15919

Evaluating the effectiveness of public support on inbound open innovation: evidence from Spanish manufacturing SMEs

Abstract

Applying several matching estimators to a sample of Spanish manufacturing SMEs, we evaluate the impact of regional and federal funding on inbound open innovation strategies. Such SMEs are more likely to respond to public support by increasing either their cooperation with government institutions or their investment in extramural R&D than by enhancing cooperative networks. A policy corollary is to promote measures to attenuate cooperation failures. Methodologically, each type of funding and each type of cooperative partner should be considered separately. Moreover, overestimation of treatment effects when firms' unobserved characteristics are not addressed suggests that sensitivity analysis should complement matching estimation.

Keywords: inbound open innovation; SMEs; evaluation of public support; sensitivity analysis; coordination failure

1 Introduction

Evaluation of innovation policies, until recently, was mainly concerned with input and output additionalities, whereby input additionality occurs when firms increase their R&D investment as a result of receiving subsidies; and output additionality results in new patents, product and process innovations or increased innovative sales (a share of turnover from new products). Focusing on innovation inputs and outputs, however, means that we stay outside the "black box" of innovation processes, rather observing the beginning (innovation inputs) or end results (innovation outputs) of processes (OECD, 2006).

The concept of behavioural additionality invites us to go beyond input and output additionality and assess the impact of public measures on firms' innovative behaviour. The literature on additionality lacks a common definition of behavioural additionality. Most empirical studies explore network additionality (Georghiou and Clarysse, 2006), which occurs when firms expand their networks and cooperative activities as a result of participation in support programmes. Given that data on other types of additionality are not available, we follow this tradition.

The narrow concept of behavioural additionality (i.e. network additionality) encompasses the impact of public funding on inbound open innovation. In 2003, open innovation emerged as a new conceptual framework, emphasizing the role of networking and knowledge exchange on firms' innovativeness, and their critical role in creating and sustaining competitive advantages (Chesbrough, 2003). Open innovation is the subject of an increasing number of empirical studies, mainly focusing on the determinants of open innovation strategies and their impact on innovation and firm performance (for a comprehensive review, see Schroll and Mild, 2012).

Evaluation of public support should take into account endogeneity and consequent selection bias arising from two sources: self-selection of programme participants; and selection of potentially successful applicants by government agencies (Busom and Fernandez-Ribas, 2008). Due to often noted factors hampering econometric analysis (such as, lack of longitudinal data and of valid instruments for selection models),[1] matching estimation has become a widely used evaluation method in the literature on the effectiveness of innovation policy. Drawing on Community Innovation Survey (CIS) 2006 data, we employed several matching estimators to investigate the impact of public support on open innovation practices in Spanish small and medium-sized enterprises (SMEs).

Our study contributes to the evaluation literature by providing the first empirical findings on the impact of public innovation support on cooperative behaviour in SMEs (behavioural additionality), and on two additional inbound open innovation practices: outsourcing R&D; and acquiring other external knowledge (e.g. patents and know-how). The treatment effects are reported for two separate sources of funding: regional and national support programmes. Following the literature on the determinants of R&D cooperation, we explicitly take into account incoming spillovers, knowledge flows from different sources (suppliers, customers, competitors, government and Higher Education Institutions) and include barriers to innovation and to cooperation in our methodological framework. Finally, particular features of this study are the conduct of sensitivity analysis to assess the impact of unobserved heterogeneity on the estimated effects of public support programmes; and, moreover, that the conclusions of this study take into account the evidence and implications of unobserved heterogeneity.

The paper is organised as follows. Section 2 defines the concepts of open innovation, while Section 3 formulates the methodological framework, discusses model specification and

[1] These limiting factors are noted in most studies on the additionality of public support (see, for instance, Busom and Fernandez-Ribas, 2008; Czarnitzki et al., 2007; Czarnitzki et al., 2011).

data used in the study. Section 4 gives the main results from matching and discusses the evidence and implications of sensitivity analysis. Section 5 concludes.

2 Open innovation

The significance of cooperation in firms' innovation activities is reinforced with the concept of open innovation. With the advent of Chesbrough's (2003) seminal work, open innovation emerged as a new conceptual framework in innovation literature, in distinction to closed innovation systems (Lichtenthaler, 2011). This new paradigm acknowledges firms' limited internal innovative capacities and suggests that generating external knowledge is necessary for innovation processes, as firms no longer can be successful innovators by relying solely on internal capabilities.

Open innovation is defined as "the use of purposive inflows and outflows of knowledge to accelerate internal innovation, and to expand the markets for external use of innovation, respectively" (Chesbrough et al., 2006, p. 1). Knowledge flows aiming at fostering internal innovation are termed *inbound* open innovation (technology exploration or acquisition), while market expansion focusing on the commercialisation phase of the innovation process is termed *outbound* open innovation (technology exploitation or commercialization) (Dahlander and Gann, 2010; Lichtenthaler, 2011; Van de Vrande et al., 2009).[2] The process of technology exploration or acquisition (i.e. inbound open innovation) encompasses the following practices (Parida et al., 2012; Van de Vrande et al., 2009):[3]

---

[2] Inbound open innovation is also referred to as the outside-in process of open innovation, whereas outbound open innovation is referred to as the inside-out process of open innovation (Enkel et al., 2009).

[3] Dahlander and Gann (2010) divided inbound and outbound open innovation practices into two categories - pecuniary and non-pecuniary - whereby revealing and selling are non-pecuniary and pecuniary outbound innovation respectively, and sourcing and acquiring are non-pecuniary and pecuniary inbound innovation respectively.

- *Technology scouting*, that is, a process of gathering information and knowledge from the technological environment (Cohen and Levithal, 1990; Dahlander and Gann, 2010; Lichtenthaler and Lichtenthaler, 2009).
- *Customer involvement*: Customers can be involved in firms' internal innovation processes, which enables firms to develop new products or to modify the existing ones according to customers' needs and preferences.
- *External networking*: Networking on innovation is an important component of open innovation, and it encompasses both formal (e.g. R&D alliances) and informal cooperation on innovation with individuals and organisations (e.g. suppliers).
- *External participation*: This form of open innovation is associated with equity investment in other companies in order to access their knowledge or to benefit from other synergies.
- *Outsourcing R&D:* Extramural R&D activities performed by other firms or private and public organizations are an important alternative to intramural R&D.
- *Inward IP licensing (licensing-in)*: Firms can benefit from external knowledge through purchasing patents, trademarks, copyrights and other forms of IPs (Dahlander and Gann, 2010).

The process of technology exploitation or commercialization (i.e. outbound open innovation) includes several strategies:

- *Venturing*: In the context of open innovation, venturing refers to spin-offs, i.e. establishing new firms based on a firm's internal knowledge.
- *Outward licensing of Intellectual Property (IP)(licensing-out)*: This practice allows companies to generate profit from selling IPs to other companies (Dahlander and Gann, 2010; Lichtenthaler, 2011).

Open innovation practices are the subject of an increasing number of empirical studies over the last few years. The main research objectives are aimed at identifying the determinants of inbound and outbound open innovation strategies, and assessing their impact on firms' innovation performance (for a comprehensive review of empirical studies, see Schroll and Mild, 2012).

Both large firms and SMEs can greatly benefit from external knowledge. For instance, Soh (2003) found that repeated partnerships with other firms are positively associated with new product performance. Open innovation is particularly relevant to SMEs, because limited human and financial resources are critical barriers to internal innovation in those firms (Parida et al., 2012; Van de Vrande et al., 2009). Yet, limited resources can also have a detrimental effect on open innovation in SMEs, for instance, in acquiring extramural R&D or maintaining collaborative networks (Huizingh, 2011). Indeed, empirical evidence suggests that large firms engage in open innovation to a larger extent than SMEs (Bianchi et al., 2011; Lihtenthaler, 2008) and, within SMEs, medium-sized firms are more prone to opening up innovation processes than are small firms (Van de Vrande et al., 2009). Furthermore, Van de Vrande et al. (2009) find that SMEs mostly engage in user innovation (customer involvement) and in external networking. Larson (1991) notes that establishing and maintaining network partnerships is critical to entrepreneurial firms' survival and success. Conversely, the least practiced open innovations are outward and inward IP licensing, venturing and external participation, which require substantial financial resources, unlike customer involvement and external networking, which are often informal and need not entail significant financial investment.

