdpi</i>	The absorptive capacity is proxied by R&D personnel intensity, measured as the ratio between the log of researchers (full time equivalent) and the size of the firm
<i>gp</i>	Dummy variable that takes value 1 if the firm is part of an enterprise group. A group consists of two or more legally defined enterprises under common ownership. Each enterprise in the group may serve different markets, as with national or regional subsidiaries, or serve different product markets. The head office is also part of an enterprise group
<i>funloc</i>	Dummy variable taking value 1 if the firm received any public financial support for innovation activities from local or regional authorities
<i>fungmt</i>	Dummy variable taking value 1 if the firm received any public financial support for innovation activities from central government (including central government agencies or ministries)
<i>funeu</i>	Dummy variable taking value 1 if the firm received any public financial support for innovation activities from the European Union
<i>HHI</i>	As a measure of industry concentration we adopt the Herfindahl-Hirschman Index, computed as the sum of the squared firms' market shares
<i>Y<sub>inno</sub></i>	Innovation output is proxied by the percentage of total turnover from product or process innovation (new to the firm and/or to the market)

**Table 5** Summary statistics

Variable	Mean	SD	Min	Max
$In_{ij}$	0.572	–	0	1
$rad_{ij}$	0.317	–	0	1
$inc_{ij}$	0.336	–	0	1
$spill_{ij}$	0.161	0.255	0.000	1.000
$size_{ij}$	4.234	1.134	0.405	9.942
$risk_{it}$	0.228	0.303	0.000	1.000
$cost_{ij}$	0.259	0.333	0.000	1.000
$funloc_{ij}$	0.069	–	0	1
$fungmt_{ij}$	0.259	–	0	1
$funeu_{ij}$	0.049	–	0	1
$lp_{ij}$	0.270	–	0	1
$HHI_j$	0.078	0.098	0.015	0.378
$gp_{ij}$	0.663	–	0	1
$rdpi_{ij}$	0.026	0.079	0.000	0.957

Size turns out to be another important determinant of R&D cooperation. Cassiman and Veugelers (2006) highlighted how the size of the firms is a control variable, traditionally used by the literature in firm-level analysis. Therefore, in line with the existing literature, we include firm size measured as the logarithm of the number of the firm's employees ( $size_{ij}$ ). On average the log of labor is 4.234. The innovativeness of firms is proxied by two dummy variables. The first dummy variable,  $rad_{ij}$ , proxying radical innovation, takes the value one if the enterprise introduced a new good or service into its market before its competitors, while the second dummy,  $inc_{ij}$ , is equal to 1 if a new good or service was already available from the competitors in the market of interest (incremental innovation). Firms introducing a radical innovations account for 32 % of the sample, while imitative innovators sum up to 34 % of the total number of firms in the sample. Among determinants of R&D cooperation, firms' absorptive capacity is considered as one of the most important. To proxy firms' absorptive capacity R&D personnel intensity ( $rdpi_{ij}$ ) is often used (Tether 2002; Belderbos et al. 2004a, b) instead of R&D expenditure. R&D personnel intensity is defined here as the ratio between R&D personnel and size of firms. On average, only 3 % of the labor force is dedicated to research activities.

The integration of the firm into a group may also have a positive influence on cooperation as it indicates access to a substantial pool of resources, which are complementary to R&D. Thus, we include the intra-group variable,  $gp$ , a control variable which is equal to 1 when the firm belongs to a group, and 66 % of the firms are part of an enterprise group.

#### 4.4 Obstacles to innovation

The variables  $risk_{ij}$ , and  $cost_{ij}$  measure the degree of importance,  $\{0, 1/3, 2/3, 1\}$ , attributed by firms to the two factors hampering innovation activities. The cost factor seems to be larger than the risk associated with market uncertainty (26 versus 23 %). The literature on R&D cooperation shows that the risks and costs of innovation and the need to exploit complementary resources are the main motivations for cooperative behavior, and therefore, that cooperative behavior may be positively related to a number of obstacles such as high risks and cost of innovation. R&D cooperations, in fact, allow firms to share costs or to reduce risks of innovation. Kim (2014) offers additional insights on the sharing rationale of cooperating parties. He distinguishes between formal and informal agreements and explores the extent to which these modes of collaboration governance are substitutes or complements. The way in which alliance governance mechanisms are employed may have implications for how the partners extract economic value from the cooperative relationship. In this regard, we hypothesize that a cooperation with customers could reduce the risk to introduce a radical innovation in the market.

#### 4.5 Public funding

The impact of public funding is quite debated. In fact, while a number of empirical studies have found a positive influence on firms' R&D expenditure, indirectly influencing the propensity to cooperate in R&D (Veugelers 1997; Piga and Vivarelli 2004; Almus and Czarnitzki 2003), there is also evidence of possible adverse effects. Catozzella and Vivarelli (2014) show that, although the public funding has a positive impact on the additional R&D expenditures, it has a negative effect on the innovative productivity. Crespi et al. (2015) investigate the role of environmental policies in the evolution of green technologies and find that, if on the one hand environmental regulation and policies improve the productivity in some specific activities in the short-run, on the other hand they might cause negative indirect effects. Arora and Cohen (2015) find that the impact of government support to product R&D is larger, the smaller is the average firm size within the industry. Hanley et al. (2015) do not find any effect of government financial system intervention on China's regional innovation performance. Therefore, to test the impact of public funding on R&D collaboration, we include dummies taking value 1 if the firm benefitted from both local, national, and European scientific and technological policies ( $funloc_{ij}$ ,  $fungmt_{ij}$ , and  $funeu_{ij}$ , respectively) sponsoring cooperative projects, as they potentially constitute an incentive to cooperate. The number of firms receiving a national funding account for 26 % of the observations, while local and European financial support is less frequent, as only 7 and 5 percent of firms received public sector support, respectively.

