Report

Firm-level treatment effects of innovation subsidies in Flanders

Date

February 2020

Submitted by

Prof. Dr. Dirk Czarnitzki
KU Leuven
Dept. of Managerial Economics, Strategy and Innovation
Naamsestraat 69
BE-3000 Leuven
and
Centre for R&D Monitoring (ECOOM)
Naamsestraat 61
BE-3000 Leuven

Executive Summary

The purpose of this project is to apply econometric methods used for treatment effects estimation to Flemish firm level data. The main goal of the analysis is to estimate input additionalities and output additionalities of Flemish innovation subsidies granted by VLAIO (formerly IWT) until the year 2016.

Therefore, a database on Flemish subsidy recipients and other (innovative) companies has been compiled. The core data stem from the Community Innovation Surveys implemented by the Centre for R&D Monitoring (ECOOM) at KU Leuven and the subsidy applicant data provided by VLAIO. In addition, these data are supplemented by patent data from the PATSTAT database and accounting data from Bureau van Dijk's BELFIRST database. The resulting data compilation is used to estimate treatment effects of the VLAIO subsidies on innovation inputs, such as R&D employment and R&D expenditure, as well as on outputs, such as new product or process introductions, new product sales, and firm growth.

In the empirical econometric applications, several methods such as multiple regressions, matching, and difference-in-difference estimators have been considered. Eventually the difference-in-difference method turned out to produce most credible results and the statistical assumptions required for the application of this method hold. Furthermore, within the difference-in-difference applications alternatively designed control groups are be used in order to test the sensitivity of results.

The main focus of the analysis is on innovation inputs, i.e. R&D employment and R&D expenditure. The final database that can be used is a firm-level panel comprising of 12,000 observations on firm-years, where about 1,300 refer to firm-years in which the companies received a VLAIO/IWT grant.

The results concerning R&D inputs are the most reliable and these clearly reject full crowing out effects of the VLAIO R&D grant programs. Instead, granting a subsidy has a positive effect on R&D employment and R&D spending. In terms of economic magnitude the effects are high, on average. With regard to R&D employment, the average treatment effect on the treated amounts to

about 3 persons per project-year. As granted projects last on average about 22 months, roughly 5.5 R&D-person-years are created with one project.

These averages should be interpreted with some care, however, as the distribution of R&D inputs is very skewed in the Flemish economy. Most firms are small and many R&D-performing companies employ only a handful of R&D employees on average. A few companies, however, have more than 100 R&D employees. Therefore, averages are somewhat sensitive to the inclusion or exclusion of some of the larger R&D performers in the sample. Even though the results reported here seem to be stable across different model and sample specifications, some caution with interpreting the exact magnitude of the treatment effects is advised. The most important result of this exercise is the finding that the IWT/VLAIO grants are effective and lead to considerable input additionality in R&D of Flemish subsidy-receiving firms.

We also find that the result of employing three R&D employees per granted project-year holds across different sizes of firms. We distinguish small, medium and larger firms as measured by their employment figures. When applying the econometric treatment effects models, we do not find that the treatment effects differ in any statistically significant way across the firm size categories.

Furthermore, it is also found that the VLAIO project stimulate R&D cooperation among the applicants. The likelihood of collaboration within R&D projects is much higher when firms apply for subsidies. In other words, R&D collaboration would take place to a much lesser extent in Flanders if there were not VLAIO programs around. However, the econometric results also show that the subsidy awardees are not necessarily maintaining active collaborations when the subsidized projects end. After the completion of the projects, the collaboration propensity drops significantly again.

Further findings refer to the so-called output additionalities of the innovation process. We also investigated the likelihood of new product introductions to the market and new product sales as well as process innovation. We find positive treatment effects on all three indicators, i.e. without IWT/VLAIO projects Flemish firms would realize lower new product sales, less improvements to their production technology, and customers would benefit from less new products.

In summary, this report finds positive and very robust effects of IWT/VLAIO subsidies on a variety of innovation indicators of Flemish companies. In other words, it implies that in absence of VLAIO subsidy schemes, Flemish firms would invest less into innovation and would thus be in danger of losing their ability of positive long-run growth and their competitive position in the local, national and also global markets. This would not only threaten the survival of the firms and a large number of jobs in non-R&D positions at these firms, but also the survival of many innovation-intensive Flemish companies.

1 Introduction

In this report, methods for econometric treatment effects estimation are applied to Flemish firm-level data. The main goal of the empirical analysis is to investigate whether R&D grants administered by VLAIO, the Flemish public agency for innovation and entrepreneurship, have a positive effect on firms' innovation performance in terms of inputs and outputs or whether these subsidies are subject to full crowding out effects.

The question on crowding out can only be answered empirically as, on the one hand, there are good reasons for governments to subsidize R&D in the private sector. It is a common economic opinion that firms cannot appropriate all returns from knowledge-creating investments such as research and development (R&D). R&D activities create information and something as intangible as information and thus knowledge can never be kept fully secret. Thus knowledge spills over to other companies that may benefit from this information. In the case where others benefit from knowledge that has been created elsewhere, the social returns of the initial investment are higher than the private returns, i.e. the innovation projects of the initial investor create positive external effects. As the initial investor can only appropriate the private returns, the firm (or the inventor) will only invest into projects where the expected private return is higher than the private cost. Thus many innovation projects may never be implemented even though the social returns are high, because the private returns do not cover the private cost. Therefore, it is justified from a societal perspective to publicly intervene into the market for R&D and to subsidize certain R&D projects because of their positive external effects.

On the other hand, however, such policies can be subject to crowding out effects. Once subsidy policies are in place, firms may have an incentive to apply for public grants with any project for subsidies, not only those where the private cost exceed the private returns. In such a situation, it may happen that the companies do not increase their R&D budgets but just substitute their private money by public R&D grants and in the worst case there is no positive effect on private R&D efforts in the economy.

In addition to the argument on positive external effects of R&D, governmental intervention is often justified because of financial constraints for R&D investments that are known to occur especially in small and/or young high-tech firms.

In order to answer the empirical question on the crowding-out or additionality effects of Flemish innovation policy, we construct a firm-level panel database in which we have information on innovation inputs and outputs and can identify subsidy recipients. The recipients are compared to different control groups of non-subsidized firms.

The estimated effects are derived from multiple regressions (as benchmark case), matching estimators and mainly difference-in-difference regressions.

2 Policy description

The Flemish Agency for Innovation & Entrepreneurship (Vlaanderen Agentschap Innoveren & Ondernemen, VLAIO) is a governmental agency, formed in 2016 from the merger of the Agency for Entrepreneurship (Agentschap Ondernemen) and the governmental agency for Innovation by Science and Technology in Flanders (Innovatie door Wetenschap en Technologie in Vlaanderen, IWT). It provides financial support and consultation services targeting Flemish small and large companies, as well as clusters and consortia involving businesses and other nonprofit innovative organizations (e.g. universities, research centers). Its purpose is to promote entrepreneurship, stimulate growth and innovation, facilitate cooperation among firms, and foster an enterprise-friendly environment. Besides, VLAIO assists the Flemish government in the development and implementation of economic policies.

In 2018, VLAIO funded 1,400 companies for a total of € 400 million, of which around € 230 million was granted to support innovation and knowledge acquisition. According to agency regulations, a single firm can receive up to € 8 million per year to carry out its R&D projects. The financial aid for innovation is provided through 27 different subsidy measures. Among those, the most important ones are two newly introduced programs targeting research and development projects, respectively. The subsidy to research activities cover minimum 50% of the project costs. Similarly, development projects enjoy a support rate of 25%. More favorable conditions are conceded to small firms (additional 20% subsidy rate), medium businesses (10%) and partnerships (10%), which can increase the subsidy rate up to a maximum of 60% for research, and of 50% for development projects. Those two instruments substituted two other R&D programs, one broad and one specific to SME's, which were discontinued at the end of 2017. The old R&D scheme provided a base subsidy rate of 25% for development and 40% for research projects with a system of rate surcharges in favor of SME's and consortia. On the other hand, the SME-specific program granted a minimum support rate of 35% on innovation projects with a maximum budget of € 250,000.

A partial replacement for the SME R&D program is constituted by the SME growth subsidy, which can be used to acquire knowledge -defined in terms of either consultancy from an external service provider or hiring of a strategic employee- necessary to ensure competitiveness, internationalization and a steady growth trajectory. Such scheme does not directly target R&D

processes and has a broader scope that, however, can greatly benefit those innovation SME's who need to develop their absorptive capacity.

Other innovation programs managed by VLAIO include subsidies to innovation clusters, support to inter-organizational cooperation (ICON), start-up accelerators, grants for PhD and post-doc research conducted inside companies, funding of Flemish colleges and universities actively involved in practice-oriented research (TETRA). Furthermore, the agency administers also the allocation of R&D subsidies within international networks, such as EUROSTARS, EUREKA, ERA-NET, and other international networks including networks co-funded by the European Commission.

3 Brief context of literature

Quantitative analyses on the impacts of public R&D and innovation subsidies have a long tradition in Flanders. Already three decades ago, Holemans and Sleuwaegen (1988) related R&D subsidies to R&D spending at the firm level and investigated whether the receipt of subsidies increase R&D efforts of firms, or whether this funding merely replaces the private investment. In the latter case, the innovation policy would be subject to crowding out effects and thus detrimental to social welfare.

In their early study, Holemans and Sleuwaegen used the Ordinary Least Squares (OLS) to estimate R&D investment equations. They found that publicly subsidized Belgian firms invest more into R&D as a response to the receipt of subsidies. However, according to todays' methodological standards OLS regressions in this context suffer from several weaknesses as this method does not take into account that subsidies are not randomly distributed across firms.