Firms' strategic decisions on whether to develop new technologies and innovations by increasing in-house R&D or by external knowledge acquisition, depend on the type of technology. Innovation processes than involve generic (standardized) technological competences, should be developed by external knowledge exploitation either through cooperation or subcontracting (Narula, 2001). However, core technological competencies, which are the main source of firms' competitive advantage, should be developed internally. Furthermore, in discriminating between cooperation and R&D subcontracting, following the argument advanced in transaction costs economics, firms have incentive to opt for the latter when opportunism and free riding are more likely to occur, thus increasing transaction costs (Dhont-Petrault and Pfister, 2011). If we assume that opportunistic behaviour decreases with the increase in the level of technology standardization, this would mean that R&D subcontracting is more suitable for developing or enhancing standardized technologies (the 'standardization' hypothesis). Moreover, standardized technologies usually lack a degree of novelty sufficient to be patentable, thus suggesting that appropriability issues are less likely to occur.

Conversely, due to potential cooperation failure, firms can opt for R&D subcontracting for developing strategic, core technologies (the 'incentive' hypothesis). Cooperation failure refers to reduced R&D effort in cooperative partnerships when cooperating firms do not clearly specify which partner will be assigned the exclusive property rights (Dhont-Petrault and Pfister, 2011). For instance, Cassiman and Veugelers (2002) report a negative relationship between vertical cooperation and the effectiveness of appropriation methods. Moreover, Leiponen and Byma (2009) argue that small firms with close links to cooperative partners might face difficulties in protecting their returns to innovation. Unlike large firms, small firms utilize formal methods of protecting IPs (such as patenting) to a lesser extent, and rely more on informal methods such as secrecy and lead time (Leiponen and Byma, 2009). Therefore, to assure the maximum level of R&D effort, firms can assign exclusive property rights to the subcontractors, thus avoiding appropriability issues.

Finally, assessing the impact of public support on open innovation strategies is closely related to the concept of behavioural additionality (BA). While input and output additionality leave the black-box of innovation process unopened, BA goes beyond innovation inputs and outputs and aims at explaining what is happening inside the box. It is associated with intermediate effects of public support on firms' innovative behaviour (Georghiou and Clarysse, 2006). Following Busom and Fernandez-Ribas (2008), BA assesses the short-term impact of public programmes. Although the literature advances a broad perspective on BA, most empirical studies investigate only one segment of BA; that is the impact of public intervention on firms' cooperative behaviour (scope additionality as defined by Falk, 2007; or network additionality following the OECD, 2006, definition).[4] This narrow concept of BA encompasses the impact of public funding on inbound open innovation, specifically the effect on external networking. As previous studies do not investigate other forms of behavioural additionality, this inquiry, unlike other studies, expands research beyond cooperative networking to include two additional inbound open innovation strategies: outsourcing R&D; and acquisition of other external knowledge.

3 Methodology

3.1 Matching estimators

The main advantage of matching estimators, compared to selection models and IV approaches, is that they do not require any distributional assumptions regarding the error terms in the selection equation and in the outcome equation. However, matching estimators control only for firms' observed characteristics. In cases when unobserved characteristics of firms influence treatment assignment, matching yields biased estimates of treatment effects.

[4] Our study suffers from the same limitation; available data do not allow for exploring other categories of behavioural additionality.

The crucial step in the matching procedure is the choice of covariates, denoted *X*. The literature suggests that all observed variables that simultaneously affect a treatment and outcome should be included in the estimation of propensity scores (the selection equation) (Austin, 2011; Caliendo and Kopeinig, 2008; Ho et al., 2007; Steiner et al., 2010). The next step in the propensity score matching is the estimation of the propensity score. Since the propensity score is a probability of receiving a treatment (in our case, public subsidies), researchers can choose any discrete choice model, because both probit and logit models tend to yield comparable results (Caliendo and Kopeinig, 2008).

For the sake of brevity, we will not review a full range of matching estimators, but instead will focus on those applied in our study (for a review of matching estimators, see Austin, 2011; Imbens, 2004; Morgan and Harding, 2006; Stuart, 2010). Nearest Neighbour (NN) matching is the most commonly used matching estimator in innovation literature (Czarnitzki et al., 2007). The propensity score can be used to construct matched pairs by applying three methods (Guo and Fraser, 2010): i) nearest matching on the estimated propensity score; ii) Mahalanobis metric matching including the estimated propensity score with other matching variables;[5] and iii) nearest Mahalanobis metric matching with calipers based on the propensity score. The third method is superior to others with respect to balancing of the covariates between a treatment and comparison group (Rosenbaum and Rubin, 1985). In choosing the optimal caliper size, Cochran and Rubin (1973) note that 98% of the bias on a normally distributed covariate is removed with the caliper of 0.2 of the standard deviation of the estimated covariate (in the case of PSM, the caliper is based on the estimated propensity score).

The purpose of matching estimators is to balance observed covariates *X* between treated and untreated units. As discussed, nearest Mahalanobis metric matching with caliper based on

[5] This matching method is termed hybrid matching (Czarnitzki et al., 2011).

the propensity score results in the best balancing quality and, for that reason, we apply this estimator. Matching arguments, besides the estimated propensity score, are binary indicators for small firms and industries. The inclusion of additional matching arguments is motivated by the arguments advanced in the literature on SME innovation, whereby SMEs are a heterogeneous group of firms and their innovative activities should be analysed at industry level (Nooteboom, 1994; for the same empirical strategy see Czarnitzki et al., 2007; Czarnitzki et al., 2011; Spithoven et al., 2012).

After the estimation of the propensity score, but prior to applying a chosen matching estimator, a balancing test should be conducted. The purpose of a balancing test before matching (stratification test) is to check how well the estimated propensity score has succeeded in balancing covariates.[6] This approach requires the division of the sample into strata conditional on the propensity score, and checking whether there are no statistically significant differences between the means of the propensity scores of the treated and non-treated firms within each stratum. If the difference in means is statistically insignificant, then covariates are well balanced between matched pairs (Austin, 2011; Caliendo and Kopeinig, 2008; Lee 2013; Stuart, 2010).

The literature identifies several approaches for assessing the matching quality after matching. The first approach consists of comparing the standardized bias before and after matching. The rule- of- thumb adopted in most empirical studies is that a standardized bias below 3% or 5% is acceptable (Caliendo and Kopeinig, 2008). The second approach is based on the t-test statistics, whereby we check whether there are statistically significant differences in the means of covariates *X* between treated and non-treated firms after the matching. Significant differences after matching imply low matching quality. Finally, the matching quality can be assessed by checking the joint significance of all covariates in the selection

[6] Balancing tests before matching should not be confused with balancing tests after matching.

equation based on the likelihood-ratio (LR) test. All variables should be jointly significant before matching and jointly insignificant after matching. Furthermore, one can estimate the propensity score only for matched treated and non-treated firms and compare the pseudo- R^2 before and after matching. Low pseudo-R^2 after matching indicates a good matching quality (Caliendo and Kopeinig, 2008; Sianesi, 2004).

For a robustness check, we use two matching estimators. The first is kernel matching, whereby weighted averages of most units in the control group are used to estimate a counterfactual outcome.[7] The major advantage of this non-parametric estimator is the reduction in variance as the entire sample of the control group is used in the matching algorithm.

The second estimator is Inverse Probability of Treatment Weighting (IPTW), which uses weights based on the propensity score to create an artificial population in which treatment assignment is independent of the exogenous covariates *X*. The purpose of weighting is similar to using survey sampling weights to obtain weighted survey samples that are representative of a population (Austin, 2011). The variance estimation of the IPTW estimator has to take into account that weights are used to create an artificial sample. It is a common practice to use robust variance estimation (Austin, 2011; Emsley et al., 2008).