#### 4.6 Sector-level characteristics

As for sectoral characteristics, we include the degree of industry concentration, measured by the Herfindahl-Hirschman index ( $HHI_{ij}$ ), as this may affect firms'

**Table 6** Estimation results: multivariate mixed logit

	Suppliers	Customers	Competitors	Institutional
<i>Intercept</i>	-3.541*** (-4.448, -2.568)	-4.691*** (-5.721, -3.637)	-5.718*** (-7.015, -4.441)	-6.085*** (-7.365, -5.045)
<i>rad</i>	0.593*** (0.307, 0.898)	0.926*** (0.621, 1.220)	0.618*** (0.221, 1.054)	0.593*** (0.255, 0.899)
<i>inc</i>	0.280** (0.014, 0.577)	0.486*** (0.172, 0.809)	0.596** (0.196, 1.043)	0.410*** (0.103, 0.756)
<i>size</i>	0.339*** (0.222, 0.461)	0.290*** (0.161, 0.428)	0.225*** (0.068, 0.379)	0.343*** (0.212, 0.489)
<i>spill</i>	0.416 (-0.145, 0.882)	0.539 (-0.053, 1.193)	1.298*** (0.321, 2.214)	1.799*** (1.021, 2.644)
<i>cost</i>	0.197 (-0.118, 0.521)	0.342* (-0.036, 0.664)	-0.145 (-0.594, 0.267)	0.630*** (0.243, 1.002)
<i>risk</i>	-0.098 (-0.397, 0.226)	0.581*** (0.243, 1.001)	0.196 (-0.240, 0.704)	-0.072 (-0.393, 0.289)
<i>rdpi</i>	2.184*** (0.908, 3.562)	0.713 (-0.237, 1.648)	0.646 (-0.540, 1.660)	2.165** (1.032, 3.464)
<i>fungnt</i>	0.532*** (0.254, 0.857)	0.565*** (0.235, 0.891)	0.353* (-0.009, 0.750)	0.724*** (0.384, 1.030)
<i>funeu</i>	0.641*** (0.153, 1.188)	0.301 (-0.218, 0.838)	1.133*** (0.622, 1.646)	0.804*** (0.332, 1.353)
<i>gp</i>	0.411*** (0.106, 0.758)	0.1189 (-0.165, 0.541)	0.141 (-0.281, 0.601)	0.494*** (0.125, 0.853)
Firm-level random intercept				
$\hat{\sigma}_\gamma$	7.532*** (6.797, 8.297)	5.324*** (4.519, 6.750)	2.970*** (2.443, 3.457)	3.671*** (3.362, 4.615)
Sector-level random effects (averages over sectors)				
$\hat{\sigma}_{\lambda_0}$	0.634*** (0.273, 1.665)	0.701*** (0.279, 1.817)	0.676*** (0.245, 1.729)	0.614*** (0.272, 1.633)
$\hat{\sigma}_{lp}$	0.326*** (0.210, 0.842)	0.329*** (0.200, 0.855)	0.382*** (0.188, 0.909)	0.324*** (0.200, 0.836)
$\hat{\sigma}_{HHI}$	0.638*** (0.257, 1.666)	0.709*** (0.301, 1.917)	0.684*** (0.225, 1.661)	0.644*** (0.279, 1.761)

95 % posterior credible interval (PCI) in brackets; significance codes for the PCIs: \*\*\* 0.01; \*\* 0.05; \* 0.1; DIC = 5516.991



motivation of combining resources with other firms. However, the impact of market concentration on the firms' propensity to form R&D alliances is a theme empirically less explored. The empirical contribution of Hernán et al. (2003) showed a positive impact of market concentration on the propensity of firms to cooperate in R&D, since a more concentrated industry offers a greater opportunity for internalization of spillovers. Wang and Zajac (2007), instead, did not achieve clear-cut conclusions since they found different results for different model specifications. The estimated Herfindahl-Hirschman index is, on average, 8 %. The index range from, 0.015, not concentrated, to 0.378, moderately/highly concentrated sector.

## 5 Estimation results

In the current section, we present the results of the multivariate mixed-effects logistic regression model as in (5), showing how R&D collective interactions are firm-level processes with high heterogeneity of actors and activities, where a strong sectoral specificity exists. In particular, we distinguish between firm-level and sectoral-level determinants of R&D cooperation. Table 6 reports results for the complete sample of 1929 observations.

As it has been shown in previous literature, some (if not all) of the variables included in this model may be endogenous. In this paper, we do not control for omitted variables, selection, or simultaneity, therefore the estimates are causally uninterpretable. The estimation approach we propose is rather aiming at the decomposition of the conditional variance structure due to firm- and sector-level heterogeneity. In fact, the expected correlation among the different firms' cooperative strategies is expected to have a nested structure that could pick the multiplier effect of innovation policies.

With the potential lack of interpretation in mind, we note that the estimated coefficients statistically differ across the equations,<sup>7</sup> indicating the appropriateness of distinguishing between cooperation types. We observed that the innovativeness of a firm plays an important role in disentangling the determinants of R&D cooperation. In particular, as expected, developing radical innovation, therefore being a 'true' innovator, has a larger impact on all four cooperation types than performing only imitative innovations. It is interesting to note that the enterprises which introduced a new good or service in their reference market before their competitors, tend to form alliances with customers in the first place. As a matter of fact, introducing a radical innovation increases the odds of cooperating with customers by more than 2.5 times (i.e., it increases the probability to cooperate with customers by almost 72 %). On the other hand, introducing a new (to the firm) good already available from the competitors in the market in which the firm operates, enhances the chance to establish a formal R&D agreement with a competitor more than with other partners.

The hypothesis that incoming spillovers positively affect the probability of cooperation is confirmed only for two cooperation types. The spillover variable

---

<sup>7</sup> We formally test differences in the estimated coefficients using a Welch two-sample *t* test.