Instead, such regressions suffer from selection bias and/or endogeneity. First, subsidized firms choose to apply for the program, i.e. they self-select and are thus often not comparable (without further adjustments) to other firms that do not apply for subsidies. Second, public authorities that administer policy schemes often follow a "picking-the-winner" strategy, i.e. the authorities have an incentive to grant subsidies to the most promising firms in order to maximize the success of their policy instruments. As a result the agencies might pick the most innovative companies for their grants which makes the receipt of subsidies an endogenous variable in regressions of innovation input or output on subsidy variables. Therefore, scholars have argued – at the latest since the early 2000s – that OLS regressions in the context of R&D subsidies are biased as they do not account for self-selection and endogeneity, respectively.

As three decades ago with the study by Holemans and Sleuwaegen, Flanders has also been active in pioneering the use of methodologies for counterfactual impact evaluations in the past. As early as 2006 the IWT published a pilot study (Aerts and Czarnitzki, 2006) where matching methods and instrumental variable regressions have been applied to Flemish firm level data stemming from the Community Innovation Surveys (CIS). These data had been merged with detailed project level data provided by the IWT. Aerts and Czarnitzki found that the IWT subsides complement the private R&D investment of the firms, or in other words that the level of innovation activity would have been significantly lower in Flanders if the policies would have not been existing. Similarly Aerts and Schmidt (2008) find comparable results, but this time by employing a conditional difference-in-difference estimator for repeated cross-sectional data. Aerts (2008) finds that subsidies do not only increase R&D spending in Flemish firms but also the R&D employment. This is by no means trivial as a concern that subsidies may only lead to higher wages of scientists but not to increased knowledge creation had been expressed in the literature.

More recently, Czarnitzki and Lopes-Bento (2013) used Matching techniques to investigate whether various policy practices are subject to crowding-out effects. For example, the IWT Flanders wondered whether funding multiple projects in the same firm is reducing the effect of the subsidy or whether funding multiple projects in a row leads to a reduction in efficacy. No such evidence could be found in the analysis. Czarnitzki and Delanote (2017) integrate innovation input and output effects of R&D subsidies into a modified Crépon–Duguet–Mairesse (CDM) model. Their results largely confirm insights of the input additionality literature, i.e. public subsidies complement private R&D investment. In addition, results point to positive output effects of both purely privately funded and subsidy-induced R&D. They do not find evidence of a premium or discount of subsidy-induced R&D in terms of its marginal contribution on new product sales when compared to purely privately financed R&D.

So far, however, the analysis of Flemish data has always been restricted to pooled cross-sections of firm-level data. The purpose of the present study is to exploit panel data for the first time in such evaluation studies for Flanders besides matching techniques or IV regressions.

4 Methods

In this report, we use different methods to identify treatment effects. The naïve benchmark case for the estimation are Pooled Ordinary Least Squares Regressions (POLS). This is the most basic and most common method used in econometrics. This method possibly overestimates the treatment effects, as it neglects

(i) that the subsidy receipt possibly depends on the outcome variables of interest. For instance, firms that generally conduct more R&D are also more likely to receive a subsidy. This means that the explanatory variables are not exogenous in the regression, and therefore a basic assumption of the statistical regression models is violated.

(ii) unobserved heterogeneity among the firms. Even though we will have a large number of firm characteristics that may determine the outcome variables of interest, there is always the remaining concern that some unobserved differences may also influence the outcomes. A common example is management quality.

In order to overcome these possible shortcomings of POLS, we also employ matching estimators and (conditional) difference-in-difference regressions.

4.1 Matching

Matching estimators have been applied and discussed by many scholars (see Smith and Todd, 2005, for an econometric overview on Matching estimators; and Zunica et al., 2014, for an overview on applications in the context of R&D subsidies).

Generally, matching estimators are used to answer the question of what treated units with a given set of characteristics would have done if they would not have received the treatment. The objective is to compare the two outcomes – when receiving and when not receiving a treatment – for the same unit. The problem is of course that we can observe at most one of these outcomes because the observed unit has either received a treatment or not. Holland (1986) refers to this as the fundamental problem of causal inference. Hence, the counterfactual situation of a treated firm (i.e. an untreated firm) is not directly observable and has to be estimated.

Our fundamental evaluation question can be illustrated by an equation describing the average treatment effect on the treated firms:

$$E\left(\alpha_{TT}\right)=E\left(Y^{T} \mid S=1\right)-E\left(Y^{C} \mid S=1\right) \tag{1}$$

where Y^T is the outcome variable. The status S refers to the group: S=1 is the treatment group and S=0 the non-treated firms. Y^C is the potential outcome which would have been realized if the treatment group (S=1) had not been treated. As previously explained, while $E(Y^T|S=1)$ is directly observable, it is not the case for the counterpart. $E(Y^C|S=1)$ has to be estimated. In the case of matching, this potential "untreated outcome" of treated firms is constructed from a control group of firms that did not receive innovation subsidies. The matching relies on the intuitively attractive idea to balance the sample of program participants and comparable non-participants. Remaining differences in the outcome variable between both groups are then attributed to the treatment.

Because of a potential selection bias due to the fact that the receipt of a subsidy is not randomly assigned, $E(Y^C|S=1) \neq E(Y^C|S=0)$ and the counterfactual situation cannot simply be estimated as average outcome of the non-participants. Rubin (1977) introduced the conditional independence assumption (CIA) to overcome this selection problem, that is, participation and potential outcome are statistically independent for individuals with the same set of exogenous characteristics X. Thus, the critical assumption using the matching approach is whether we can observe the crucial factors determining the entry into the programme. If this assumption is valid, it follows that

$$E\left(Y^{C} \mid S=1, X\right)=E\left(Y^{C} \mid S=0, X\right) \tag{2}$$

Provided that there no systematic differences in the observed characteristics between both groups, the treatment effect can be written as:

$$E\left(\alpha_{TT}\right)=E\left(Y^{T} \mid S=1, X=x\right)-E\left(Y^{C} \mid S=0, X=x\right) \tag{3}$$

In the present analysis, we conduct a nearest neighbour matching. More precisely, we pair each subsidy recipient with the single closest non-recipient. The pairs are chosen based on the similarity in the estimated probability of receiving such a subsidy, meaning the propensity score stemming from a probit estimation on the dummy indicating the receipt of subsidies *S*. Matching on the propensity score has the advantage not to run into the "curse of dimensionality" since we use only one single index as matching argument (see Rosenbaum and Rubin, 1983).

Last but not least, it is essential that there is enough overlap between the control and the treated group. We thus calculate the minimum and the maximum of the propensity scores of the potential control group, and delete observations on treated firms with probabilities larger than the maximum and smaller than the minimum in the potential control group.

Table 1: Matching protocol

Step 1	Specify and estimate a probit model to obtain the propensity score $\hat{P}(X)$.
Step 2	Restrict the sample to common support: delete all observations on treated firms with probabilities larger than the maximum and smaller than the minimum in the potential control group. (This step is also performed for other covariates that are possibly used in addition to the propensity score as matching arguments.)
Step 3	Choose one observation from the subsample of treated firms and delete it from that pool.
Step 4	Calculate the Mahalanobis distance between this firm and all non-subsidized firms in order to find the most similar control observation. $MD_{ij} = \left(Z_j - Z_i\right)' \Omega^{-1} \left(Z_j - Z_i\right)$ where Ω is the empirical covariance matrix of the matching arguments based on the sample of potential controls. If only the propensity score is used, there is no need to calculate a multidimensional distance. In that case, e.g. a Euclidian distance is sufficient.
Step 5	In this application of the matching, we restrict the group of potential neighbors to firms active in the same industry as the particular treated firm. Select the observation with the minimum distance from the remaining sample. (Do not remove the selected controls from the pool of potential controls, so that it can be used again.)
Step 6	Repeat steps 3 to 5 for all observations on subsidized firms.
Step 7	Using the matched comparison group, the average effect on the treated can thus be calculated as the mean difference of the matched samples: $$\hat{\alpha}_{TT} = \frac{1}{n^T}\left(\sum_i Y_i^T - \sum_i \widehat{Y_i^C}\right)$$ with $\widehat{Y_i^C}$ being the counterfactual for *i* and n^T is the sample size (of treated firms).
Step 8	As we perform sampling with replacement to estimate the counterfactual situation, an ordinary *t*-statistic on mean differences is biased, because it does not take the appearance of repeated observations into account. Therefore, we have to correct the standard errors in order to draw conclusions on statistical inference. We follow Lechner (2001) and calculate his estimator for an asymptotic approximation of the standard errors.

4.2 Difference-in-difference

In addition to Matching, we also employ difference-in-difference (DID) estimations which account for the panel structure of the data.

The treatment group's outcome variables of interest are observed before they participated in the program and afterwards. The control group's outcome variables is observed for the same time period. The idea of the difference-in-difference (DiD) estimator is based on exploiting this "panel structure", i.e. different firms can be traced over time. The DiD estimator works as follows: One could calculate the difference in outcomes for each observed firm over time, i.e. for both the treated firms and the control group. Suppose period t_1 is the treatment period and t_0 a year before program participation:

$$\Delta_i^T = Y_{i,t1} - Y_{i,t0}$$

$$\Delta_j^C = Y_{j,t1} - Y_{j,t0}$$

where T denotes the treatment group and C the control group, and Y is an outcome variable of interest, such as R&D employment or expenditure. One thus calculates the change of Y over time. As the change in Y may well be subject to economic shocks that concern the whole economy, one relates the change in Y of the treatment group to the change in Y of the control group. An underlying assumption is that both treated and control group would be affected by economic shocks in the same manner. Thus the treatment effect, α, can be estimated as difference in the both differences:

$$\alpha^{DiD} = E(\Delta_i^T) - E(\Delta_j^C).$$

The expected value would simply be estimated as the sample averages of the changes in Y in the treatment and control group respectively.