3.2 Model specification

Available data allows us to explore how public support affects several inbound open innovation strategies: customer involvement (cooperation with customers); external networking [i.e. aggregate cooperation as well as cooperation with suppliers, competitors, consultants, Higher Education Institutions (HEIs; e.g. universities and polytechnics) and government institutions]; outsourcing R&D; and acquisition of other external knowledge. As the CIS data do not contain information on outbound open innovation, we are not able to

[7] How many comparison units will be used depends on the choice of bandwidth.

assess the effectiveness of public support on those open innovation practices. Further, we separately analyse receipt of local/regional support (*FUNLOC*) and of national support (*FUNGMT*).[8]

Outcome variables are defined as follows (see Table A.1 for variable definitions and descriptive statistics):

- Aggregate cooperation (*COOPERATION*): DV=1 if firms cooperate with any partner: consumers, suppliers, universities or other higher education institutions (HEIs), consultants, government or competitors; otherwise zero;
- Cooperation with consumers (*COOP_CUSTOMERS*): DV=1 if firms cooperate with clients or customers, otherwise zero;
- Cooperation with suppliers (*COOP_SUPPLIERS*): DV=1 if firms cooperate with suppliers, otherwise zero;
- Cooperation with competitors (*COOP_COMPETITORS*):[9] DV=1 if firms cooperate with competitors or other firms in the sector, otherwise zero;
- Cooperation with consultants (*COOP_CONSULTANTS*): DV=1 if firms cooperate with consultants, commercial labs or private R&D institutes, otherwise zero;
- Cooperation with HEIs (*COOP_HEI*): DV=1 if firms cooperate with universities or other higher education institutions, otherwise zero;
- Cooperation with government (*COOP_GOVERNMENT*): DV=1 if firms cooperate with government or public research institutes, otherwise zero.
- Outsourcing R&D (*OUTSOURCING_RD*): DV=1 if firms conduct extramural R&D activities, otherwise zero;

[8] For the sake of brevity, in further text we use the term regional support, but this refers to either local or regional public funding.

[9] Cooperation with competitors is termed horizontal cooperation (Busom and Fernandez-Ribas, 2008) or co-opetition networks (Lechner et al., 2006).

- Acquisition of other external knowledge (*EXTERNAL_KNOWLEDGE*): DV=1 if firms purchase or license patents, know-how, and other types of knowledge from other firms, otherwise zero.

Although our sample is restricted to SMEs, we further include a dummy variable for small firms (*SM*) with more than 10 and fewer than 50 employees. SMEs are a heterogeneous category, and public support could have a differential effect on small firms relative to medium-sized firms (Curran, 2000).

A novelty of the study is the inclusion of barriers to innovation in the estimation of propensity scores (Becker and Dietz, 2004). The correlation matrix between seven variables measuring barriers to innovation indicates that multicollinearity might exist between these constraining factors.[10] Thus, to avoid multicollinearity, we omit four and include three variables: too high innovation costs (*BARRIER3*); a lack of qualified personnel (*BARRIER4*) and difficulties in finding cooperative partners (*BARRIER7*) (the variables measured as scores: 0 - no importance; 1 - low importance; 2 - medium importance; and 3 - high importance). The resource-based theory of the firm posits that resources are a crucial determinant of firms' competitive advantages (Barney, 1991; Peteraf, 1993). For SMEs, limited human and financial resources are critical factors in hampering innovation activities (Hewitt-Dundas, 2006; Madrid-Guijarro et al., 2009), and justifies the inclusion of the aforementioned barriers to innovation. In addition, limited internal resources and competencies can, at least partially, be compensated through cooperation with network partners (Lee et al., 2010; Parida et al., 2012).

[10] Seven barriers are as follows: lack of funds within enterprise or group; lack of finance from sources outside a firm; innovation costs too high; lack of qualified personnel; lack of information on technology; lack of information on markets; and, difficulty in finding cooperation partners for innovation.

The following variables are included to control for firms' absorptive capacity: patent activities (*PROPAT*); and whether firms continuously innovate (*CONTINUOUS_RD*). The reason to model these variables is that public agencies could adopt a strategy of picking the winners (Czarnitzki et al., 2007; Spithoven et al., 2012). In that case, government selects those firms that have a record of successful innovation.

Our model also includes a dummy variable for belonging to a group (*GP*). This variable can capture a twofold effect; a directly positive effect on cooperation, as firms that are a part of the enterprise group could be more likely to cooperate with other firms within a group (Czarnitzki et al., 2007); and/or, an indirect effect via an adverse effect on the probability of receiving support. Some support measures are restrictive insofar as SMEs that are part of a group are not eligible to apply. Thus, belonging to a group can be a barrier to participation in support programmes (Almus and Czarnitzki, 2003).

We model exporting activities (*EXPORT*) as a binary indicator equal to one if firms export and zero otherwise. Exporting can have a positive impact on cooperation, given that exporters potentially have a larger and/or more diverse network of cooperation partners than do non-exporting firms. Furthermore, exporting firms might have more incentive to innovate as a result of competitive pressure on international markets (Busom and Fernandez-Ribas, 2008; Czarnitzki and Lopes-Bento, 2013).

Another novelty of our study is the inclusion of sources of information in the selection equation. Following the literature on determinants of R&D cooperation (Cassiman and Veugelers, 2002; Chun and Mun, 2012), we model incoming spillovers proxied by the importance of various sources of information, such as: (a) conferences, trade fairs and exhibitions (*INCOMING1*); (b) scientific journals and publications (*INCOMING2*) and (c) professional and industry associations (*INCOMING3*). Furthermore, we include:

- Internal source of information, to measure the importance of information within a firm or enterprise group (*INFO_INTERNAL*);
- Market sources of information: from customers (*INFO_CUSTOMERS*), from suppliers (*INFO_SUPPLIERS*), competitors (*INFO_COMPETITORS*) and consultants, commercial labs or private R&D institutes (*INFO_CONSULTANTS*); and
- Institutional sources: from universities (*INFO_HEI*) and from government or public research institutes (*INFO_GOVERNMENT*);

All variables are measured as scores (0 - no importance; 1 - low importance; 2 - medium importance; and 3 - high importance).

In addition, the balancing test before matching reported that two variables (*INFO_INTERNAL* and *INFO_SUPPLIERS*) were not balanced in the propensity score model where the treatment variable is government support (*FUNGMT*). Following the literature on matching estimators discussed in Section 3.1, if the propensity score model is not balanced before matching, it should be re-specified by adding interaction terms and/or polynomials. We added two covariates (*INTERNAL_SM and SUPPLIERS_SM*), created as interaction terms between a binary indicator for small firms and the two unbalanced covariates (*INFO_INTERNAL* and *INFO_SUPPLIERS*). After these additional covariates were added to the propensity score model, covariate balance before matching was achieved. We used this specification of the propensity score model for each treatment variable, which will enable us to compare the treatment effects of both sources of funding.

To control for industry heterogeneity, based on the NACE classification at the 2-digit industry level, we include sectoral dummy variables for each of the fourteen manufacturing industries (see Table 1 for variable definition).[11] The base category is *INDUSTRY9* (sector 25 - Manufacture of rubber and plastic products).

3.3 Data

Our study employs Spanish CIS2006 survey data covering the period 2004-2006.[12] The sample consists of 8,022 small and medium-sized enterprises (SMEs) in manufacturing sectors, from which 5,115 are small and 2,907 are medium-sized firms.[13] Around a quarter of the sample participated in regional programmes (1,854 SMEs or 23.1 per cent) and less than 20 per cent received federal government support (1,312 firms or 16.4 per cent). Furthermore, 534 firms (6.7 per cent) received both regional and government support.

Descriptive statistics are presented in Table A.1 (see Appendix A). Only one-fifth of SMEs cooperate on innovation (22.2 per cent).[14] Regarding cooperation partners, the largest number of firms cooperate with suppliers (10.7 per cent) followed by government institutions (8.8 per cent) and universities (7.0 per cent). The smallest numbers of firms engage in

---

[11] There are two discrepancies between the NACE two-digit classification and the CIS anonymised microdata. Firstly, sector 31 - Manufacture of electrical machinery and apparatus is a medium-high tech sector, but it is aggregated with three high-tech sectors: 30 - Manufacture of electrical and optical equipment; 32 - Manufacture of radio, television and communication equipment; and 33 - Manufacture of medical, precision and optical instruments. Secondly, sector 23 - Manufacture of coke, refined petroleum products and nuclear fuel is a medium low tech sector but is aggregated with sector 24 - Manufacture of chemicals and chemical products, which is a medium high tech industry.

[12] Anonymised data in a micro-aggregated form are provided by Eurostat. Several studies use anonymised micro-data: for instance, Mohnen and Hoareau (2003) from the second wave of the CIS; and Grimpe and Sofka (2008) from the third wave of the CIS. Spithoven et al. (2012) use anonymised micro-data from the third CIS wave to analyse five EU countries. In addition, the authors compare the empirical results for Belgium from both the original and the anonymized data and report that 'the results showed very similar parameter estimates' (p. 69).

[13] Small firms are defined as those employing more than 10 and less than 50 workers, while medium-sized firms employee between 50 and 250 workers.