(*spill*), measured as the total pool of external knowledge that is potentially available for a firm, has a high and significantly positive impact only on competitors and institutional cooperation strategies (odd-ratios of 3.661 and 6.043, respectively). Higher incoming spillovers positively affect the probability of cooperating with research institutes and competitors, but have no effect on cooperation with customers or suppliers. Cassiman and Veugelers (2002) also do not find evidence of a significant impact of incoming spillovers on vertical cooperation, but do find statistical evidence of the positive relation between appropriability and the probability of cooperating with customers or suppliers. As a matter of fact, our results slightly differ from the existing studies in that, while they find a significant increase in the probability of cooperating with research institutes due to incoming spillovers, they do not find any significant effect on the propensity to collaborate with competitors. We suggest two different but correlated explanations to this phenomenon. The first concerns the construction of our spillover variable, the second relates to the multilevel structure of our model where firms are nested in sectors. The incoming spillovers are measured by the importance of publicly available information for the firm's innovation process, but consider as a source of information also the question relative to the importance of professional or industry associations for innovation activities. Such a construction of the spillover variable implicitly calls for a higher probability of horizontal cooperation. Indeed, a firm can exploit much better the information coming from an industry association in the context of a cooperative alliance within the same sector. On the other hand, we hypothesize that the coefficient of our spillover variable is not significant for vertical cooperation types in that, by definition, the variable does not consider inter-industry collaboration.

Given the multilevel structure of our model where firms are supposed to be nested in sectors, the spillover variable plays the most important role. Indeed, the synergy between this particular source of information and our innovative multilevel structure might be at the root of such a significant regression coefficient.

In line with empirical findings, firm size is positive and significant in all four cooperation strategies. Larger firms are more likely to have the right absorptive capacity required to engage in R&D cooperation, and this effect is stronger for cooperation with universities and suppliers. In line with previous studies we used the logarithm form, since it is natural that this effect is attenuated when the number of employees grows large.

Empirical literature generally found a positive impact of barriers to innovation such as costs and risks connected to the innovation process on the propensity of firms to cooperate. The propensity of firms in engaging in R&D cooperations with universities or other research institutions are expected to be positively correlated with the costs of innovation (Cassiman and Veugelers 2002). Indeed, the cost sharing motive is found to be an important determinant for firms when deciding to cooperate with customers or research institutes. On the other hand, cooperations with customers, other than to access to complementary knowledge, are aimed to reduce the risk associated with bringing an innovation to the market. As a matter of fact, the risk-sharing variable is found to be significantly positive and the magnitude

**Table 7** Firm- and sector-level random effects correlation matrix

Coop. type	Suppliers	Customers	Competitors	Institutional
Suppliers	1.000	0.962*** (0.908, 0.988)	0.987*** (0.966, 0.999)	0.889* (0.839, 0.948)
Customers		1.000	0.926*** (0.882, 0.966)	0.740*** (0.611, 0.860)
Competitors			1.000	0.931*** (0.874, 0.987)
Suppliers	1.000	0.723*** (0.346, 0.974)	0.724*** (0.306, 0.965)	0.714* (0.310, 0.973)
Customers		1.000	0.708*** (0.286, 0.975)	0.741*** (0.372, 0.979)
Competitors			1.000	0.690*** (0.274, 0.963)

95 % posterior credible interval (PCI) in brackets; significance codes for the PCIs: \*\*\* 0.01; \*\* 0.05; \* 0.1

of the regression coefficient is much larger than the one of the cost-sharing variable. Further, the risk factor is not significant for the other three collaborative agreements.

Furthermore, as expected, R&D personnel intensity has a positive effect on the probability of cooperation with all sorts of partners. In line with Veugelers (1997) and Piga and Vivarelli (2004), we find that the variables for public financial support for innovation activities have a positive and significant impact on almost all cooperation strategies. This may suggest that subsidies, especially those from the European Union, promote competitive R&D partnerships, in particular with research institutions. Lastly, we find that firms belonging to a domestic group are more likely to cooperate with suppliers or research institutes.

### 5.1 Firm- and sector-level heterogeneity

The hypothesis of a more complex structure of the heterogeneity of cooperation determinants is confirmed by our results. Both firm- and sector-level variances and covariances are found to be significant, meaning that enterprises within the same industry share similar characteristics (same random effects), which lead to correlation between research partners' choices.

The proportion of variation explained by the firm- and sector-level random intercepts varies across cooperation strategies. The firm- and sector-level variations of cooperating with suppliers account for the 86.5 and the 2.8 %, respectively, 82.0 and 6.6 % for cooperating with customers, 71.2 and 3.9 % for horizontal research alliances, and 75.5 and 3.1 % are the firm- and sector-level proportions of residual variance specific of cooperating with a public research institution.

In general, we find that the variance of the firm-level intercept is much larger than that of the sector-level. Concerning the coefficient estimates of sector-level variables, namely the Herfindhal index and the measure of legal protection, are reported at the bottom of Table 6. Both the legal protection measures adopted by the

**Table 8** Estimation results: multivariate probit

Variable	Customers	Suppliers	Competitors	Institutional
<i>rad</i>	0.351*** (0.072)	0.243*** (0.074)	0.198* (0.108)	0.161 (0.104)
<i>inc</i>	0.242*** (0.054)	0.090 (0.061)	0.170* (0.091)	0.138* (0.073)
<i>size</i>	0.110** (0.043)	0.160*** (0.023)	0.143*** (0.050)	0.190*** (0.028)
<i>spill</i>	0.389** (0.183)	0.474*** (0.076)	0.482* (0.264)	0.942*** (0.198)
<i>cost</i>	0.153** (0.076)	0.157* (0.080)	-0.083 (0.121)	0.293** (0.115)
<i>risk</i>	0.259*** (0.072)	-0.135*** (0.052)	0.010 (0.099)	-0.053 (0.127)
<i>rdpi</i>	0.173 (0.233)	0.574*** (0.213)	0.273 (0.295)	0.744*** (0.163)
<i>fungmt</i>	0.168** (0.067)	0.152** (0.074)	0.085 (0.112)	0.236** (0.095)
<i>funeu</i>	0.199 (0.151)	0.420*** (0.079)	0.557*** (0.143)	0.438*** (0.113)
<i>gp</i>	0.114 (0.094)	0.208** (0.097)	0.044 (0.140)	0.277*** (0.105)
<i>lp</i>	0.151*** (0.051)	0.090 (0.055)	0.047 (0.069)	0.141* (0.074)
<i>HHI</i>	0.984*** (0.346)	0.449** (0.183)	0.341 (0.366)	0.792*** (0.232)
<i>Intercept</i>	-2.276*** (0.158)	-1.866*** (0.155)	-2.601*** (0.193)	-3.086*** (0.243)
Correlations matrix				
Customers	1	0.761*** (0.041)	0.637*** (0.021)	0.640*** (0.037)
Suppliers		1	0.652*** (0.040)	0.686*** (0.027)
Competitors			1	0.615*** (0.051)

Standard errors in parentheses; sample of 1354 firms; Log pseudolikelihood = -1959.293; Wald  $\chi^2(6) = 824.763$ , Prob >  $\chi^2 = 0.000$

firm and the degree of industry concentration affect their motivation of combining resources with other firms.