A test whether the treatment effect is positive in statistical terms, that is, the program increases Y in the funded firms, could simply be implemented by a two-sample t-test on mean differences in this example. In a regression context, it would simply mean that one regresses changes in Y on the treatment dummy variable, [D(TREAT)]. This would be numerically equivalent to conducting a two-sample t-test. The regression, however, allows easily for the inclusion of other control variables that could affect Y besides the treatment.

4.3 Difference-in-difference estimation with multiple time periods

Since our panel database has more than two time periods, we will not implement the difference-in-difference estimation via t-tests as described above, but rather by fixed effects "within" panel regressions. Let Y_{it} be the outcome variable of interest, e.g. R&D employment, of firm i in year t (for i = 1,…,N; and t = 1,…,T). Re-writing the equation of the previous subsection as a regression, we would obtain

$$\Delta Y_{it} = \alpha_0 S_{it} + \varepsilon_{it}$$

where S is a dummy variable and ε the commonly used i.i.d. statistical error term. In the case of two periods, the dummy S would be equal to zero for all firms in the first period, and then switch to the value 1 in the second period for the treated firms. Te coefficient α_0 would be numerically be equivalent to the α^{DiD} defined in the subsection above.

As our database has multiple time periods, however, we estimated the model not in the first-differenced form as written above, but used a so-called "within" fixed effects regression. This can be written as

$$Y_{it} = c_i + \alpha S_{it} + \varepsilon_{it}$$

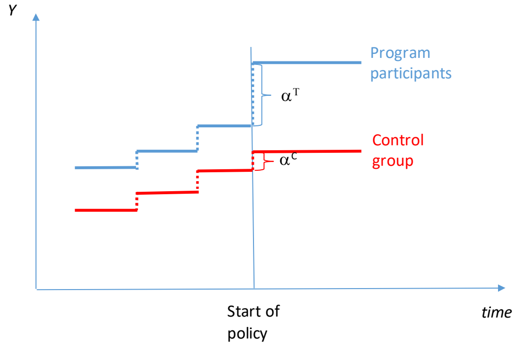
Now we would estimate a separate intercept c_i for each firm which would level out average differences in Y across firms, and the coefficient α will identify the deviation from that average during the program participation of the participating firms relative to the development of employment in the control group (those observations where S never switches to 1).

If our data would comprise again only of two time periods, all approaches would lead to exactly the same numerical parameter estimations of interest: $\alpha^{DiD} = \alpha_0 = \alpha$. As we have in fact, however, multiple periods, i.e. the year 2004 to 2016, the fixed effects within regression is the preferred implementation for a difference-in-difference approach. Of course, it is also possible to include other covariates next to the subsidy dummy variable. For more information on the implementation of standard difference-in-difference regressions, see any modern econometric textbook, such as Wooldridge (2009) or Angrist and Pischke (2009).

4.4 The common trend assumption

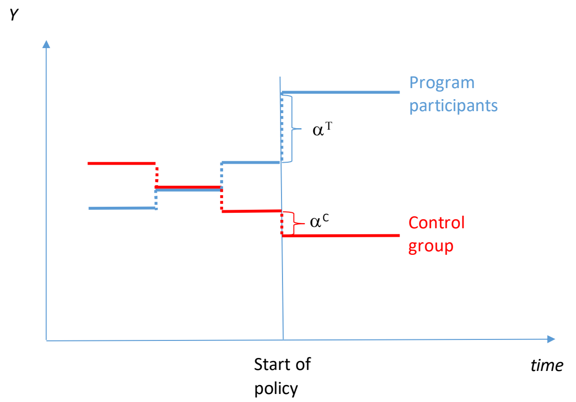
When applying the difference-in-difference methodology, it is implicitly assumed that the treatment group and the control group follow the same trend before the treatment takes place, and the identification of the treatment effects then depends on the "jump" the treatment group and the control group make at the time when the treatment group enters the program. In the figure below, the hypothetical data shows that the program participants and the control group evolve in the same way over time, however, when the participants enter the program, their increase in the dependent variable *Y*, e.g. R&D employment, is larger than that of the control group. The estimated treatment effect α^{DiD} would amount to $\alpha^{DiD} = \alpha^{T} - \alpha^{C}$. The treatment effect would not be α^{T} as this would only be a before-after comparison of the participants, but as we can see in the graph, the economy develops positively irrespectively of the program which can be seen in the positive trend in *Y* of the control group. Therefore, subtracting α^{C} from α^{T} adjusts the estimated treatment effect by netting out the general economic development and thus isolates the true program effect under the assumption that both the treatment group and the control group would have evolved similarly in the absence of the policy. This assumption becomes credible because of the common trend that can be observed before the policy had been in place.

Figure 1: A sketch of the DiD methodology with a common trend



If the common trend assumption, however, would be violated, the estimated treatment effect obtained by a difference-in-difference approach is typically not be regarded as trustworthy. Such a situation is depicted in the figure below. In this graph, the two groups of firms never exhibit the same trend and therefore the application of the difference-in-difference methodology would mistakenly result in a treatment effect that is again equal $\alpha^{DiD} = \alpha^{T}$ - α^{C}, but in this case it seems not to be a credible. *In fact,* $\alpha^{DiD} > \alpha^{T}$ as the treatment effect even gets boosted by the negative trend of α^{C}. The graph rather suggests that the two groups are not comparable at all, and thus using the control group in a standard difference-in-difference framework is an invalid approach in this case.

Figure 2: A sketch of the DiD methodology with a violation of the common trend assumption



4.5 Conditional difference-in-difference

A possible solution to the violation of the common trend assumption in the context of difference-in-difference is the combination of the methodology with so-called matching estimator. In very simple terms, this means that control group one would not use all firms that did not participate in the program, but only firms that are similar to the participant groups in some observable characteristics. For instance, the control group could be adjusted by discarding firms that are not in the same industries and regions as the treatment group. In addition, other firm level variables could be considered. The choice should be made in an economically meaningful way for each application. In our study, we will eventually have to revert to conditioning on similar pre-treatment growth rates between control and treatment group, as in this case the treatment group appears to be a very select type of firm, i.e. firms with a high desire for fast and large growth. As a "growth intention" is unobserved to every researcher who is interested in deriving treatment effects, we will have to rely on calculating pre-treatment growth rates and to pick the control group on this. As it will be shown later, the common trends across the two groups can then be restored and therefore a policy treatment effect can credibly be derived.

In order to implement the conditional difference-in-difference method, one has to find comparable firms in the control group based on observable characteristics. A common method used is the so-called propensity score matching: to find a control group consisting of firms similar to the treatment group, one specifies a Probit model where the likelihood of program participation is estimated based on observable attributes of the firms. For instance, firm size before participation, growth rates, industry affiliation and regional location etc. Rosenbaum and Rubin (1983) have shown that one can draw control observations from the pool of potential control firms based on the propensity score (=the estimated participation probability in this case). Doing so will ensure that the treatment group and the selected control group will be comparable in the covariates that were used to estimate the participation probability.

Subsequently the DiD regression is only conducted on the matched samples rather than using all potential control firms.

In summary, the CDiD methodology can be seen as an attempt to move from a situation as sketched in Figure 1~~Figure 1~~, that is, a violation of the common trend assumption, to a scenario as drawn in Figure 2~~Figure 2~~, a common trend across the treatment and control groups.

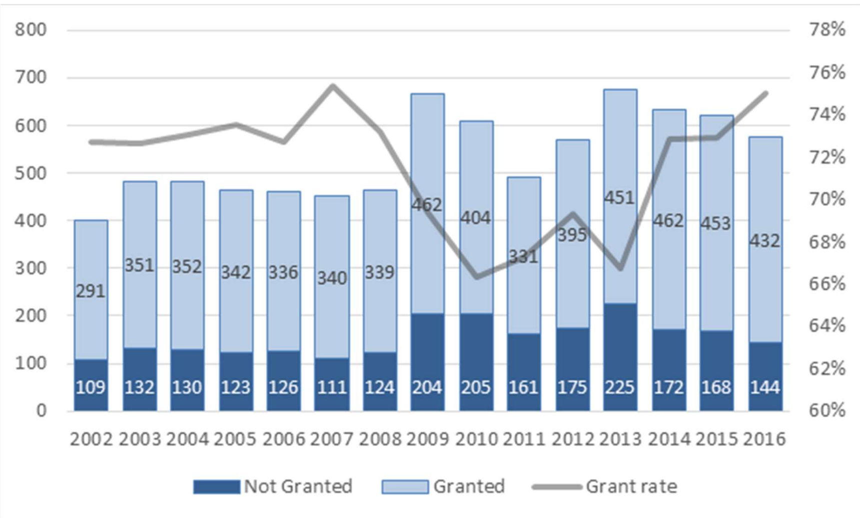
5 Data

This section illustrates the data used for the analysis. First, the program participation data are described. Successively, the procedure used to construct the final dataset is presented, followed by an overview of the main statistics of interest.

5.1 Program participation data

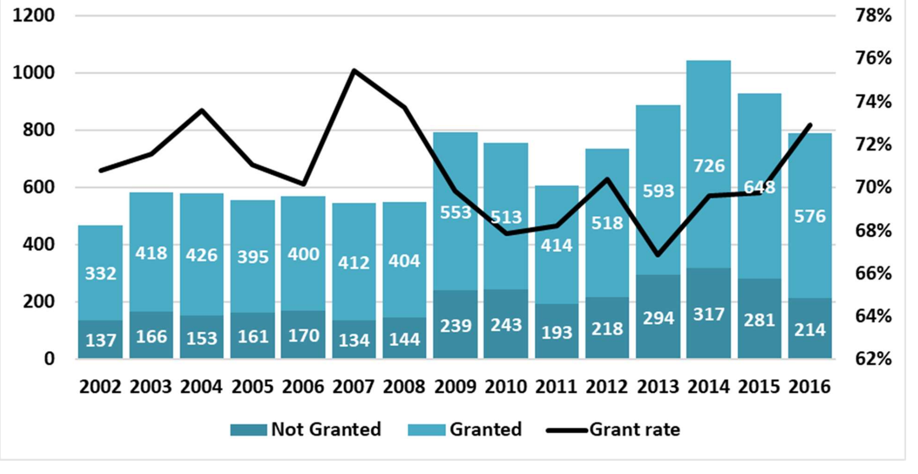
The data provide information on 10,392 firm-level project applications by 4,388 different companies over the 2002-2016 period. It is important to note here that the term "firm-level project application" refers to parts of project that are executed by firms. If one project application for an R&D grant is filed jointly by three companies and one university, we only consider the firms for the purpose of this report, and would count this application as three firm-level project applications. Figure 3~~Figure 3~~ shows the number of filed applications by consortia split into granted and not granted projects. Before 2008 the number of applications was always below 500 per year but the number increased at the time of the financial crisis in 2009 to more than 650 and remained at a higher level than before the crisis.