[14] We refer to inter-organizational cooperation for improving firms' innovative capabilities (Faems et al., 2005).

horizontal cooperation with competitors (3.6 per cent). With respect to innovation activities, only 11.3 per cent applied for a patent in the period covered by the survey, while 34.5 per cent of firms continuously engage in R&D activities, and one-fourth of SMEs undertook extramural R&D activities (24.7 per cent). Furthermore, a large number of SMEs are exporters (68.6 per cent). Among various sources of information, the most important are internal sources (mean value of 2.1), followed by customers and suppliers (mean values of 1.4 and 1.5 respectively). The least important source of information is from government and public research institutes (mean value of 0.4).

Table A.2 (see Appendix A) presents numbers and percentages of SMEs according to their cooperative behaviour and participation in support programmes. Out of 8,022 firms, more than two-thirds of firms neither cooperate on innovation nor participate in support programmes (63.5 per cent of firms from the perspective of regional support; and 67.6 per cent of firms from that of government support). By contrast, the percentage of firms that both cooperate and participate in public funding is rather low (with 8.9 per cent receiving local/regional support; and 6.1 per cent receiving national support). A similar pattern is found for participating firms that undertake extramural R&D activities (10.1 per cent participating in regional support measures; and 7.1 per cent receiving federal government support). Only a very modest share of participating firms acquires other types of external knowledge (0.8 per cent of those participating in regional funding; and 0.7 per cent receiving national support).

4 Results and discussion

4.1 Treatment effects

We estimated the impact of public support on various types of cooperation (vertical, horizontal, and private-public partnerships etc.), as well as on two additional open innovation practices: outsourcing R&D; and acquisition of other external knowledge.[15] Our objective is to estimate the Average Treatment Effect on the Treated (ATT), which indicates the difference in outcomes of the treated firms with and without treatment (the counterfactual outcome). The choice of three matching estimators is motivated by suggestions advanced in the literature on matching. We estimated the Nearest Neighbour (NN) matching with the Mahalanobis metric and a caliper of 0.2 of the standard deviation of the propensity score, because, as discussed in Section 3.1, this estimator results in the best covariate balance after matching (D'Agostino, 1998; Cochran and Rubin, 1973). However, in our study, the number of matching arguments in the Mahalanobis metric amounted to eleven, which could be the reason why the matching balance was worse than found after kernel matching (Stuart and Rubin, 2008).[16] For robustness checks, we applied two additional matching estimators: kernel matching; and the IPWT estimator.

Table 1 presents the estimated Average Treatment Effect on the Treated (ATT) for three sources of funding.[17] With respect to behavioural additionality, the overall results strongly indicate a positive but differential impact of public support for each source of funding. Although estimated ATT effects are fairly consistent across the three matching estimators, we will interpret the results from kernel matching, because the latter resulted in the best balance

[15] For the sake of brevity, we do not report the estimated coefficients and marginal effects from the probit models. However, they are available upon request. The purpose of the probit models is to estimate propensity scores, and even those variables that are statistically insignificant should remain in the model, if we suspect that they might affect the outcome variables. The critical step is to check the balancing property after the estimation of the probit models.

[16] NN matching with the Mahalanobis metric might result in a lower matching quality if the number of matching arguments is larger than eight or the covariates are not normally distributed.

[17] The ATT effects are estimated in the region of common support, as discussed in Section 4.1. Very few observations are lost due to the common support restriction, and this indicates a large overlap of estimated propensity scores among participating and non-participating SMEs. Regions of common support are not reported, but are available upon request.

after matching for each source of funding. The results of the balancing tests are reported in Table 2 below. The ATT effect of regional programmes on aggregate cooperation is 14.1 percentage points (p.p.) of an increase in the probability of cooperating; and of national programmes 8.5 percentage points.

A comparison between treatment effects of regional and government support reveals that participation in regional programmes has a larger effect on most types of cooperation than does participation in national programmes; the exceptions are cooperation with competitors (horizontal cooperation) and cooperation with HEIs. Moreover, the largest ATT effect is found for cooperation with government institutions for both sources of funding (for regional programmes, 11.8 p.p. increase in the probability of cooperating with government institutions; and for national support, 8.4 p.p. increase).

[Insert Table 1 here]

On the other hand, the smallest ATT effect of regional support is reported for cooperation with competitors (2.7 p.p. increase of the probability of cooperating with competitors), and of national support for vertical cooperation (2.9 p.p. increase in the probability of cooperating with customers and 2.6 p.p. increase in the probability of cooperating with suppliers) together with cooperation with consultants (2.8 p.p. increase in the probability of cooperating with consultants).

[Insert Table 2 here]

Turning to open innovation strategies other than cooperation, the most emergent finding is reported for outsourcing R&D. Participation in both regional and government support programmes results in a larger effect on extramural R&D activities than on either the aggregated or the disaggregated categories of cooperation. In contrast, receiving public support from regional programmes has no effect on the acquisition of external knowledge, and for SMEs participating in government programmes only a very small effect (1.2 p.p. increase in the probability of acquiring other external knowledge).

4.2 Sensitivity analysis

As discussed in Section 3.1, the main drawback of matching as an evaluation method is that it controls only for selection on observables. Firms' innovative behaviour as well as the selection process can be affected by unobserved characteristics, such as managerial attitude toward innovation (Busom and Fernandez-Ribas, 2008). This unobserved heterogeneity is referred to in evaluation literature as hidden bias. The presence of hidden bias indicates a failure of the identifying assumption on unconfoundedness or the selection on observables (the conditional independence assumption - CIA).[18]

The evaluation literature proposes several tests for the presence of hidden bias. The results of these tests should be taken with caution, as they cannot directly confirm whether the CIA holds; rather, they can indicate whether hidden bias arises or not. However, testing for unobserved heterogeneity should always complement a propensity score analysis, as the assumption on unconfoundedness cannot be tested directly (Guo and Fraser, 2010). Naturally, the ideal robustness check would be to apply those evaluation methods that control for unobserved heterogeneity. However, as discussed in the introductory section, the lack of valid instruments precludes this empirical strategy.

[18] The assumption of unconfoundedness refers to the condition that potential outcomes are independent of a treatment assignment, conditional on the observed covariates *X*, which are not affected by treatment (pre-treatment variables).

Sensitivity analysis is not common in empirical studies on the additionality of innovation policy. No study on behavioural additionality reports any type of sensitivity analysis. Moreover, to our knowledge, only the study on input additionality by Alecke et al. (2011) reports the results of sensitivity analysis.[19] The authors adopted the same Rosenbaum bound approach as in our study (Rosenbaum, 2002). The idea behind the Rosenbaum bounds approach is to determine how large the impact of an unobserved variable has to be to render the treatment effect statistically insignificant, under the assumption that this variable simultaneously affects both treatment assignment and the outcome variable (DiPrete and Gangl, 2004).

Values of gamma (Γ) show the magnitude of the factor by which unobserved heterogeneity may cause matched pairs to differ in their odds of treatment assignment. When gamma has a value of 1, the treatment effect is free of hidden bias. If unobserved characteristics have no influence on the causal inference, then again Γ=1 and estimated ATT effects and associated confidence intervals are unbiased (Li, 2012). Higher values of gamma indicate a departure from random assignment (selection) on observables.[20]

Table 3 reports the results of a sensitivity analysis of the main empirical results. Because our approach to sensitivity analysis cannot be applied to the kernel matching estimator, we begin by estimating the ATT effects from NN matching without replacement and with a caliper of 0.2 of the standard deviation of the estimated propensity score. Besides the ATT

[19] However, we believe that the authors did not correctly apply the test. The user-written command for Stata, *mhbounds*, can only be used for two types of matching estimators: NN matching without replacement; and stratification. Alecke et al. (2011) employ kernel matching. To our understanding, *mhbounds* cannot be applied to kernel matching.