The values of correlations between the different cooperation strategies at the both firm- and sector-level are summarized in Table 7. The upper part of the table reports the correlation between cooperation choices at the level of the enterprise. The firm-specific correlation coefficients are on average 30 % higher the sector-specific correlation coefficients.

Additionally, comparing the estimates of the multivariate mixed logit (Table 8) with the estimates of a multivariate probit in Table 8, we notice that, in general, for all cooperation strategies, developing radical innovation or performing imitative innovations do not have the same strong effect on the propensity to cooperate as with the multivariate mixed logit estimates. Regression coefficients' estimates are statistically significant at the 5 % only for cooperations with customers or suppliers. Another distinguishing feature of the multivariate probit consists in the fact that industry legal protection does not seem to play any remarkable role in enhancing the chance to cooperate with suppliers, competitors or research institutes. On the other hand, the level of industry concentration (HHI) appears to have a positive and significant impact on each of the R&D alliances, except for the one with competitors. This result is counterintuitive as we would expect a significant and high effect of the market concentration especially on the probability to cooperate horizontally with competitors, as such an alliance would offer a greater scope for internalization of spillovers (Hernán et al. 2003).

## 6 Evaluating the impact of public funding on innovative output

The success of all stages of innovation should be perceived as the outcome of a collaborative occurrence, arising from cognitive proximity (Cohen and Levinthal 1990). Proximity can be defined, other than from a geographical perspective, in terms of organizational and institutional proximity. In particular, Dosi (1999) claims that the production of innovative output is influenced by the “social embeddedness” of firms' routines and strategies. Simply put, the innovativeness of a firm is likely to be driven by socially specific factors, such as the nature of the local labor markets, workforce training institutions, financial institutions, mechanisms governing the support of business start-ups and development, science and technology policies, inter-firm and firm-institutional interactions.

Large part of European policy measures supporting innovation activities focus on both the promotion of R&D cooperation between the actors of the innovation system (which includes enterprises, universities, and research institutes), and on the improvement of the conditions for the uptake of innovations and/or to improve the articulation of demand in order to spur innovations and the diffusion of innovations.

Policy evaluation of European initiatives, regional, or national innovation policies usually relies on counterfactual analysis, and on the implementation of econometric techniques such as matching (Heckman et al. 1998; Klette et al. 2000; Almus and Czarnitzki 2003; Catozzella and Vivarelli 2014). Undertaking such a systematic approach goes beyond the scope of this paper. We rather want to show

**Table 9** Estimation results: Mixed linear model

Dep. var.: $Y_{Inno}$	Fixed coefficients	Random coefficients
$CO^{supp}$	-0.011 (-0.030, 0.012)	-0.008 (-0.029, 0.012)
$CO^{cust}$	0.030*** (0.008, 0.051)	0.025** (0.004, 0.047)
$CO^{comp}$	-0.007 (-0.030, 0.020)	-0.007 (-0.032, 0.015)
$CO^{inst}$	0.022** (0.002, 0.043)	0.027*** (0.006, 0.048)
$rdpi$	0.007*** (0.002, 0.013)	0.008*** (0.003, 0.014)
$funloc$	0.006 (-0.017, 0.028)	0.050*** ( $\hat{\sigma}_{loc}$ ) (0.032, 0.072)
$fungmt$	0.020*** (0.003, 0.035)	0.045*** ( $\hat{\sigma}_{gmt}$ ) (0.029, 0.062)
$funeu$	0.007 (-0.018, 0.033)	0.054*** ( $\hat{\sigma}_{eu}$ ) (0.034, 0.078)
<i>Intercept</i>	0.045*** (0.034, 0.057)	0.070*** ( $\hat{\sigma}_{int}$ ) (0.039, 0.107)
DIC	-1480.567	-1477.612

95 % posterior credible interval (PCI) in brackets; significance codes for the PCIs: \*\*\* 0.01; \*\* 0.05; \* 0.1

how the proposed hierarchical structure has an impact on the measured effects innovation policies on both cooperation propensity and innovative output.

In particular, in this paragraph, we investigate the relationship between innovative output, measured as the percentage of firm’s total turnover from goods and service innovations introduced during 2004 to 2006 that were new to the firm’s market ( $Y_{Inno}$ ), and innovative input, simply measured by the R&D personnel intensity ( $rdpi$ ). We then control for public financial support for innovation activities at different levels of government (regional,  $funloc$ ; central government,  $fungmt$ ; European Union,  $funeu$ ) and for cooperation with four types of partners. Therefore, the linear mixed model describing the impact of public funding on innovative sales is:

$$Y_{Inno,ij} = \beta_0 + \beta'_c y_{ij} + \beta_r rdpi_{ij} + \alpha'_j F_{ij} + v_{ij}, \tag{2}$$

where  $y_{ij} = (y_{ij}^1, \dots, y_{ij}^4)'$  is the vector of 4 binary variables taking value 1 when one of the 4 cooperation partner is selected. As before, we define  $\alpha_j$  as a normally distributed sector-level random effect, and  $F_{ij} = (funloc_{ij}, fungmt_{ij}, funeu_{ij})'$  as the vector of dummies proxying whether the firm received a funding from a regional, governmental, or European institution. The error term  $v_{ij}$  is assumed to be normally distributed.