Figure 3: Number of project applications with firm participation per year at the consortium level



For the purpose of this report, it is more instructive to look at the number of firm-level project application, i.e. if a project application consisted of a consortium of three firms, it is counted as three firm-level applications. The number of processed firm-level applications increased from 469 in 2002 up to 1043 in 2014, then dropped to 790 in 2016 (see Figure 4~~Figure 4~~). The average yearly grant rate, i.e. the proportion of subsidized projects over total number of applications, fluctuated between 66.9% in 2013 and 75.5% in 2007.

Figure 4: Number of firm-level project applications per year



The applications refer to a large variety of VLAIO-administered support programs, which have been grouped in four main categories specified in Table 2~~Table 2~~. The program for small and medium enterprises has been the most popular over the years with a total of 5,360 firm-level project applications by 2,971 different companies, followed by R&D support schemes, cluster initiatives and Flemish-funded European programs, respectively.

Met opmaak: Lettertype: Niet Vet

Table 2: Grant programs by category

Category	Programs	Firm-level applications	Different firms
SME	KMO	5360	2971
R&D	O&O Bedrijfssteun, BAEKELAND, Innovatiemandaten, LURU, Sprint, transformationeel Geneeskundig Onderzoek	2797	1024
Cluster policy	Lichte structuur, Poeftuin bouw, Proeftuin zorg, SOC-MAAK, Speerpuntclusters, eMedia, IMEC, MIP.	1377	789
Flemish-funded European	Ambient Assisted Living, EUREKA, EUREKA-KMO, EUREKA-CELTIC, EUREKA-EURIMUS, EUREKA-PIDEA, EUREKA-MEDEA+, EUREKA-ITEA, EURIPIDES, ARTEMIS, CATRENE, EUROSTARS, Electronic Components and Systems	858	347

In Figure 5~~Figure 5~~ to Figure 8~~Figure 8~~ the firm-level application count and the grant rate are displayed for each program category. The yearly number of demands rarely exceeds 250 for all programs except for the SME subsidy scheme. However, the average grant rate vary over years between a minimum of around 60% to a maximum of 85%, save for Flemish-funded European programs whose grant rate went below 50% in the years following 2009.

Each applicant proposed nearly two projects per year, on average, while the maximum number per firm oscillates between seven in 2007 and 16 in 2015. The average number of projects per firm does not change if considering only the subsidized applicants.

In addition, 62.7% of firm-level subsidy applications (6516) concerns projects that firms intended to carry out individually. The remaining amount are joint applications whereby two or more firms decided to apply for a grant on the same project.

Figure 5: Number of firm-level project applications per year, R&D program

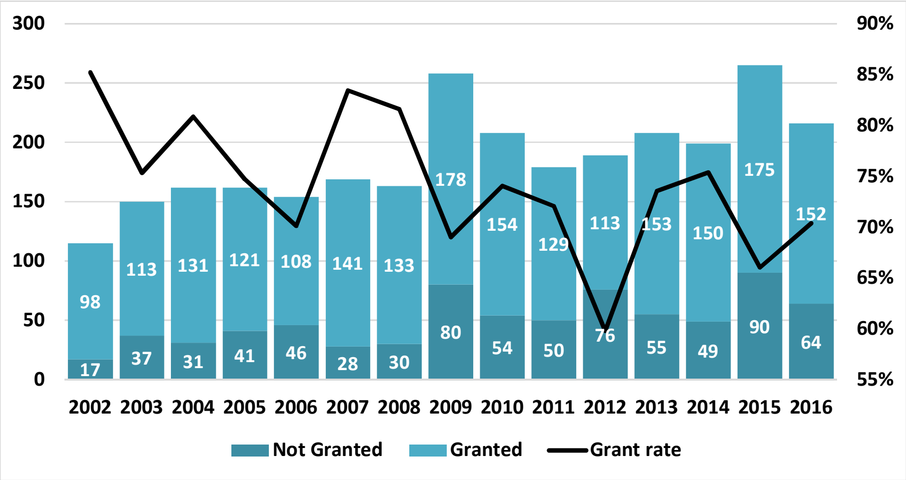


Figure 6: Number of firm-level project applications per year, SME program

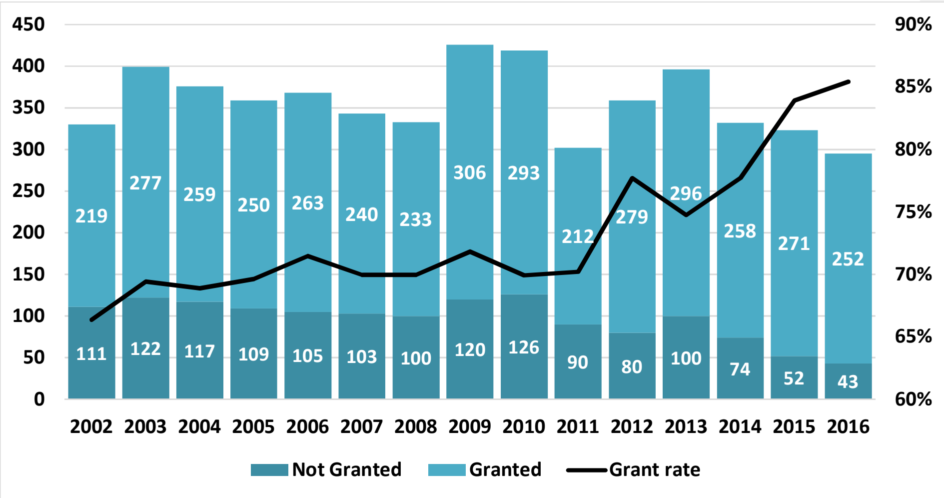


Figure 7: Number of firm-level project applications per year, Flemish-funded European programs

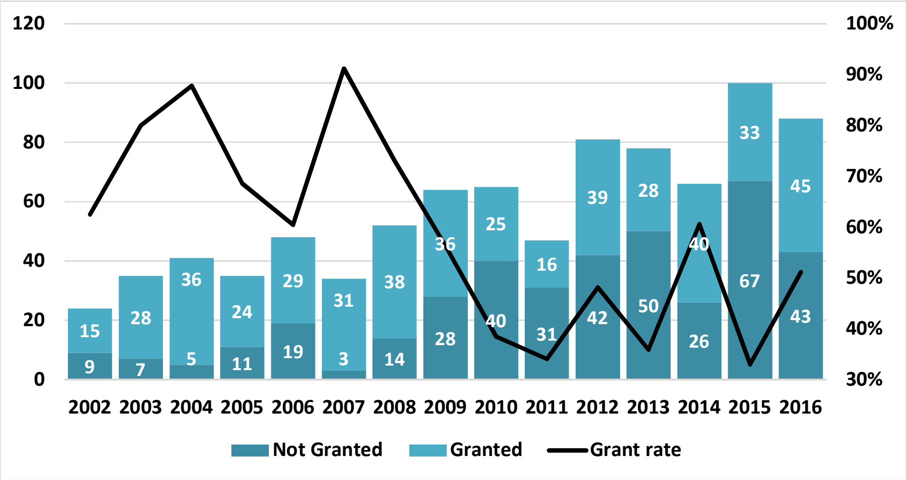
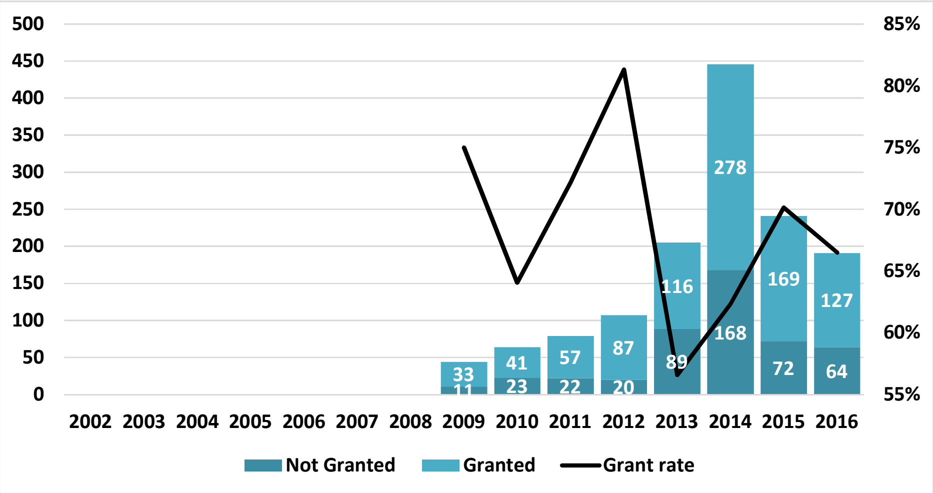


Figure 8: Number of firm-level project applications per year, Cluster programs



On average the applied budget is 537.4 thousand € (median: 230.8), while the mean revised budget is 511.2 thousand € (median: 224.2). The average subsidy is 213.8 thousand € per project (median:

98.2). Out of all program categories, R&D initiatives display the largest amounts and the highest variability in terms of standard deviation (see **Table 3**~~Table 3~~).