[20] For instance, if gamma is equal to two, treated units are twice as likely to receive treatment as untreated (control) units. Furthermore, causal inference is sensitive to hidden bias for large gamma. Keele (2010) notes that using gamma between 1 and 2 is sufficient for sensitivity analysis, as for larger values of gamma most treatment effects are not robust to hidden bias.

effects estimated by applying NN matching without replacement, the table reports those gamma values for which the 5% significance levels of the upper bounds indicate whether results are sensitive to unobserved heterogeneity. The null hypothesis is no treatment effect (columns titled Hidden bias at 5 %); that is, an unobserved covariate renders the ATT insignificant. The implication of a non-rejection of the null means that the reported ATT effect is spurious, as it does not take into account variations in unobservables. For positive treatment effects, we are interested in the upper bounds indicating a possible overestimation of the true treatment effects (Becker and Caliendo, 2007).[21]

Unfortunately, the literature on sensitivity analysis does not provide clear guidance as to which value of gamma should be taken as a threshold for concluding whether or not a study is robust to hidden bias. Based on the proposal advanced by DiPrete and Gangl (2004), that a critical value of gamma depends on the research question, Lee and Lee (2009, p. 103) argue as follows their labour market study: "If more track records for the sensitivity parameters are established in future through more applications so that researchers can agree on how big is big for sensitivity analysis parameters, then the sensitivity analysis may become useful tools in dealing with unobserved confounders." Given that only one study in the literature on R&D and innovation policy includes a sensitivity analysis (that of Alecke et al., 2012),[22] we consulted empirical studies in labour market economics (Aakvik, 2001; Caliendo et at., 2005; Hujer et al., 2004) and adopt the recommended threshold of $\Gamma=1.5$. Therefore, if the significance level is below a p-value of 5% for $\Gamma\leq 1.5$, we report that a model is sensitive to unobserved heterogeneity. Conversely, if the significance level is above 5% for $\Gamma>1.5$, we conclude that the model is robust to hidden bias. In the analysis, we set the maximum value for Γ to 2 with increments of 0.05.

[21] The null hypothesis of underestimated effects is rejected at the 1 % significance level in most cases.

[22] The threshold in their study is $\Gamma=3$.

Sensitivity analysis reveals that most of the estimated treatment effects are sensitive to hidden biases. Analysing each source of funding separately, the Rosenbaum bound approach suggests the following:

- In the case of regional support, the models that are less sensitive to unobserved heterogeneity are those with the following outcome variables: aggregate cooperation; cooperation with government (the least likely to be affected by hidden bias); and outsourcing R&D. The remaining models are rather sensitive to selection bias.
- In the case of national treatment assignment, deviations from the underlying conditional independence assumption (CIA) are less likely to occur in the models with horizontal cooperation and with public institutions. For the remaining models, Rosenbaum's bounds indicate that ATT effects are sensitive to hidden bias.

It is important to note that the results from a sensitivity analysis adopting the Rosenbaum bounds are the worst-case scenarios (DiPrete and Gangl, 2004). For instance, in the model with cooperation with suppliers (for regional support), the estimated ATT effect is sensitive to hidden selection bias for $\Gamma \geq 1.25$. However, this does not necessarily mean that there is no true positive effect of public support on cooperation with suppliers. The result suggests that, if there were to be a confounding variable with a large effect on both treatment assignment and the outcome variable and if that variable were to increase the odds ratio of receiving a treatment for participating firms by 25 per cent (i.e. $\Gamma=1.25$) then the confidence interval for the ATT effect would include zero (DiPrete and Gangl, 2004).

The overall conclusion from sensitivity analysis suggests that hidden bias is unlikely to occur only in the case of cooperation with government agencies; and, to a lesser extent, in models of aggregate cooperation (regional support); cooperation with customers (borderline for regional support), cooperation with competitors (government support); and the outsourcing of R&D (regional support and borderline for government support). On the other side, hidden bias is most likely to arise in modelling cooperation with suppliers, consultants and Higher Education Institutions (HEIs). Finally, the models in which the outcome variable is the acquisition of external knowledge are least robust to unobserved heterogeneity, as hidden bias arises even at gamma equal to 1.

Our findings raise several issues. First, sensitivity analysis should be a necessary step when the effectiveness of R&D and innovation policy is assessed with the PSM analysis, as the findings indicate that treatment effects could be overestimated when firms' unobserved characteristics are not controlled for. Although a sensitivity analysis is considered to be an integral part of the PSM analysis (Caliendo and Kopeinig, 2008; Guo and Fraser, 2010), it is not adopted as a common practice in empirical innovation studies. However, a lack of sensitivity analysis is not only pertinent to innovation studies; Pearl (2009) points out that researchers often assume that the assumption of strong ignorability (i.e. CIA) holds because a large number of covariates is included in estimating a propensity score. However, it is not enough to recognize the major limitation of the PSM analysis; we should also examine whether selection on observables is likely to be satisfied. Although a sensitivity analysis cannot directly test the assumption, it can gauge the level of robustness of empirical findings to hidden bias.

Second, given the dominance of matching estimators in empirical studies, empirical evidence should be treated with caution. Most empirical studies reviewed in Section 3 report a positive impact of public support on firms' cooperation for innovation. Our results suggest that, depending on the type of cooperative partners, particular treatment effects could be overestimated: with respect to regional support, these are the estimates for cooperation with customers (on the borderline of our threshold), with suppliers, with competitors, with consultants and with HEIs as well as for acquisition of other external knowledge; regarding national funding, these are the models for aggregate cooperation, cooperation with customers, with suppliers, with consultants and with HEIs as well as for outsourcing of R&D and acquiring other external knowledge.

[Insert Table 3 here]

For a robustness check, following Busom and Fernandez-Ribas (2008), we restricted the sample to those firms that reported positive intramural R&D expenditures, which enables us to focus on innovative firms. The estimated treatment effects are consistent with those reported for the whole sample, although slightly increased. Furthermore, subsequent sensitivity analysis indicates that six models from this sub-sample are robust to hidden bias, compared to five models for the whole sample.[23]

5 Conclusions

Our study reports a positive, but heterogeneous impact of public support on open innovation in Spanish SMEs. However, sensitivity analysis suggests that many of the programme effects could be overestimated due to unobserved heterogeneity, which matching estimators cannot account for. Notably, the results for two cooperative partners - cooperation with suppliers and with HEIs - seem to be highly sensitive to hidden bias. This is not to say that there is an issue of unobserved heterogeneity from the perspective of either suppliers or HEIs. On the contrary, through cooperative networking, they may obtain all the necessary information about the firm. The issue of hidden bias is associated with unobserved firm

[23] We do not report the results for the subsample of innovative firms, but they are available on request. Six robust treatment effects are found: for innovative firms funded by regional support programmes - aggregate cooperation, cooperation with government agencies, and outsourcing R&D; and for innovative firms funded by national support programmes, robust treatment effects are estimated for cooperation with competitors, cooperation with government agencies, and outsourcing R&D.

characteristics, such as managerial abilities and attitudes, which are generally inaccessible to researchers.

Our general conclusion from the Rosenbaum bound approach that matching estimates of the effects of public support are likely be subject to hidden bias is consistent with Busom and Fernandez-Ribas (2008), who reach this conclusion by comparing estimates from matching and bivariate probit estimation. However, the particular estimates most subject to hidden bias are difficult to compare, because Busom and Fernandez-Ribas (2008) investigate more aggregate categories. For example, Busom and Fernandez-Ribas (2008) find pronounced hidden bias in estimates of the effectiveness of public support for private-public partnerships in general, whereas we find hidden bias for estimates of the effects on cooperation with HEIs but not for estimates of the effects on cooperation with government institutions; similarly, Busom and Fernandez-Ribas (2008) do not find hidden bias in estimates of the effectiveness of public support on vertical cooperation, whereas we find uniform evidence for hidden bias in the estimated effects on cooperation with suppliers but mixed evidence with respect to effects on cooperation with customers. We conclude that while hidden bias may be endemic in matching studies, there is no evidence that hidden bias is consistent across different studies of the effectiveness of public support on cooperation. A corollary is the usefulness of investigating the effects of public support for different types of cooperative partners separately, in which we depart from some previous studies (e.g. Fier et al. 2006; Busom and Fernandez-Ribas, 2008; Spithoven et al., 2012, p.171 and p. 181). Similar reasoning leads us also to the usefulness of investigating the effects of support from different levels of government separately (Busom and Fernandez-Ribas, 2008). A similar conclusion is advanced by Spithoven et al. (2012, p. 170), who investigated network additionality in Belgian firms and found that “there are, indeed, substantial differences in impact between different types of funding”.

Given the lack of sensitivity analysis in empirical studies, empirical evidence from matching studies should be treated with caution. The issue of unobserved heterogeneity is further exacerbated by the absence of valid instruments in available datasets (prominently the CIS data), which precludes researchers from applying other evaluation methods, not only as a robustness check but also as a way of controlling for selection on unobserved firm characteristics. In the absence of a robustness check in this context, the importance of a sensitivity analysis is even more pronounced.

Taking into account the results of sensitivity analysis, we proceed with the concluding remarks. In total, 18 treatment effects were estimated from the whole sample and the same number from the subsample of innovative SMEs. Five estimated effects in the whole sample are rather robust to selection bias; and six estimates in the subsample (perhaps due to a more homogenous sample). In total, out of 36 treatment effects, only 11 are not likely to be overestimated. Finally, across both the whole sample and the subsample of innovative firms, five ATT effects are robust to hidden bias:

- For local/regional support, three effects on the following open innovation activities - aggregate cooperation, cooperation with government institutions, and outsourcing R&D;
- For national (government) support, two effects - on horizontal cooperation and cooperation with government institutions.