Table 9 reports the estimates of two models. The first column displays estimated mean values and 95 % posterior intervals of a linear mixed model without sector-level random components.<sup>8</sup> The second column presents estimates of a linear mixed model with random effects and random intercept. Specifically, the effect of variables proxying for regional, central government and European Union R&D

<sup>8</sup> Since the MLE estimator and the mean of the posterior are asymptotically equivalent and their difference depends on the inverse of the square root of the sample size, the larger the sample size the narrower this difference. As our sample is pretty large (1929 observations), this difference is likely to be negligible.

support (*funloc*, *fungmt*, *funeu*) is considered to be random at the sectoral level. We also allow for the sector-specific intercept to be random.

Results confirm our beliefs. When considering a hierarchical framework, the policy interventions are relevant for innovative output. The realizations of sectoral intercept and slopes of the policy coefficients are expected to lie in 95 % positive intervals. That is, the marginal effect of a regional innovation policy on the percentage of innovative turnover will be between 0.032 and 0.072. The same way of reasoning applies for the other policies.

It is also interesting to note that, when disregarding the sector-specific random effects (first column), the central government R&D policy resulted to have the higher (and statistically significant at 1 % level) marginal effect on innovative output (0.020, compared with 0.006 for regional funding, and 0.007 for European funding). However, when turning to our multilevel approach that can take into account organizational proximities within sectors, the impact of European Union fundings has the biggest HPD interval (0.034, 0.078). European scientific and technological policies increase the innovative sales by a value falling in the aforementioned interval.

## 7 Conclusions

Using data from the last available 2006 edition of the Community Innovation Survey for the Netherlands, this paper contributes to the existing empirical literature, by proposing a methodology to study the determinants of innovative collaborative agreements and to assess the impact of public financial support to R&D. In particular, we explore the firm- and the sector- level heterogeneity of the determinants of either forming an R&D alliance, or selling innovative products, by considering a (generalized) linear mixed model.

The two steps of our analysis can be summarized as follows. In the first stage, we investigate the relationship between the factors driving the propensity to collaborate with different research partners, assuming a multivariate hierarchical logit model. The second step confirms the key role of the assumed multilevel structure, by considering the relationship between innovative output and innovative input, controlling for public financial support to innovation activities at different levels of government.

To our knowledge, this is the first attempt to model the both firm- and sector-level heterogeneity in the determinants of R&D partner's choices and innovation output. Our hypothesis of a heterogeneity across firms and sectors is confirmed by the results. All covariances are found to be significant. In other words, firms within the same industry share similar characteristics (same random effects), which lead to correlation between research partners' choices. This confirms that R&D cooperation, as well as the innovative production, is a firm-level process, where a strong sectoral specificity exists.

Taking into account this sectoral-organizational proximity can help assessing the right impact of R&D policies on innovative output. We have seen that, when using the multilevel approach the impact of public fundings has a positive and significant



sign, while, when omitting this nested framework the policies have a poor effect on innovative turnover.

Moreover, our suggested empirical framework can be brought to a deeper level of analysis, if data on markets were observed. As a matter of fact, understanding market dynamics could be the key to create more innovation-friendly market conditions that are necessary to reduce the time-to-market of new goods and to enable emerging sectors and/or markets to grow faster.

In these markets, for example, the removal of barriers would essentially contribute to the competitive process and lead to the emergence of new markets. Competitive-ness is here meant not only as the ability of the firm to come up with innovation from its internal technological strength, but also on its ability to access the innovation networking, that, as we have shown, depends on sector-specific networking and proximity (and, presumably, also on market-specific characteristics).

### Appendix 1: Multilevel heteroskedastic choice model

#### Appendix 1.1: Covariance structure

The firm- and sector-level random coefficients  $\gamma_i^c$  and  $\alpha_j^c \equiv (\alpha_{1j}, \dots, \alpha_{qj}^c)'$  are

$$\alpha_j^c \sim N_q(0, W) \quad \text{and} \quad \gamma_i^c \sim N(0, r),$$

where  $q$  is the number of random effects included in the model. The variances of the firm- and sector-level random components,  $r$  and  $W = \text{diag}(w_1, \dots, w_q)$ , respectively, are assumed to be invariant to cooperation choice  $c$ . We define the random effects for all sectors as  $\alpha^c \equiv ((\alpha_1^c)', \dots, (\alpha_j^c)')$ , for all firms as  $\gamma^c \equiv (\gamma_1^c, \dots, \gamma_N^c)'$ , and for all  $c$  cooperation strategies as  $\alpha \equiv ((\alpha^1)', \dots, (\alpha^C)')$ , and  $\gamma \equiv ((\gamma^1)', \dots, (\gamma^C)')$ . We then assume that the vector of firm- and sector level random intercepts and slopes,  $\gamma$  and  $\alpha$  have the following covariance structure:

$$\gamma \sim N(\mathbf{0}, \mathbf{G}_1) \quad \text{and} \quad \alpha \sim N(\mathbf{0}, \mathbf{G}_2).$$

$\mathbf{G}_1$  and  $\mathbf{G}_2$  are defined as the Kronecker product between matrices  $\mathbf{A}_1$ , and  $\mathbf{A}_2$ , and  $\mathbf{V}_1$ , and  $\mathbf{V}_2$ , i.e.,  $\mathbf{G}_1 = \mathbf{V}_1 \otimes \mathbf{A}_1$ , and  $\mathbf{G}_2 = \mathbf{V}_2 \otimes \mathbf{A}_2$  where

$$\mathbf{V}_1 = \begin{pmatrix} \varsigma_{11}^2 & \varsigma_{12} & \cdots & \varsigma_{1c} \\ \varsigma_{21} & \varsigma_{22}^2 & \cdots & \varsigma_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ \varsigma_{c1} & \varsigma_{c2} & \cdots & \varsigma_{cc}^2 \end{pmatrix} \quad \text{and} \quad \mathbf{V}_2 = \begin{pmatrix} \sigma_{11}^2 & \sigma_{12} & \cdots & \sigma_{1c} \\ \sigma_{21} & \sigma_{22}^2 & \cdots & \sigma_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{c1} & \sigma_{c2} & \cdots & \sigma_{cc}^2 \end{pmatrix} \tag{3}$$

are the firm- and sector-level cooperation strategy-specific covariance matrices with elements  $\varsigma_{c\tilde{c}} = \varsigma_{\tilde{c}c}$ , and  $\sigma_{c\tilde{c}} = \sigma_{\tilde{c}c}$ , for  $c \neq \tilde{c}$ . In our application, these are  $4 \times 4$  matrices, as we consider 4 types of R&D cooperation ( $C = 4$ ), and where

$$\mathbf{A}_1 = \text{diag}(r_1, \dots, r_N) \quad \text{and} \quad \mathbf{A}_2 = \text{diag}(W_1, \dots, W_J).$$