Table 3: Descriptive statistics of applied, accepted and subsidized budget, overall and by program categories (in thousands €)

		Applied budget	Accepted budget	Subsidy
Overall	Mean	537.4	511.2	213.8
	Standard Dev.	1231.6	1023.9	377.0
	Median	230.8	224.2	98.2
R&D	Mean	1244.5	1179.0	478.1
	Standard Dev.	2093.6	1659.1	591.1
	Median	560.7	617.9	256.6
SME	Mean	193.3	175.5	74.2
	Standard Dev.	201.9	177.7	71.0
	Median	104.0	100.0	45.5
Flemish-funded	Mean	915.7	931.0	385.2
European programs	Standard Dev.	1145.6	935.3	349.9
	Median	564.0	634.9	280.0
Cluster policy	Mean	235.3	243.8	137.6
	Standard Dev.	260.0	258.1	150.1
	Median	150.6	152.7	86.1

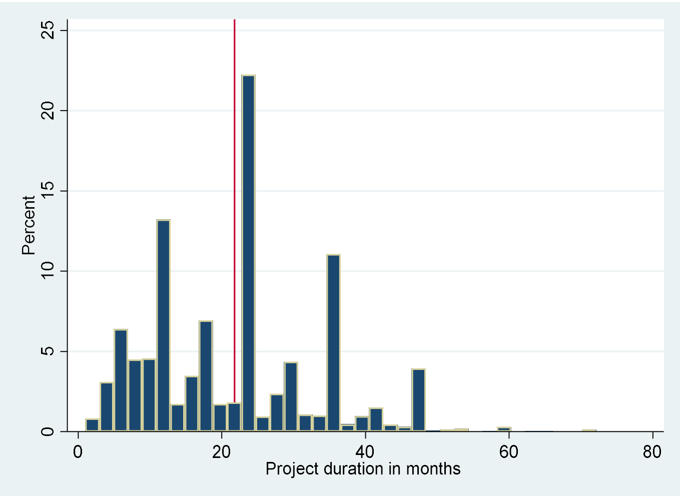
As to the average subsidy rate, i.e. the ratio between the accorded subsidy and the revised budget for that project, it is 47 percent overall (Table 4~~Table 4~~). By program, the mean stays almost the same, except for cluster initiatives, whereby the percentage is nearly 10 percent points higher (54).

Table 4: Descriptive statistics of subsidy rate in percentage, overall and by program

	Mean	Standard dev.	1st Quartile	Median	3rd Quartile
Overall	47.04	13.23	37.50	47.79	55.00
R&D	43.46	14.73	35.00	45.00	50.00
SME	47.42	10.09	41.52	45.00	53.53
Flemish-funded European	44.38	15.78	35.00	47.00	55.00
Cluster	54.14	16.03	46.04	55.00	60.00

Another important characteristic of the projects is their duration. This will become important for the interpretation of the estimated treatment effect in the subsequent econometric study. As can be seen in Figure 9~~Figure 9~~, most common project durations are 12, 24 and 36 months, but there is quite some variation in the duration among projects. The average duration amounts to 21.8 months.

Figure 9: Project duration



5.2 Merged dataset

For the purpose of our analysis, data on characteristics, performance and R&D activities have been added from the Belfirst database and from multiple waves of the Community Innovation survey (CIS). The new data include information concerning not only firms who have been participating in the VLAIO programs, but also unrelated firms that constitute the comparison group through which the effectiveness of the program can be assessed.

The new information was merged with VLAIO data by using the VAT number of firms as matching variable. In the process, the original data were reshaped from project level to firm level. Moreover, as the CIS is administered biyearly, the final dataset features only the even years from 2004 to 2016.

The final dataset features 8,219 firms who never took part in the VLAIO programs for a total of 15,056 more observations (see Table 5~~Table 5~~). On the other hand, after the merging process information of 1,740 VLAIO participants could be retained. If considering only those companies that are observed for more than one year, then the number drops to 1,130 participants and 3,495 non-participants.

Table 5: Number of firms and observations in the merged dataset

	VLAIO				Merged		
No. Firms	Recipients	Non-recipients	Total	% Recipients	Non-VLAIO	Total	% VLAIO
All	1429	311	1740	82.2	8232	9972	17.4
>1 obs/year	966	164	1130	85.5	3495	4625	24.4
No. Observations							
All	4034	671	4705	85.7	15056	19761	23.8
>1 obs/year	3571	524	4095	87.2	10319	14414	28.4

The additional data provide information about several firm characteristics, namely age, size, turnover, ownership, exporting, assets, cash and debt. As to firms' R&D activities, information is provided regarding expenditures for in-house R&D, size of R&D personnel, cooperation, innovation performance and commercialization (see Table 6~~Table 6~~).

Table 6: Descriptive statistics of the variables in the merged dataset

	VLAIO Grant recipients		VLAIO Non-recipients		Non-VLAIO firms	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Characteristics						
Age	27.69	20.11	27.02	20.13	26.79	17.16
Employees	241.40	808.07	136.20	220.28	74.74	197.84
D. Group (%)	62.46	48.43	61.25	48.75	52.81	49.92
D. Foreign (%)	26.00	43.87	23.55	42.46	25.91	43.82
D. Export (%)	72.93	44.44	65.54	47.60	53.46	49.88
Assets (m€)	98.79	685.48	28.46	96.94	15.88	266.71
Cash (m€)	3.70	19.98	2.02	7.72	1.72	13.69
Debt (m€)	78.91	484.66	34.99	108.90	16.45	218.94
R&D activity						
Int. Expense (m€)	378.20	7144.3	22.31	156.24	17.77	446.68
Employees (hc)	23.21	89.67	5.21	21.28	2.02	14.05
D. Cooperat. (%)	64.78	47.77	43.46	49.61	20.00	40.00
D. Innovator (%)	86.85	33.78	75.30	43.16	45.42	49.79
D. New mkt (%)	55.57	49.70	38.81	48.78	17.16	37.71
Sales newmkt (%)	6.07	15.96	4.29	14.04	1.79	8.77
No. firms	1429		311		8232	

Note: the prefix "D." denotes dummy variable which take the values 0 or 1.

A first look on the descriptive statistics evinces how on average characteristics of subsidized firms are different from the other two groups of rejected and non-applicant firms. The most striking differences regard the size (nearly 241 employees versus 136 and 75), turnover (11.9 million € vs. 6.2 and 3.8), percentage of sales from exports (29.2 vs. 20.3 and 16.6), assets (98.8 million € vs.

28.5 and 15.9), cash (3.7 million € vs. 2 and 1.7), debt (78.9 million € vs. 35 and 16.5), internal R&D expenditures (378.2 million € vs. 22.3 and 17.8), size of R&D staff (23 employees versus 5 and 2), frequency of cooperation (64.8 percent vs. 43.5 and 20), as well as likelihood of developing a completely new product or service (55.6 percent versus 38.8 and 17.2).

Those patterns, however, may reflect the beneficial effect of the program on recipients compared to the non-supported groups. For this reason, some of the firm characteristics are analyzed in more detail considering the subsidized firms only in the period before receiving the first grant from VLAIO.

6 Econometric Analysis

In this chapter, we present the econometric estimation results. We first walk the reader through the example on R&D employment as outcome variable in detail. R&D intensity in terms of R&D expenditure is then considered subsequently as alternative innovation input variable as robustness test. Other analyses on innovation output variables are considered as supplemental analysis as these measures are not affected by the subsidy directly but only indirectly through the innovation inputs of the firm.

6.1 R&D employment

6.1.1 R&D employment intensity - Full sample

The first outcome variable that is considered is the R&D employment intensity of the firm measured as R&D employment divided by total employment. We apply a Pooled OLS regression, Matching and Difference-in-Difference regressions. First, all firms in the Community Innovation Survey are used. Subsequently, we experimented with using only innovators and finally only using firms in the control group that have at least had one R&D employee in any of the years in the survey.

The analysis starts with a "naïve" benchmark. In that naïve benchmark, an OLS regression is run only with a "treatment" dummy variable that indicated that a certain firm got a VLAIO grant in a certain year. In addition, the "post treat" dummy indicated that the firm has been a successful VLAIO awardee before but has no longer an ongoing subsidized project. This is done to make sure that a formerly subsidized firm is not mistakenly put into the control group of non-subsidized firms. If the latter would be done, the treatment effect could be underestimated if there is a "memory effect', i.e. the firms keep the R&D employment or investment at higher levels after the subsidy ends when compared to the time period before they got a subsidy. Furthermore, the coefficient of the post-treatment dummy variable shows whether the recipient firms reduce their R&D inputs again after the period of being subsidized ended.

In the"naïve" benchmark regression using the method of POLS without further control variables, we find that the R&D employment intensity is 17.6%-points higher for treated firms when compared to the control group. This effect, however, is confounded by firm heterogeneity, i.e.

systematic differences among subsidy recipients and other firms. Once, covariates are taken into account, the estimated treatment effect reduces 11%-points. The covariates that are taken into account are firm size, economic sectors, years, the prior experience with the VLAIO system as measured by the prior application stock per employee that the firms had before the corresponding time period (and its squared value to control for possible non-linear effects), the patent application stock per employee prior to the corresponding year (and its squared value), an export dummy as exporters might be more R&D-intensive, a dummy indicating whether a firm belongs to a group and another dummy variable indicating whether the group's parent company is a foreign firm, and balance sheet information, i.e. the capital intensity of the firm (=total assets per employee), the debt ratio (= debt/total assets) and cash-flow per employee.