Overall, we find that public support most robustly increases SME cooperation with government institutions; only slightly less robust is that the largest treatment effects of public support - both regional (a robust finding) and federal (borderline robust) - are for outsourcing R&D activities. Yet there is not so much robust evidence that public support increases cooperative and innovative behaviour more generally. Recent work on cooperation failure can help us to make sense of this contrast, suggesting that it may be of systematic rather than merely contingent significance.

By analysing treatment effects of different types of inbound open innovations, our study discriminates between the effects of public intervention on cooperation for innovation and on R&D and innovation outsourcing (extramural R&D investments and acquiring other external knowledge). Our results suggest that, depending on the source of funding, SMEs are more likely to respond to public support by increasing either their cooperation with government institutions or their investment in extramural R&D than by establishing and maintaining cooperative networks. Following our discussion in Section 2, acquiring external knowledge through cooperation could be subject to cooperation failure. In this case, compared to cooperation with other firms, either increased cooperation with government institutions may be facilitated by greater trust that these are unlikely to appropriate the firm's intellectual property; or/and R&D subcontracting is a more viable option. This issue deserves further attention from both practitioners and policy-makers. For example, to increase the effectiveness of public support for cooperation between firms – including customers and suppliers – policy makers should place particular emphasis on measures designed to attenuate cooperation failures (Zeng et al., 2010).

Empirical investigation into behavioural additionality is still in its nascent years. Our study is the first to investigate the impact of public innovation measures on open innovation practices other than cooperative behaviour. However, available data does not allow for assessing the effectiveness of public support on other categories of firms' behaviour, such as changes in managers' competencies and expertise (Busom and Fernandez-Ribas, 2008; Fier et al., 2006). Moreover, the effectiveness of public support on outbound open innovation (such as venturing or outward licensing of IPs) could also be a subject of future research. And finally, the absence of longitudinal data continues to inhibit the exploration of medium-tolong-run effects of programme participation on cooperative behaviour (Busom and Fernandez-Ribas, 2008).

References

Aakvik, A., 2001. Bounding a matching estimator: the case of a Norwegian training program. Oxford Bulletin of Economics and Statistics, 63(1), 115-143.

Alecke, B., Mitze, T., Reinkowski, J., Untiedt, G., 2012. Does Firm Size make a Difference? Analysing the Effectiveness of R&D Subsidies in East Germany. German Economic Review, 13 (2), 174-195.

Almus, M., Czarnitzki, D., 2003. The Effects of Public R&D Subsidies on Firms' Innovation Activities: The Case of Eastern Germany. Journal of Business and Economic Statistics, 21 (2), 226-236.

Austin, P.C., 2011. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behavioral Research, 46 (3), 399-424.

Barney, J. B., 1991. Firm resources and sustained competitive advantage. Journal of Management, 17 (1), 99-120.

Becker, S.O., Caliendo, M., 2007. mhbounds - Sensitivity Analysis for Average Treatment Effects. Stata Journal, 7 (1), 71-83.

Becker, W., Dietz, J., 2004. R&D cooperation and innovation activities of firms-evidence for the German manufacturing industry. Research Policy, 33 (2), 209-223.

Bianchi, M., Cavaliere, A., Chiaroni, D., Frattini, F., Chiesa, V., 2011. Organisational modes for open innovation in the bio-pharmaceutical industry: an exploratory analysis. *Technovation*, 31 (1), 22-33.

Busom, I., Fernandez-Ribas, A., 2008. The impact of firm participation in R&D programmes on R&D partnerships. Research Policy, 37 (2), 240-257.

Caliendo, M., Hujer, R., Thomsen, S.L., 2005. The employment effects of job creation schemes in Germany: a microeconometric evaluation. IZA Discussion Paper 1512.

Caliendo, M., Kopeinig, S., 2008. Some Practical Guidance for the Implementation of Propensity Score Matching. Journal of Economic Surveys, 22 (1), 31 - 72.

Cassiman, B., Veugelers, R., 2002. R&D cooperation and spillovers: Some empirical evidence from Belgium. American Economic Review, 92 (4), 1169-1184.

Chesbrough, H., 2003. Open Innovation: The New Imperative for Creating and Profiting from Technology, Harvard Business School Press, Boston, MA.

Chesbrough, H., Vanhaverbeke, W., West, J., 2006. Open innovation: Researching a New Paradigm, Oxford University Press, Oxford.

Chun, H., Mun, S.-B., 2012. Determinants of R&D cooperation in small and medium-sized enterprises. Small Business Economics, 39 (2), 419-436.

Cochran. W.G., Rubin, D.B., 1973. Controlling bias in observational studies: A review. Sankhya: The Indian Journal of Statistics, Series A, 35 (4), 417-446.

Cohen, W.M., Levithal, D.A., 1990. Absorptive Capacity: A New Perspective on Learning and Innovation. Administrative Science Quarterly, 35 (1), 128-152.

Curran, J., 2000. What is Small Business Policy in the UK for? Evaluation and Assessing Small Business Policies. International Small Business Journal, 18 (3), 36-50.

Czarnitzki, D., Ebersberger, B., Fier, A., 2007. The relationship between R&D collaboration, subsidies and R&D performance: Empirical evidence from Finland and Germany. Journal of Applied Econometrics, 22 (7), 1347-1366.

Czarnitzki, D., Hanel, P., Miguel-Rosa, J., 2011. Evaluating the impact of R&D tax credits on innovation: A microeconometric study on Canadian firms. Research Policy, 40 (2), 217-229.

Czarnitzki, D., Lopes-Bento, C., 2013. Value for money? New microeconometric evidence on public R&D grants in Flanders. Research Policy, 42 (1), 76-89.

D'Agostino, R. B. Jr., 1998. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Statistics in Medicine, 17 (19), 2265–2281.

Dahlander, L., Gann, D. M., 2010. How open is innovation? Research Policy, 39 (6), 699-709.

Dhont-Petrault, E., Pfister, E., 2011. R&D cooperation versus R&D subcontracting: empirical evidence from French survey data. Economics of Innovation and New Technology, 20(4), 309-341.

DiPrete, T. A., Gangl, M., 2004. Assessing Bias in the Estimation of Causal Effects: Rosenbaum Bounds on Matching Estimators and Instrumental Variables Estimation with Imperfect Instruments. Sociological Methodology, 34 (1), 271-310.

Doloreux, D., 2004. Regional networks of small and medium sized enterprises: evidence from the metropolitan area of Ottawa in Canada. European Planning Studies, 12(2), 173-189.

Emsley, R., Lunt, M., Pickles, A., Dunn, G., 2008. Implementing double-robust estimators of causal effects. Stata Journal, 8 (3), 334-353.

Enkel, E., Gassmann, O., Chesbrough, H., 2009. Open R&D and open innovation: exploring the phenomenon. R&D Management, 39 (4), 311-316.

Faems, D., van Looy, B. and Debackere, K. (2005): "Interorganizational Collaboration and Innovation: Toward a Portfolio Approach", Journal of Product Innovation Management, 22 (3), pp. 238-250.

Falk, R., 2007. Measuring the effects of public support schemes on firms' innovation activities: Survey evidence from Austria. Research Policy, 36 (5), 665-679.

Fier, A., Aschhoff, B., Löhlein, H., 2006. Behavioural Additionality of Public R&D Funding in Germany, in: OECD (Eds.), Government R&D Funding and Company Behaviour: Measuring Behavioural Additionality, OECD Publishing, Paris, pp. 127-149.

Georghiou, L., Clarysse, B., 2006. Behavioural Additionality of R&D Grants: Introduction and Synthesis, in: *OECD (Eds.), Government R&D Funding and Company Behaviour: Measuring Behavioural Additionality*, OECD Publishing, Paris, pp. 9-38.

Grimpe, C. ,Sofka, W. 2008. Search patterns and absorptive capacity: Low- and high-technology sectors in European countries. Research Policy, 38 (3), 495-506.

Guo, S., Fraser, M.W., 2010. Propensity Score Analysis: Statistical Methods and Applications, Sage Publications Inc.

Hewitt-Dundas, N., 2006. Resource and Capability Constraints to Innovation in Small and Large Plants. Small Business Economics, 26 (3), 257-277.

Ho, D.E., Imai, K., King, G., Stuart, E.A., 2007. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. Political Analysis, 15 (3), 199-236.