The matrices  $W_1, \dots, W_J$  have dimension  $q \times q$ , so that the dimension of  $\mathbf{A}_2$  is  $qJ \times qJ$ .<sup>9</sup>  $\mathbf{G}_1$  and  $\mathbf{G}_2$  are block matrices of dimension  $4N \times 4N$  and  $4qJ \times 4qJ$ , respectively. When the hypothesis of constant coefficients across firms and sectors is supported, it could seem reasonable to allow variations in parameters across cross-sectional units in order to take into account the firm and sectoral heterogeneity. The model specification can be generalized so as to take into account for both random coefficients and heteroskedasticity, by using the following specification for the random intercept and coefficients:

$$\alpha_j^c \sim N_q(0, W_j) \quad \text{and} \quad \gamma_i^c \sim N(0, \mathbf{r}_i).$$

The variances of the firm- and sector-level random components,  $r_i$  and  $W_j = \text{diag}(w_{1j}, \dots, w_{qj})$ , would then measure the degree of heterogeneity of each firm, nested in each sector. In this paper, the only source of heteroskedasticity which is explicitly taken into account is the one deriving from different cooperation alternatives,  $\zeta_{c\bar{c}}$  and  $\sigma_{c\bar{c}}$ .

### Appendix 1.2: Bayesian estimation approach

Assuming *conditional independence* of firm’s choice probabilities given the covariates and the random effects, we can write the unconditional marginal probability<sup>10</sup> of the response block matrix,  $\mathbf{Y} \equiv \left[ y_{ij}^c \right]_{4 \times N \times J}$ , where  $y_{ij}^c \in \{0, 1\}$  is the observed research cooperation choice, as

$$L(\mathbf{Y}|\mathbf{G}_1, \mathbf{G}_2) = \int \int \prod_j \prod_c \prod_i f_{ij}^c(y_{ij}^c|\theta) \pi_1(\gamma|\mathbf{G}_1) \pi_2(\alpha|\mathbf{G}_2) d\gamma d\alpha. \tag{5}$$

The maximum likelihood method is the standard approach for statistical inference in the mixed effects model. In order to maximize the sample likelihood, integration over the random-effects distribution must be performed. Yet, there exists no analytical solution for the intractable integral in Eq. (5). As a result, estimation

<sup>9</sup> If we set the error component  $z'_{ij}\alpha_j^c = d'_{ij}\alpha_j^c$ , where  $d_{ij}$  is a dummy variable that takes the value 1 if firm  $i$  is nested in sector  $j$  and zero otherwise,  $\alpha_j^c$  is reduced to a category-specific random intercept. In such a case  $W = w$  and  $\mathbf{A}_2 = \text{diag}(w_1, \dots, w_J)$  would be a simple diagonal matrix of dimension  $J \times J$ .

<sup>10</sup> Coull and Agresti (2000) derive a multivariate Binomial logit-normal distribution, where the  $c$  responses  $Y_i = (Y_{i1}, \dots, Y_{ic})$  with index vector  $m_i = (m_{i1}, \dots, m_{ic})$  are assumed to be independent binomial distributions, with success parameter vector  $\pi_i$ . Then the multivariate Binomial logit-normal model is expressed by incorporating a random effect, such that  $\text{logit}(\pi_i) = X_i\beta + z_i$ , where  $X_i$  is a  $c \times p$  covariate matrix and  $z_i$  is a  $c \times 1$  vector of random effects and is distributed as a multivariate normal distribution with mean vector  $\mathbf{0}$  and covariance matrix  $\Sigma$ . Then the probability density function of  $y$  is written as

$$p(y; \pi, m, \Sigma) = \int_{[0,1]^c} f_B(y|\pi, m) f_N(z; \Sigma) dz \tag{4}$$

where  $f_B(y|\pi, m)$  denotes the binomial probability mass function with  $m$  trials and success probability  $\pi$  and  $f_N(z; \Sigma)$  denotes the multivariate normal density function of  $z$ .

is much more complicated than in models for continuous normally distributed outcomes where the solution can be expressed in closed form. Various approximations for evaluating the integral over the random-effects distribution have been proposed in the literature; many of these are reviewed in Rodríguez and Goldman (1995).

Simulation methods are also popular techniques to estimate mixed effects models (Train 2009). The unconditional probabilities in equation (5) are approximated through simulation for any given value  $\theta$  of the parameters of the mixing distribution  $f(\alpha|\theta)$ . Such methods fall under the rubric of Markov Chain Monte Carlo (MCMC) algorithms.

In this paper we adopt a Bayesian approach and explore the MCMC fitting of the multivariate mixed logit model. One advantage of the Bayesian approach over its frequentist counterpart includes the fact that the Bayesian procedures do not require maximization of any function. For complicated random effects structures, computation of a single maximum likelihood fit can be expensive, making the simulation of statistics of interest computationally prohibitive. Second, with Bayesian procedures, estimation properties, such as consistency and efficiency, can be attained under more relaxed conditions than with classical procedures. As shown in Train (2009) (Chapter 10), consistency of the Maximum Simulated Likelihood (MSL) estimator depends on the relationship between the number of draws that are used in the simulation and the sample size. If the number of draws is considered fixed, then the MSL estimator does not converge to the true parameters, because of the simulation bias. The simulation bias disappears as the sample size rises without bound together with the number of draws. In contrast, the Bayesian estimators are consistent for a fixed number of draws used in simulation and are efficient if the number of draws rises at any rate with sample size.