The variable POST-TREAT indicates that the treatment effect of 11% shrinks by 5% once the subsidy period ends in the POLS model. This means that the firms do partly reduce their R&D inputs again once they do not longer receive public resources from VLAIO.

An alternative to the multiple regression method is the matching estimator, where for each treated firm-year observation and control observation with most similar characteristics is drawn from the control group in order to make the treatment group and the control group even more comparable than the regression does by controlling parametrically for the set of different covariates. The results are very similar to the multiple POLS regression (not displayed in detail).

As a further method, the difference-in-difference estimator is applied. Now the treatment effect is not only estimated as difference between the treatment and control group but as change in R&D employment intensity for treated firm over time relative to the control group. This means the identified treatment effect is the change in R&D employment intensity as response to a subsidy in the receiving firms compared to the change in R&D employment intensity in the control group in the same time period. This method is thus a much more rigorous approach in statistical terms, as it controls for permanent, unobserved differences between the firms that cannot be observed by the researcher, such as (R&D) management quality.

When applying the DiD method, the estimated treatment effect reduces to 2.9%-points (column 3 in Table 7~~Table 7~~). This is a large difference to 11% as estimated by the POLS approach and highlights the importance of unobserved heterogeneity among treated and non-treated firms.

A test also shows that the common trend among the treatment and the control group is not rejected in the DiD application (column 4 in Table 7~~Table 7~~).

In order to interpret this finding in terms of its economic significance, one can put the estimated treatment effect into perspective. The VLAIO subsidy recipients show an R&D employment intensity of 8.8% before they win an R&D grant, on average. A firm with 100 employees in total, had thus about 9 R&D employees. As a response to the subsidy, the intensity increases by 2.9% points, i.e. a firm with 100 employees would hire almost 3 R&D employees (in headcounts, not necessarily full-time equivalents), all else constant.

On average, the firms that have a granted, ongoing project in any given receive about € 121,000 per year. This means per headcount, about € 40,000 are necessary to create a workplace. According to the R&D survey 2012, an average R&D workplace per year did cost € 76,000 in Flanders. This means that the cost per R&D workplace are roughly split half between the VLAIO subsidy and additional private funds that the awardee firm invests.

As the post-treatment effect amounts to -0.013 and is statistically significant, the DiD models also show that after the subsidy period ends, the firms would on average reduce their R&D employment again by about one R&D employee.

Table 7: Regression results on R&D employment intensity – full sample

VARIABLES	(1) POLS no controls	(2) POLS with controls	(3) DID	(4) DID common trend
TREAT	0.176***	0.110***	0.029***	0.032***
	(0.007)	(0.006)	(0.006)	(0.007)
POST-TREAT	-0.085***	-0.050***	-0.013**	-0.013**
	(0.011)	(0.008)	(0.005)	(0.005)
PRE-TREAT				0.006
				(0.007)
Ln(employment)		-0.044***	-0.138***	-0.138***
		(0.002)	(0.007)	(0.007)
Application stock/employee		0.748***	-0.092	-0.092
		(0.076)	(0.080)	(0.080)
(Application stock/employee)^2		-0.476***	0.007	0.007
		(0.062)	(0.047)	(0.047)
Patent stock/employee		0.012***	-0.005***	-0.005***
		(0.001)	(0.002)	(0.002)
(Patent stock/employee)^2		-0.000***	0.000**	0.000**
		(0.000)	(0.000)	(0.000)
Export dummy		0.020***	0.010***	0.010***
		(0.003)	(0.002)	(0.002)
Group member dummy		0.042***	0.007**	0.007**
		(0.003)	(0.003)	(0.003)
Foreign parent dummy		0.019***	-0.000	-0.000
		(0.003)	(0.005)	(0.005)
Assets/employee		0.000***	-0.000**	-0.000**
		(0.000)	(0.000)	(0.000)
Debt/assets		-0.001***	-0.000**	-0.000**
		(0.000)	(0.000)	(0.000)
Cash flow/employee		-0.000	-0.000***	-0.000***
		(0.000)	(0.000)	(0.000)
Observations	12,010	12,010	12,010	12,010
R-squared	0.129	0.416	0.357	0.357

Robust standard errors in parentheses; sample: 3958 different firms; all regressions include a full set of time dummies; POLS regressions also include industry dummies.
*** p<0.01, ** p<0.05, * p<0.1

The control variables have the expected signs. Larger firms have a lower R&D employment intensity. The more experience the firms have with the VLAIO funding system in the past, the higher is their current R&D intensity. Similarly, firms with a higher patent stock, i.e. past successful R&D activities that led to inventions, are also engaging more into R&D currently. Exporting companies are showing a higher R&D intensity that non-exporting firms, and firms that belong to a group do also invest more into R&D as measured by the R&D employment intensity than stand-alone companies. This could point to a resource effect, as group members may benefit financially from the consortium, or the consortium can realize higher economies of scope from R&D than stand-alone firms. The results on foreign parents within firm consortia is more mixed. While the effect is positive in the POLS regressions, it disappears in the DiD regressions. Similarly, there is no clear conclusion regarding the capital intensity and cash flow. The signs turn around between POLS and DID.

6.1.2 R&D employment intensity: control group reduced to R&D performers

Even though the control group in the previous analysis does not violate the common trend assumption that is underlying the DiD approach, it is interesting to see how the estimated treatment effects change if a more rigorous control group is used. In this case, the control group is reduced to firms that at least once perform R&D in the observed time period, i.e. they have at least 1 R&D employees once in the panel.

The results show that the sample reduces from about 12,000 observations to 7,489 firms that have at least once performed R&D in the sample between 2004 and 2016. This control group might thus constitute a more credible sample for estimating the relevant counterfactual of what the R&D intensity of the treated firms would have been if they had not gotten a subsidy.

The results are remarkably robust when compared to the full sample. The treatment effect derived from the DiD still amount to 2.7%-points and does thus only reduce by 0.2%-points when compared to the full sample (see column 3). Thus, a firm with 100 employees hires almost 3 R&D employees as a response to the subsidy receipt.

This conclusion is not rejected by the test on common trends (column 4). The pre-treatment variable is not significant in the regression.

Table 8: Regression results on R&D employment intensity – sample of R&D performers

VARIABLES	(1) POLS no controls	(2) POLS with controls	(3) DID	(4) DID common trend
TREAT	0.152***	0.102***	0.027***	0.028***
	(0.008)	(0.006)	(0.006)	(0.007)
POST-TREAT	-0.088***	-0.052***	-0.012**	-0.012**
	(0.012)	(0.008)	(0.006)	(0.006)
PRE-TREAT				0.002
				(0.008)
Ln(employment)		-0.068***	-0.166***	-0.166***
		(0.003)	(0.009)	(0.009)
Application stock/employee		0.452***	-0.159**	-0.159**
		(0.073)	(0.078)	(0.078)
(Application stock/employee)^2		-0.297***	0.036	0.036
		(0.051)	(0.045)	(0.045)
Patent stock/employee		0.010***	-0.005**	-0.005**
		(0.001)	(0.002)	(0.002)
(Patent stock/employee)^2		-0.000***	0.000**	0.000**
		(0.000)	(0.000)	(0.000)
Export dummy		0.015***	0.013***	0.013***
		(0.004)	(0.004)	(0.004)
Group member dummy		0.053***	0.011**	0.011**
		(0.004)	(0.004)	(0.004)
Foreign parent dummy		0.035***	-0.001	-0.001
		(0.004)	(0.006)	(0.006)
Assets/employee		0.000***	-0.000*	-0.000*
		(0.000)	(0.000)	(0.000)
Debt/assets		-0.001***	-0.000	-0.000
		(0.000)	(0.000)	(0.000)
Cash flow/employee		0.000	-0.000***	-0.000***
		(0.000)	(0.000)	(0.000)
Observations	7,489	7,489	7,489	7,489
R-squared	0.091	0.480	0.435	0.435

Robust standard errors in parentheses; sample: 2259 different firms; all regressions include a full set of time dummies; POLS regressions also include industry dummies.
*** p<0.01, ** p<0.05, * p<0.1

6.1.3 R&D employment in levels (headcount); control group: R&D performers

This subsection presents a robustness check where R&D employment is used as dependent variable, i.e. the headcount of R&D personnel. The advantage of this specification is that one can directly read the treatment effect from the estimated coefficient in the regression table. A disadvantage is that such a specification in levels is likely to be sensitive to a few large values in the dependent variable. Most firms have very few R&D employees, but some have more than thousand. This implies that the average of the R&D employment distribution may be very affected by the few large observations. In order to mitigate this problem to some extent, we have excluded the largest 2% of the observed R&D employment numbers. We experimented with different exclusion rules and the estimated effects behave as expected. The more large numbers are excluded the more does the estimated average treatment effect reduce. The regressions below are performed with an R&D distribution having an average value of 7.42 and a maximum of 138 R&D employees. In the original sample the maximum is above 2,000 R&D employees.

As the table below shows the treatment effects in the POLS regressions are implausibly high with values around 14 employees. The DID regressions, however, account for the average values at the firm level over time and identify changes in the firm-specific R&D labor force. These deliver credible treatment effects. The estimated average treatment effect on the treated amounts to about 3.2 R&D employees as response to receiving a VLAIO grant.

It should be noted, however, that these results are still somewhat sensitive to the exclusion of the largest numbers in the R&D employment distribution. If the largest 52 numbers (equal to 1% of the positive observations on R&D employment) are dropped from the sample, the estimated treatment effect in column (3) drops to 2.45. Even though the effects are somewhat sensitive to the inclusion or exclusion of large numbers, always positive and statistically significant treatment effects are found. The estimates are thus very robust also in comparison to the specifications discussed above where R&D employment intensity has been used instead of headcounts.