Huizingh, E.K.R.E., 2011. Open innovation: State of the art and future perspectives, Technovation, 31 (1), 2-9.

Hujer, R., Caliendo, M., Thomsen, S.L., 2004. New evidence on the effects of job creation scheme in Germany- a matching approach with threefold heterogeneity. Research in Economics, 58 (4), 257-302.

Imbens, G. W., 2004. Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review. Review of Economics and Statistics, 86 (1), 4-29.

Imbens, G.W., Wooldridge, J.M., 2009. Recent Developments in the Econometrics of Program Evaluation. Journal of Economic Literature, 47 (1), 5-86.

Keele, L., 2010. An overview of rbounds: An R package for Rosenbaum bounds sensitivity analysis with matched data", available online at http://www.personal.psu.edu/ljk20/rbounds%20vignette.pdf

Lee, M.J., Lee, S.J., 2009. Sensitivity analysis of job-training effects on reemployment for Korean women. Empirical Economics, 36 (1), 81-107.

Lee, S., Park, G., Yoon, B., Park, J., 2010. Open innovation in SMEs - An intermediated network model. Research Policy, 39 (2), 290-300.

Lee, W-S., 2013. Propensity score matching and variations on the balancing test. Empirical Economics, 44 (1), 47-80.

Leiponen, A., Byma, J., 2009. If you cannot block, you better run: Small firms, cooperative innovation, and appropriation strategies. Research Policy, 38 (9), 1478-1488.

Li, M., 2012. Using the Propensity Score Method to Estimate Causal Effects: A Review and Practical Guide. Organizational Research Methods, 00 (0), 1-39.

Lichtenthaler, U., 2008. Open innovation in practice: an analysis of strategic approaches to technology transactions. IEEE Transactions on Engineering Management, 55 (1), 148–157.

Lichtenthaler, U., 2011. Open Innovation: Past Research, Current Debates, and Future Directions. Academy of Management Perspectives, 25 (1), 75-93.

Lichtenthaler, U., Lichtenthaler, E., 2009. A Capability-Based Framework for Open Innovation: Complementing Absorptive Capacity. Journal of Management Studies, 46 (8), 1315-1338.

Madrid-Guijarro, A., Garcia, D., Van Auken, H., 2009. Barriers to Innovation among Spanish Manufacturing SMEs. Journal of Small Business Management, 47 (4), 465-488.

Mohnen, P., Hoareau, C., 2003. What type of enterprise forges close links with universities and government labs? Evidence from CIS 2. Managerial and decision economics, 24 (2-3), 133-145.

Morgan, S.L., Harding, D.J., 2006. Matching Estimators of Causal Effects: Prospects and Pitfalls in Theory and Practice. Sociological Methods & Research, 35 (1), 3-60.

Narula, R., 2001. Choosing Between Internal and Non-internal R&D Activities: Some Technological and Economic Factors. Technology Analysis & Strategic Management, 13 (3), 365-387.

Nichols, A., 2008. Eratum and discussion of propensity-score reweighting. Stata Journal, 8 (4), 532-539.

Noteboom, B., 1994. Innovation and diffusion in small firms: theory and evidence. Small Business Economics, 6(5), 327-347.

OECD, 2006. Government R&D Funding and Company Behaviour: Measuring Behavioural Additionality, OECD Publishing, Paris.

Parida, V., Westerberg, M., Frishammar, J., 2012. Inbound Open Innovation Activities in High Tech SMEs: The Impact on Innovation Performance. Journal of Small Business Management, 50 (2), 283-309.

Pearl, J., 2009. Causality: Models, Reasoning, and Inference, Cambridge University Press, New York.

Peteraf, M., 1993. The cornerstones of competitive advantage: a resource-based view. Strategic Management Journal, 14 (3), 179-191.

Rosenbaum, P.R., 2002. Observational Studies, second ed. Springer, New York.

Rosenbaum, P. R., Rubin, D., 1985. Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score. The American Statistician, 39 (1), 33- 38.
Schroll, A., Mild, A., 2012. A Critical Review of Empirical Research on Open Innovation Adoption. Journal fur Betriebswirtschaft, 62 (2), 85-118.
Sianesi, B., 2004. An evaluation of the Swedish system of active labour market programs in the 1990s. The Review of Economics and Statistics, 86 (1), 133-155.
Soh, P.-H., 2003. The role of networking alliances in information acquisition and its implications for new product performance. Journal of Business Venturing, 18 (6), 727-744.
Spithoven, A., Teirlinck, P., Frantzen, D., 2012. Managing Open Innovation: Connecting the Firm to External Knowledge, Edward Elgar Publishing.
Steiner, P.M., Cook, T.D., Shadish, W.R., Clark, M.H., 2010. The Importance of Covariate Selection in Controlling for Selection Bias in Observational Studies. Physiological Methods, 15 (3), 250-267.
Stuart, E.A., Rubin, D.B., 2008. Best Practices in Quasi–Experimental Designs: Matching Methods for Causal Inference, in: Osborne, J. (Eds.), Best Practices in Quantitative Method. SAGE Publications Inc., pp. 155-176.
Stuart, E.A., 2010. Matching methods for causal inference: A review and a look forward. Statistical Science, 25 (1), 1-21.
Van de Vrande, V. , de Jong, J. P.J., Vanhaverbeke, W., de Rochemont, M., 2009. Open innovation in SMEs: Trends, motives and management challenges. Technovation, 29 (6-7), 423-437.
Zeng, S.X., Xie, X.M., Tam, C.M., 2010. Relationship between cooperation networks and innovation performance of SMEs. Technovation, 30 (3), 181-194.

Tables

Table A.1: Variable definition, mean and standard deviation of dependent and independent variables

Variable	Variable definition	Mean	Standard deviation
FUNLOC	DV=1 if a firm received local/regional support; 0 otherwise;	0.231	0.422
FUNGMT	DV=1 if a firm received government support; 0 otherwise;	0.164	0.370
COOPERATION	DV=1 if a firm cooperates with suppliers, customers, competitors, consultants, universities and government; 0 otherwise;	0.222	0.416
COOP_CUSTOMERS	DV=1 if a firm cooperates with customers; 0 otherwise;	0.061	0.240
COOP_SUPPLIERS	DV=1 if a firm cooperates with suppliers; 0 otherwise;	0.107	0.309
COOP_COMPETITORS	DV=1 if a firm cooperates with competitors; 0 otherwise;	0.036	0.187
COOP_CONSULTANTS	DV=1 if a firm cooperates with consultants, commercial labs or private R&D institutes; 0 otherwise;	0.057	0.232
COOP_HEI	DV=1 if a firm cooperates with universities or other Higher Education Institutions (HEI); 0 otherwise;	0.070	0.255
COOP_GOVERNMENT	DV=1 if a firm cooperates with government or public research institutes; 0 otherwise;	0.088	0.284
OUTSOURCING_RD	DV=1 if a firm conducts extramural R&D; 0 otherwise;	0.247	0.431
EXTERNAL_KNOWLEDGE	DV=1 if a firm purchases or licenses patents, know -how and other types of knowledge from other firms; 0 otherwise;	0.025	0.157
SMALL_FIRMS	DV=1 if a firm has between 10 and 50 employees;	0.638	0.481
BARRIER3	Importance of too high innovation costs as a barrier to innovation (score between 0- no importance; 1 - low importance; 2 -medium importance; and 3 -high importance);	1.842	1.090
BARRIER4	Importance of lack of qualified personnel as a barrier to innovation (score between 0- no importance; 1 - low importance; 2 -medium importance; and 3 -high importance);	1.441	1.006
BARRIER7	Importance of difficulties in finding cooperative partners as a barrier to innovation (score between 0- no importance; 1 - low importance; 2 -medium importance; and 3 -high importance);	0.996	1.041