Following the Bayesian approach, the model parameters  $\beta$ ,  $\alpha$ ,  $\mathbf{G}$ , summarized in the vector  $\theta$ , are treated as random variables. The assumed distributions for the parameters, called *prior distributions* and denoted by  $f(\theta)$ , borrow information from past studies, logic, or from the researcher's ideas about the values of these parameters. Therefore, the prior distribution represents how likely the researcher thinks it is for the parameters to take a particular value, over all possible values that the parameters can take. Bayesian inference is based on the *posterior distribution*,  $f(\theta|\mathbf{y})$ , which is the conditional distribution of the conjectured, but unknown, parameters  $\theta$ , given the observed data  $\mathbf{y} = \mathbf{y}_1, \dots, \mathbf{y}_n$ .

The choice of a prior distribution  $f(\theta)$  affects Bayesian estimation. In other words, Bayesian inference may be influenced by a “strong” prior. In absence of any prior information, a non-informative prior is chosen ( $f(\theta) \propto 1$ ) and Bayesian inference is asymptotically equivalent to likelihood inference. In practice, we always specify a diffuse prior for  $\beta$ , and try different values of the set of parameters  $\alpha$ ,  $\mathbf{G}$ , as a sensitivity analysis.

To estimate the parameters of the Generalized Linear Mixed Model (GLMM) defined in Sect. 4 following a Bayesian approach (Zeger and Karim 1991; Gelman et al. 2003), we use the R package `MCMCglmm` (Hadfield and Kruuk 2010). The default prior chosen by `MCMCglmm` for the regression model parameters  $\beta^c$  is a non-

informative, normal distribution  $N(0, 1e + 10)$ ,<sup>11</sup> while for both the residual and random-effect variance matrices a diffuse inverse-Wishart distribution is assumed, which is commonly used in practice. Then, assuming that the priors are independent,

$$f(\boldsymbol{\beta}, G) = f(\boldsymbol{\beta})f(G), \quad (6)$$

the posterior distribution can be written as

$$f(\boldsymbol{\beta}, G, \boldsymbol{\alpha} | \mathbf{y}) \propto \prod_{j=1}^n \prod_{c=1}^4 \prod_{i=1}^{n_j^c} \mathbf{f}_{ij}^c(\mathbf{y}_{ij}^c | \boldsymbol{\alpha}_j^c, \boldsymbol{\beta}^c) \mathbf{f}(\boldsymbol{\beta}^c) \prod_{j=1}^n \prod_{c=1}^4 \mathbf{f}(\boldsymbol{\alpha}_j^c | \mathbf{G}) \mathbf{f}(\mathbf{G}). \quad (7)$$

The R package `MCMCglmm` generate samples from the posterior distribution using Metropolis–Hastings updates (for more details on the sampling schemes, see Hadfield and Kruuk 2010). Beginning with the starting values  $(\boldsymbol{\beta}^{(0)}, \boldsymbol{\alpha}^{(0)}, W^{(0)})$ , after a warm-up (also called “burn-in”) period, we store a sample of  $(\boldsymbol{\beta}, \boldsymbol{\alpha}, W)$  from the posterior distribution. Once we generate a large number of samples, the posterior mean and posterior covariance can be approximated by the sample mean and the sample covariance based on the simulated samples. Convergence of the MCMC sampling scheme was assessed using empirical and test-based approaches (Heidelberger and Welch 1983; Geweke 1992). Results from convergence diagnostics indicated that it was sufficient to burn-in the first 15,000 samples and take the subsequent 1600 samples for inference.

## References

- Almus, M., & Czamitzki, D. (2003). The effects of public R&D subsidies on firms’ innovation activities: The case of Eastern Germany. *Journal of Business & Economic Statistics*, 21(2), 226–236.
- Arora, A., & Cohen, W. M. (2015). Public support for technical advance: the role of firm size. *Industrial and Corporate Change*, 24(4), 791–802.
- Belderbos, R., Carree, M., Diederer, B., Lokshin, B., & Veugelers, R. (2004a). Heterogeneity in R&D cooperation strategies. *International Journal of Industrial Organization*, 22(8–9), 1237–1263.
- Belderbos, R., Carree, M., & Lokshin, B. (2004b). Cooperative R&D and firm performance. *Research Policy*, 33(10), 1477–1492.
- Belderbos, R., Carree, M., & Lokshin, B. (2006). Complementarity in R&D cooperation strategies. *Review of Industrial Organization*, 28(4), 401–426.
- Busom, I., & Fernández-Ribas, A. (2008). The impact of firm participation in R&D programmes on R&D partnerships. *Research Policy*, 37(2), 240–257.
- Cassiman, B., & Veugelers, R. (2002). *R&D cooperation and spillovers: Some empirical evidence from Belgium*. Open access publications from katholieke universiteit leuven, Katholieke Universiteit Leuven.

<sup>11</sup> One of the many advantages of the package `MCMCglmm` resides in the great flexibility in the specification of various residual and random-effect variance structures. `MCMCglmm` allows variance structures of the form  $\mathbf{G} = \mathbf{V} \otimes \mathbf{A}$ : unstructured and completely parameterized covariance matrices. However, binary responses pose a special problem because the residual variance cannot be estimated because the variance is uniquely determined by the mean. Therefore, following Hadfield and Kruuk (2010), we apply restrictions on the prior distribution of the residual covariance matrix. In particular, we fix the parameters of the prior distribution at some value (1 for variances and 0 for covariances).