Table 9: Regression results on R&D employment (headcounts) – sample of R&D performers

VARIABLES	(1) POLS no controls	(2) POLS with controls	(3) DID	(4) DID common trend
TREAT	14.653***	13.458***	3.176***	3.596***
	(0.795)	(0.838)	(0.692)	(0.839)
POST-TREAT	-8.170***	-7.454***	-0.303	-0.263
	(1.107)	(1.015)	(0.745)	(0.741)
PRE-TREAT				0.882
				(0.758)
Ln(employment)		2.238***	-2.092***	-2.088***
		(0.280)	(0.724)	(0.724)
Application stock/employee		-24.830***	-25.171***	-25.101***
		(4.811)	(5.174)	(5.159)
(Application stock/employee)^2		7.789***	13.089***	13.047***
		(2.985)	(3.037)	(3.033)
Patent stock/employee		0.906***	-0.527**	-0.524**
		(0.127)	(0.230)	(0.230)
(Patent stock/employee)^2		-0.018***	0.009**	0.009**
		(0.003)	(0.004)	(0.004)
Export dummy		0.720*	0.772**	0.764**
		(0.408)	(0.328)	(0.329)
Group member dummy		3.118***	0.364	0.363
		(0.373)	(0.314)	(0.314)
Foreign parent dummy		1.419**	0.219	0.230
		(0.555)	(0.777)	(0.776)
Assets/employee		0.009***	-0.001	-0.001
		(0.002)	(0.004)	(0.004)
Debt/assets		-0.006	-0.008	-0.008
		(0.013)	(0.015)	(0.015)
Cash flow/employee		-0.002	-0.012**	-0.012**
		(0.005)	(0.006)	(0.006)
Observations	7,489	7,489	7,489	7,489
R-squared	0.111	0.232	0.049	0.050

Robust standard errors in parentheses; sample: 2259 different firms; all regressions include a full set of time dummies; POLS regressions also include industry dummies.

*** p<0.01, ** p<0.05, * p<0.1

6.1.4 Firm size differences in R&D employment effects

(measured as headcount in levels; control group: all firms)

In this final analysis on R&D employment we again use the R&D headcounts as dependent variable. Thus we can directly read the treatment effects from the regression coefficients displayed in the table below. We no consider different firm size categories as it is often suspected that small firms benefit more from R&D subsidies than larger firms, i.e. smaller firms react more to a subsidy receipt than larger firms.

In order to conduct this analysis we have to split our sample into firm size categories here. Rather than applying a European standard definition of firm size categories, we generate a custom definition of firm size categories that fit our data best. Therefore we define three different size categories:

- Small firms: their total employment is on average throughout the observed time period lower than 25 employees;
- Medium-sized firms: their total employment is larger or equal to 25 employees but less than 75 employees on average in the observed time periods;
- Larger companies have on average more than 75 employees during the observed time periods of our panel.

We use the 25 and 75 threshold values to define the three firm size categories as these split our sample of observations nicely into groups that consist of roughly one third of the data. According to these firm size definitions, 33% of our firm-year observations belong to small firms, 33% refer to medium-sized firms and 34% are larger companies.

In order to now derive treatment effects for each category separately, we apply the different econometric models using the full sample of firms, but estimate three treatment effects, one for each category: TREAT_x_SMALL FIRM, TREAT_x_MEDIUM FIRM, TREAT_x_LARGER FIRM.

Table 10: Regression results on R&D employment by firm-size class (headcounts) – full sample

VARIABLES	(1) POLS no controls	(2) POLS with controls	(3) DID	(4) DID common trend
TREAT_x_SMALL FIRM	4.961***	1.190	2.560***	3.035***
	(0.444)	(0.745)	(0.573)	(0.740)
TREAT_x_MEDIUM FIRM	7.957***	5.272***	2.927***	3.340***
	(0.611)	(0.582)	(0.859)	(0.958)
TREAT_x_LARGER FIRM	26.941***	22.672***	3.572***	3.935***
	(1.227)	(1.235)	(1.098)	(1.157)
POST-TREAT	-7.837***	-6.593***	-0.267	-0.234
	(1.010)	(0.951)	(0.714)	(0.710)
PRE-TREAT				0.854
				(0.717)
Ln(employment)		0.963***	-1.800***	-1.800***
		(0.201)	(0.597)	(0.598)
Application stock/employee		17.823***	-23.626***	-23.679***
		(4.448)	(4.990)	(4.974)
(Application stock/employee)^2		-17.411***	12.193***	12.211***
		(4.355)	(2.932)	(2.928)
Patent stock/employee		0.945***	-0.516**	-0.514**
		(0.109)	(0.219)	(0.219)
(Patent stock/employee)^2		-0.018***	0.009**	0.009**
		(0.002)	(0.004)	(0.004)
Export dummy		0.815***	0.536***	0.531**
		(0.224)	(0.207)	(0.207)
Group member dummy		1.741***	0.295	0.294
		(0.240)	(0.217)	(0.217)
Foreign parent dummy		0.718*	0.173	0.180
		(0.367)	(0.543)	(0.542)
Assets/employee		0.005***	-0.001	-0.001
		(0.001)	(0.003)	(0.003)
Debt/assets		-0.019**	-0.005	-0.005
		(0.008)	(0.008)	(0.008)
Cash flow/employee		0.000	-0.009**	-0.009**
		(0.003)	(0.004)	(0.004)
F-Test on treatment effect heterogeneity across firm size categories	**F(2, 12005) = 182.27*** **	**F(2,11977) = 113.15*** **	**F(2,3957) = 0.37**	**F(2,3957) = 0.29**
Observations	12,010	12,010	12,010	12,010
R-squared	0.230	0.295	0.040	0.040

Robust standard errors in parentheses; sample: 3958 different firms; all regressions include a full set of time dummies; POLS regressions also include industry dummies.

*** p<0.01, ** p<0.05, * p<0.1

As before the POLS models show non-credible results – the estimated treatment effects are much too high. The DID model, however, shows again more reasonable results that are in line with previous findings. In addition, we see that the estimated treatment effects are quite stable across the firm size categories: 2.56 for small firms, 2.93 for medium-sized firms and 3.57 for larger firms. While these numbers suggest that the treatment effects are slightly higher the larger the firm size category is, a statistical test on equality of these estimated does not reject the hypothesis that the coefficients are equal (the F-Test is displayed at the bottom of the table).

We thus conclude that there are not statistically significant treatment effects across the firm size classes.

6.2 Estimations on behavioral additionality and innovation output

In this final section on the empirical results, we perform regressions on other innovation characteristics of the firm. Three different variables are considered:

1. A dummy variable indicating whether a firm collaborates within its innovation projects with other partners; for example, other firms such as suppliers and customers or consulting companies or public research institutions such as universities.
   Collaboration is often used as measure on "behavioral additionality" in the context of public subsidy studies, as grant programs are often designed (or give preferential treatments) to applications of consortia. The reason is that it is expected that the collaborating partners realize knowledge spillovers effects among each other.
2. Furthermore, an innovation outcome variable is considered. The most standard variable in the context of innovation surveys is an indicator variable on whether the firm as introduced a new product or service to the market (product innovation) or has implemented a new technology in its production process (process innovation).
3. As further innovation outcome variable, new product sales can be considered. A standard variable in the context of the Community Innovation Survey is the question on the share of market novelties in the firms' total sales. This variable describes the degree of innovativeness of a firm's product portfolio and at the same time reflects the success of its product innovations on the market.

Below we only display the DiD regression as they have turned out to be the most reliable estimates among the methods applied in this study. We show two versions of each regression. First the full sample of firms is considered, and second the control group only consisting of R&D-performing companies is used.

On average, 34% of firms collaborate within their innovation projects. If only R&D performers are considered this share increase to 49%. At the DID regressions show, we find positive behavioral additionality of the VLAIO grants. On average, the likelihood to collaborate increases by 18%-points as response to the receipt of a VLAIO project (see columns 1 and 2).

The likelihood to realize a product of process innovation amounts to 57% in the full sample and to 77% in the sample of R&D performers. This probability increases by 6.4%-points or 8%-points, respectively, if firms obtain a VLAIO project (see columns 3 and 4).

The share of sales with market novelties amounts to 2.8% in the full sample and 4.3% in the subsample of R&D performers. These shares increase by 2.1%-points and by 1.5%-points, on average, when the firm is awarded a VLAIO grant. These results should be interpreted with care though because of two reasons: the effect of the grant is very indirect as the grant first and foremost increases innovation inputs at the firm-level, and only these higher inputs may translate into successfully completed R&D projects, product development and subsequent market introductions of new products. Here we estimated the effect of the subsidies directly on the product success. A second reason of concern is the time lag between R&D that the firms conduct and the time when a new product may reach the market. In addition these time lags may be very heterogeneous among industries or technologies. For instance, the innovation cycle in the ICT industry can be expected to be much shorter than in the pharmaceutical industry. Introducing a new app to the market may only take a few months from the original idea. In contrast, introducing a new drug may take up to 15 or 20 years after the original discovery of a new active ingredient. Unfortunately, the time series structure of our data is not rich enough to account for heterogeneous development times of inventions and to experiment with different lag structures. Therefore, the results on the innovation outcomes should be interpreted very cautiously.