PROPAT	DV=1 if a firm applied for a patent; zero otherwise;	0.113	0.316
CONTINOUS_RD	DV=1 if a firm continuously perform R&D activities during 2004-2006; 0 otherwise;	0.345	0.475
GP	DV=1 if a firm belongs to enterprise group; zero otherwise;	0.258	0.438
EXPORT	DV=1 if a firm is exporter; zero otherwise;	0.686	0.464
INCOMING1	Importance of following sources of information: conferences, trade fairs and exhibitions (score between 0- no importance; 1 - low importance; 2 -medium importance; and 3 -high importance);	1.051	1.041
INCOMING2	Importance of following sources of information: scientific journals and publications (score between 0- no importance; 1 - low importance; 2 -medium importance; and 3 -high importance);	0.861	0.930
INCOMING3	Importance of following sources of information: professional and industry associations (score between 0- no importance; 1 - low importance; 2 -medium importance; and 3 -high importance);	0.688	0.867
INFO_INTERNAL	Importance of the information generated within the firm or enterprise group (score between 0- no importance; 1 - low importance; 2 -medium importance; and 3 -high importance);	2.135	1.006
INFO_CUSTOMERS	Importance of customers as a source of information (score between 0- no importance; 1 - low importance; 2 -medium importance; and 3 -high importance);	1.363	1.145
INFO_SUPPLIERS	Importance of suppliers as a source of information (score between 0- no importance; 1 - low importance; 2 -medium importance; and 3 -high importance);	1.541	1.102
INFO_COMPETITORS	Importance of competitors as a source of information (score between 0- no importance; 1 - low importance; 2 -medium importance; and 3 -high importance);	1.059	1.034
INFO_CONSULTANTS	Importance of consultants as a source of information (score between 0- no importance; 1 - low importance; 2 -medium importance; and 3 -high importance);	0.791	0.977
INFO_HEI	Importance of HEIs as a source of information (score between 0- no importance; 1 - low importance; 2 -medium importance; and 3 -high importance);	0.515	0.851
INFO_GOVERNMENT	Importance of government as a source of information (score between 0- no importance; 1 - low importance; 2 -medium importance; and 3 -high importance);	0.348	0.667

Table A.2. Cooperation and programme participation in regional and government support programmes (N=8,022)

	Cooperation		Outsourcing R&D		Acquisition of other external knowledge	
Regional support	Yes	No	Yes	No	Yes	No
Yes	711 (8.9 %)	1,071 (13.4 %)	809 (10.1 %)	1,045 (13.0 %)	63 (0.8 %)	1,791 (22.3 %)
No	1,143 (14.2 %)	5,097 (63.5 %)	1,171 (14.6 %)	4,997 (62.3 %)	141 (1.8 %)	6,027 (75.1 %)
Government (national) support						
Yes	493 (6.1 %)	819 (10.2 %)	570 (7.1 %)	742 (9.2 %)	54 (0.7 %)	1,258 (15.7 %)
No	1,289 (16.1 %)	5,421 (67.6 %)	1,410 (17.6 %)	5,300 (66.1 %)	150 (1.9 %)	6,560 (81.7 %)

Note: For each stream of funding and each type of open innovation (cooperation, outsourcing R&D and acquisition of other external knowledge), the percentage total constitutes 100.

Table 1: Average Treatment Effects on the Treated (ATTs) of two sources of funding

Dependent variable	Regional support			Government support		
	NN matching with Mahalanobis metric and caliper 0.02	Kernel matching (Epanechnikov kernel, bw=0.06)	IPWT estimator	NN matching with Mahalanobis metric and caliper 0.02	Kernel matching (Epanechnikov kernel, bw=0.06)	IPWT estimator
	ATT (sub-sampled SEs)	ATT (bootstrapped SEs)	ATT (robust SEs)	ATT (sub-sampled SEs)	ATT (bootstrapped SEs)	ATT (robust SEs)
Aggregate cooperation	0.157*** (0.020)	0.141*** (0.014)	0.142*** (0.014)	0.099*** (0.026)	0.085*** (0.014)	0.089*** (0.016)
Cooperation with customers	0.054*** (0.012)	0.053*** (0.009)	0.054*** (0.009)	0.030* (0.017)	0.029*** (0.010)	0.031*** (0.010)
Cooperation with suppliers	0.047*** (0.016)	0.039*** (0.010)	0.042*** (0.010)	0.040** (0.020)	0.026** (0.013)	0.028** (0.012)
Cooperation with competitors	0.025** (0.010)	0.027*** (0.007)	0.026*** (0.007)	0.045*** (0.013)	0.045*** (0.008)	0.045*** (0.008)
Cooperation with consultants	0.031** (0.012)	0.037*** (0.009)	0.037*** (0.009)	0.034** (0.017)	0.028*** (0.010)	0.030*** (0.010)
Cooperation with HEI	0.047*** (0.013)	0.042*** (0.010)	0.044*** (0.010)	0.032* (0.019)	0.047*** (0.011)	0.051*** (0.012)
Cooperation with government	0.115*** (0.014)	0.118*** (0.012)	0.117*** (0.011)	0.069*** (0.019)	0.084*** (0.012)	0.086*** (0.013)
Outsourcing R&D	0.163*** (0.020)	0.168*** (0.013)	0.167*** (0.014)	0.117*** (0.025)	0.122*** (0.013)	0.124*** (0.016)
Acquisition of other external knowledge	-0.002 (0.009)	0.007 (0.005)	0.007 (0.005)	0.013 (0.009)	0.012* (0.006)	0.012* (0.006)

__Table 2: Balancing tests after matching for NN and kernel matching__

Matching estimators	Regional support			
	Pseudo-R^2	p-value of LR test	Mean bias	t-test
NN matching without replacement and caliper	0.002	1.000	1.8	Yes
NN matching with Mahalanobis metric and caliper	0.001	1.000	0.8	Yes
Kernel matching (Epanechnikov kernel)	0.000	1.000	0.9	Yes
	Government support			
NN matching without replacement and caliper	0.004	0.999	2.0	Yes
NN matching with Mahalanobis metric and caliper	0.004	0.996	1.5	No at the 5% level of significance
Kernel matching (Epanechnikov kernel)	0.001	1.000	1.7	Yes

Table 3: Sensitivity analysis - Rosenbaum bound approach

Open innovation strategies	Regional support		Government support	
	NN without replacement and caliper 0.02	Hidden bias at 5 %	NN without replacement and caliper 0.02	Hidden bias at 5 %
	ATT (subsampled SEs)[a]	(overestimation)[b]	ATT (subsampled SEs)	(overestimation)
Aggregate cooperation	0.135*** (0.016)	No when Γ≤1.70	0.079*** (0.020)	Yes when Γ≥1.25
Cooperation with customers	0.047*** (0.010)	Yes when Γ≥1.50	0.042*** (0.014)	Yes when Γ≥1.35
Cooperation with suppliers	0.041*** (0.013)	Yes when Γ≥1.25	0.022 (0.015)	Yes when Γ≥1.00 At Γ≥1.45 changes sign
Cooperation with competitors	0.024*** (0.008)	Yes when Γ≥1.35	0.047*** (0.011)	No when Γ≤1.85
Cooperation with consultants	0.031*** (0.010)	Yes when Γ≥1.25	0.023 (0.014)	Yes when Γ≥1.05 At Γ≥1.65 changes sign
Cooperation with HEI	0.043*** (0.011)	Yes when Γ≥1.35	0.039* (0.015)	Yes when Γ≥1.15
Cooperation with government	0.110*** (0.012)	No when Γ≤2.00	0.086*** (0.016)	No when Γ≤1.65
Outsourcing R&D	0.169*** (0.017)	No when Γ≤1.90	0.112*** (0.018)	Yes when Γ≥1.45 At Γ≥1.90 changes sign
Acquisition of other external knowledge	0.007 (0.006)	Yes when Γ≥1.00 At Γ≥1.85 changes sign[c]	0.012 (0.008)	Yes when Γ≥1.00

Notes: [a] *** ATT estimated at the one per cent level of significance; ** ATT estimated at the five per cent level of significance; * ATT estimated at the ten per cent level of significance. [b] Interpretation as follows: for example, in the case of "No when Γ≤1.70", the upper bound is not significant at the 5 per cent level when Γ is below or equal to 1.7 (so Γ at the threshold level of 1.5 is not statistically significant, meaning that the reported ATT effect would be sensitive only to very high levels of unobserved heterogeneity and that, hence, we can reject the null hypothesis of no treatment effect); "Yes when Γ≥1.50" means that the upper bound becomes significant at the 5 per cent level when Γ is 1.5 (so Γ at the threshold level of 1.5 is statistically significant, meaning that the reported ATT effect would be sensitive to relatively low levels of unobserved heterogeneity and that, hence, we cannot reject the null hypothesis of no treatment effect); and "Yes when Γ≥1.25" means that the upper bound becomes significant at the 5 per cent level when Γ exceeds 1.25 (so Γ at the threshold level of 1.5 is statistically significant). [c] Interpretation as follows: In five of the 27 cases reported in Table 3 not only is the effect of unobserved heterogeneity such that an estimated positive ATT effect can never be statistically significant at the five per cent level, but also that at very high levels the estimated ATT effect may become significantly negative.