- Cassiman, B., & Veugelers, R. (2006). *In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition*. Open access publications from katholieke universiteit leuven, Katholieke Universiteit Leuven: Open access publications from katholieke universiteit leuven.
- Catozzella, A., & Vivarelli, M. (2014). The possible adverse impact of innovation subsidies: some evidence from Italy. *International Entrepreneurship and Management Journal*, 1–18.
- Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35(1), 128–152.
- Coull, B., & Agresti, A. (2000). Random effects modeling of multiple binomial responses using the multivariate binomial logit-normal distribution. *Biometrics*, 56(1), 73–80.
- Crespi, F., Ghisetti, C., & Quatraro, F. (2015). Environmental and innovation policies for the evolution of green technologies: A survey and a test. *Eurasian Business Review*, 5(2), 343–370.
- d'Aspremont, C., & Jacquemin, A. (1988). Cooperative and Noncooperative R&D in Duopoly with Spillovers. *American Economic Review*, 78(5), 1133–1137.
- Dewar, R. D., & Dutton, J. E. (1986). The adoption of radical and incremental innovations: An empirical analysis. *Management science*, 32(11), 1422–1433.
- Dosi, G. (1999). Some notes on national systems of innovation and production, and their implications for economic analysis. In D. Archibugi, J. Howells, & J. Michie (Eds.), *Innovation policy in a global economy*. Cambridge University Press.
- Gelman, A., Carlin, J., Stern, H., & Rubin, D. (2003). *Bayesian data analysis* (2nd ed.). UK: Chapman and Hall.
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In J. Berger, J. Bernardo, A. Dawid, & A. Smith (Eds.), *Bayesian statistics* (pp. 169–194). Oxford: Oxford University Press.
- Goldstein, H. (1995). *Multilevel statistical models* (2nd ed.). New York: Halstead Press.
- Hadfield, J., & Kruuk, L. (2010). MCMC methods for multi-response generalised linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, 33(2), 1–22.
- Hanley, A., Liu, W.-H., & Vaona, A. (2015). Credit depth, government intervention and innovation in China: Evidence from the provincial data. *Eurasian Business Review*, 5(1), 73–98.
- Heckman, J. J., Lochner, L., & Taber, C. (1998). General-equilibrium treatment effects: A study of tuition policy. *American Economic Review*, 88(2), 381–86.
- Hedeker, D., & Gibbons, R. D. (1996). MIXOR: A computer program for mixed-effects ordinal regression analysis. *Computer Methods and Programs in Biomedicine*, 49, 157–176.
- Heidelberger, P., & Welch, P. D. (1983). Simulation run length control in the presence of an initial transient. *Operations Research*, 31(6), 1109–1144.
- Henderson, R. (1993). Underinvestment and incompetence as responses to radical innovation: Evidence from the photolithographic alignment equipment industry. *RAND Journal of Economics*, 24(2), 248–270.
- Hernán, R., Marín, P. L., & Siotis, G. (2003). An empirical evaluation of the determinants of Research Joint Venture Formation. *Journal of Industrial Economics*, 51(1), 75–89.
- Kaiser, U. (2002). An empirical test of models explaining research expenditures and research cooperation: Evidence for the German service sector. *International Journal of Industrial Organization*, 20(6), 747–774.
- Kamien, M. I., Muller, E., & Zang, I. (1992). Research joint ventures and R&D cartels. *American Economic Review*, 82(5), 1293–1306.
- Katz, M. L. (1986). An analysis of cooperative research and development. *RAND Journal of Economics*, 14(4), 527–543.
- Kim, J. (2014). Formal and informal governance in biotechnology alliances: Board oversight, contractual control, and repeated deals. *Industrial and Corporate Change*, 23(4), 903–929.
- Kirat, T., & Lung, Y. (1999). Innovation and proximity. *European Urban and Regional Studies*, 6(1), 27–38.
- Klette, T. J., Moen, J., & Griliches, Z. (2000). Do subsidies to commercial R&D reduce market failures? Microeconomic evaluation studies. *Research Policy*, 29(4–5), 471–495.
- Korotayev, A. V., & Tsirel, S. V. (2010). A spectral analysis of world GDP dynamics: Kondratieff waves, Kuznets swings, Juglar and Kitchin cycles in global economic development, and the 2008–2009 economic crisis. *Structure and Dynamics*, 4(1), 3–57.
- Kultti, K., Takalo, T., & Tanayama, T. (2015). R&D spillovers and information exchange: a case study. *Eurasian Economic Review*, 5, 63–76.

- Leifer, R., Gina Colarelli, O., Rice, M., & Gina Colarelli, O. (2001). Implementing radical innovation in mature firms: The role of hubs. *The Academy of Management Executive (1993–2005)*, 15(3):102–113.
- Lopez, A. (2008). Determinants of R&D cooperation: Evidence from Spanish manufacturing firms. *International Journal of Industrial Organization*, 26(1), 113–136.
- Mohnen, P., & Röller, L.-H. (2005). Complementarities in innovation policy. *European Economic Review*, 49(6), 1431–1450.
- OECD, E. (1997). Proposed guidelines for collecting and interpreting technological innovation data: Oslo manual.
- Piga, C. A., & Vivarelli, M. (2004). Internal and external R&D: A sample selection approach. *Oxford Bulletin of Economics and Statistics*, 66(4), 457–482.
- Reinganum, J. (1983). Uncertain innovation and the persistence of monopoly. *The American Economic Review*, 73(4), 741–748.
- Rodríguez, G., & Goldman, N. (1995). An assessment of estimation procedures for multilevel models with binary responses. *J. Royal Statistical Society*, 158(1), 73–90.
- Schmitz, H. (1999). Collective efficiency and increasing returns. *Cambridge Journal of Economics*, 23(4), 465–483.
- Tether, B. (2002). Who co-operates for innovation, and why: An empirical analysis. *Research Policy*, 31(6), 947–967.
- Train, K. (2009). *Discrete choice methods with simulation* (2nd Edn.). Online economics textbooks: Cambridge University Press.
- Veugelers, R. (1997). Internal R&D expenditures and external technology sourcing. *Research policy*, 26(3), 303–315.
- Wang, L., & Zajac, E. (2007). Alliance or acquisition? A dyadic perspective on interfirm resource combinations. *Strategic Management Journal*, 28(13), 1291–1317.
- Zeger, S., & Karim, M. (1991). Generalized linear models with random effects; A Gibbs sampling approach. *Journal of the American statistical association*, 86(413), 79–86.