Table 11: Behavioral additionality and innovation output: DID regressions

VARIABLES	(1) Collaboration; full sample	(2) Collaboration; RD perform.	(3) Innovation; full sample	(4) Innovation; RD performers	(5) New prod. Sales; full sample	(6) New prod. sales; RD perform.
TREAT	0.176***	0.183***	0.064***	0.080***	2.106***	1.536**
	(0.029)	(0.030)	(0.021)	(0.021)	(0.698)	(0.738)
POST-TREAT	-0.161***	-0.166***	-0.078***	-0.067***	0.771	0.339
	(0.026)	(0.027)	(0.018)	(0.018)	(0.650)	(0.675)
Ln(employment)	-0.008	-0.010	-0.044***	-0.055***	0.094	0.039
	(0.012)	(0.014)	(0.011)	(0.012)	(0.416)	(0.494)
App. stock/empl.	-0.068	-0.060	-0.324*	-0.413***	11.131	11.116
	(0.208)	(0.219)	(0.167)	(0.157)	(10.197)	(10.600)
(App. stock/empl.)^2	-0.043	-0.048	0.133	0.177**	-0.793	-1.018
	(0.121)	(0.128)	(0.082)	(0.078)	(5.117)	(5.314)
Patent stock/empl.	0.001	0.000	0.002	0.000	0.061	-0.008
	(0.005)	(0.006)	(0.004)	(0.004)	(0.200)	(0.207)
(Pat. stock/empl.)^2	0.000	0.000	0.000	0.000	-0.000	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)	(0.004)
Export dummy	0.020*	0.026	0.065***	0.073***	0.880***	1.532***
	(0.012)	(0.017)	(0.012)	(0.015)	(0.225)	(0.367)
Group member d.	0.075***	0.095***	0.052***	0.066***	0.251	0.310
	(0.016)	(0.022)	(0.016)	(0.019)	(0.321)	(0.472)
Foreign parent d.	-0.036	-0.036	-0.056**	-0.079***	0.308	0.216
	(0.022)	(0.030)	(0.023)	(0.028)	(0.445)	(0.623)
Assets/employee	-0.000	-0.000	0.000	-0.000	0.003*	0.004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.002)
Debt/assets	0.001	0.001	0.000	-0.000	0.002	-0.005
	(0.001)	(0.001)	(0.001)	(0.001)	(0.014)	(0.027)
Cash flow/employee	-0.000*	-0.000	-0.000***	-0.000	-0.002	-0.002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.005)	(0.006)
Observations	11,835	7,476	12,121	7,694	11,337	7,010
R-squared	0.018	0.023	0.026	0.037	0.060	0.083
Number of firms	3,977	2,315	3,980	2,317	3,967	2,304

Robust standard errors in parentheses; all regressions include a full set of time dummies.
*** p<0.01, ** p<0.05, * p<0.1

7 Conclusions

In this report, we have evaluated the treatment effect of VLAIO grants on firm level innovation and R&D inputs as well as outputs. Using different econometric methodologies for the estimation of treatment effects (Pooled OLS regressions, Matching, Difference-in-difference models), we find positive treatment effects throughout the study. The difference-in-differences regressions where a control group of R&D performers (which are not VLAIO grant awardees) are used seem to deliver the most conservative and reliable estimates.

The results concerning R&D inputs clearly reject full crowing out effects of the VLAIO R&D grant programs. Instead, granting a subsidy has a positive effect on R&D employment and R&D spending. In terms of economic magnitude the effects are high, on average. The average treatment effect on the treated amounts to about 3 persons; somewhat less if large outliers are excluded from the data. "Large outliers" here refers to unique large companies. Thus the estimated averages should be generally interpreted with care, as the distribution of R&D inputs is very skewed in the Flemish economy and this results in some sensitivity of estimated average treatment effects. Thus the exact magnitude of the estimated treatment effects should not necessarily be generalized to other context when other samples of firms are considered; neither should it be assumed that the treatment effects remain around three R&D employees when other firms and projects are funded in the future. Thus, the main finding concerning R&D employment is not the exact number of three R&D employees but that the IWT/VLAIO grants surely lead to so-called innovation input additionality. In the absence of IWT/VLAIO grants, Flemish firms would invest significantly less resources into R&D which would probably harm their long-run efficiency and competitive position not only in Flanders but in the global market.

We also find that the estimated treatment effects are stable across different firm size categories. When splitting our sample into small, medium-sized and larger companies, we do not find any statistically significant variation in the treatment effects across the different firm sizes. The VLAIO grants thus have a positive impact on firms of any size and small firms seem to benefit in the same way as medium and larger companies with respect to the absolute magnitude of the treatment effect. Of course, it can be argued that not-having three R&D employees that could be partially financed through A IWT/VLAIO grant, weighs much worse in the innovation input process of a smaller firm than for a large firm.

Furthermore, we find that firms also realize a higher behavioral additionality as measured by collaboration with other companies or research institutions with the innovation projects. There is also evidence on output additionality with regard to realized product and process innovations as well as new product sales.

In future research, the interaction of VLAIO grants with other policy measures present in Flanders could be considered. Examples are public procurement contracts that might involve R&D and innovation, or R&D tax credits and tax breaks on returns to intellectual property.

References

Aerts, K. And D. Czarnitzki (2006), The impact of public R&D-Funding in Flanders, *IWT Study 54*, Brussels.

Aerts, K. (2008), Who writes the pay slip? Do R&D subsidies merely increase researcher wages?, MSI Working Paper 0806, Leuven.

Aerts, K. and Schmidt, T. (2008), Two for the price of one?: Additionality effects of R&D subsidies: A comparison between Flanders and Germany, *Research Policy* 37(5), 806-822,

Angrist, J.A. and S.J. Pischke (2009), *Mostly harmless econometrics*, Princeton University Press.

Czarnitzki, D. and J. Delanote (2017), Incorporating innovation subsidies in the CDM framework: empirical evidence from Belgium, *Economics of Innovation and New Technology* 26(1-2), 78-92.

Czarnitzki, D., and C. Lopes-Bento (2013), Value for Money? New Microeconometric Evidence on Public R&D Grants in Flanders, *Research Policy* 42 (1), 76–89.

Holemans, B. and Sleuwaegen, L. (1988), Innovation expenditures and the role of government in Belgium, *Research Policy* 17(6), 375-379.

Rubin, D. B. (1977), Assignment to treatment group on the basis of a covariate, *Journal of Educational Statistics* 2(1), 1-26

Smith, J.A., and P.E. Todd (2005), Does matching overcome LaLonde's critique of nonexperimental estimators?, *Journal of Econometrics* 125(1-2), 305-353.

Wooldridge, J.A. (2010), Econometric Analysis of Cross-section and Panel Data, 2nd edition, Cambridge, MA: MIT Press.

Zúñiga-Vicente, J.A, C. Alonso-Borrego, F.J. Forcadell, J.I. Galán (2014), Assessing the effect of public subsidies on firm R&D investment: a survey, *Journal of Economic Surveys* 28(1), 36-67.

Appendix: Using R&D expenditure as outcome variable

In addition to R&D employment, we also estimate regression models using an alternative R&D input measure, R&D expenditure. These regressions should be interpreted carefully though as the distribution of R&D expenditure is even more skewed than R&D employment and therefore the results might be somewhat more sensitive to outliers. As R&D spending in Euros has even are more skewed distribution than R&D employment, we scale the expenditure by a firm size measure. Sales are the most commonly used firm size denominator in this case, and we therefore measure R&D intensity in percent as R&D expenditure divided by sales times 100. The average R&D intensity in the sample is 2%.

In the results' table below we see the same pattern as for R&D employment intensities. The naïve POLS regression without controls results in a treatment effect of 3.2%-points, but the POLS regression with control variables accounting for firm heterogeneity already reduces the effect to 2.2%-points. The DID regression in column (3) shows a treatment effect of 0.45%-points.

The average R&D spending in the sample of R&D performers as used in the regressions is 2.99 million Euros. Thus the increase of 0.45% in R&D intensity as response to the receipt of a VLAIO grant roughly corresponds € 150,000.

The average size of a granted subsidy per year in our sample is € 121,000. Therefore, the firms increase their total investment more than the grant size, i.e. in the common understanding of an economist there is no crowding out. The firms increase their R&D spending more than the amount of the public subsidy. In terms of the usual interpretation of policy makers, however, the estimated treatment effect is below expectations. Usually the projects are only partially funded by the public agency. The average subsidy rate of the total project cost is about 50%. This means that a firm would need to increase R&D spending by about € 242,000 (= € 121,000 * 2) for full project additionality. As € 150,000 is clearly below full project additionality, this result suggests a positive treatment effect, but it does not yield full project additionality.

Table 12: Regression results on R&D expenditure (measured as intensity) – sample of R&D performers

VARIABLES	(1) POLS no controls	(2) POLS with controls	(3) DID	(4) DID common trend
TREAT	3.219***	2.192***	0.448***	0.465**
	(0.230)	(0.232)	(0.165)	(0.190)
POST-TREAT	-2.118***	-1.723***	-0.316**	-0.315**
	(0.306)	(0.275)	(0.140)	(0.140)
PRE-TREAT				0.036
				(0.186)
Ln(employment)		-0.446***	-0.336***	-0.336***
		(0.066)	(0.116)	(0.116)
Application stock/employee		28.174***	-0.904	-0.895
		(4.549)	(3.489)	(3.496)
(Application stock/employee)^2		-22.916***	5.277	5.269
		(5.218)	(3.844)	(3.846)
Patent stock/employee		0.136***	-0.004	-0.003
		(0.046)	(0.067)	(0.067)
(Patent stock/employee)^2		-0.004***	-0.000	-0.000
		(0.001)	(0.001)	(0.001)
Export dummy		0.654***	0.284***	0.284***
		(0.135)	(0.085)	(0.085)
Group member dummy		0.451***	-0.045	-0.045
		(0.139)	(0.110)	(0.110)
Foreign parent dummy		0.610***	-0.239	-0.239
		(0.139)	(0.147)	(0.147)
Assets/employee		0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)
Debt/assets		-0.023***	0.000	0.000
		(0.005)	(0.005)	(0.005)
Cash flow/employee		-0.001	-0.003**	-0.003**
		(0.002)	(0.001)	(0.001)
Observations	4,939	4,939	4,939	4,939
R-squared	0.081	0.252	0.031	0.031

Robust standard errors in parentheses; sample: 1,574 different firms; all regressions include a full set of time dummies; POLS regressions also include industry dummies.
*** p<0.01, ** p<0.05, * p<0.1