BANCA D'ITALIA
EUROSISTEMA

Temi di discussione

(Working Papers)

Bridging the gap in innovation: the case of regional R&D subsidies in Southern Italy

by Antonio Veronico

April 2026

Number 1526

BANCA D'ITALIA
EUROSISTEMA

Temi di discussione

(Working Papers)

Bridging the gap in innovation: the case of regional R&D subsidies in Southern Italy

by Antonio Veronico

Number 1526 - April 2026

ISSN 2281-3950 (online)

Designed by the Printing and Publishing Division of Banca d'Italia

BRIDGING THE GAP IN INNOVATION: THE CASE OF REGIONAL R&D SUBSIDIES IN SOUTHERN ITALY

by Antonio Veronico*

Abstract

This paper estimates the causal impact of a regional R&D subsidy funded with EU cohesion funds in favour of small and medium firms operating in Puglia, a region in Southern Italy that has historically lagged behind. I leverage firms' balance sheets and a unique regional administrative dataset on subsidy payments to study how this policy affected the R&D investments, productivity, and survival probability of recipients. By adopting a recently developed match-in-diff approach and using a control group of eligible but unsubsidized firms, I first show that the treated firms display higher survival compared with control firms. Then, I find that the policy was successful in stimulating additional investments in intangible assets. The policy also had positive effects on the number of workers employed in the treated firms. However, I do not find any impact on tangible assets, value of production, and productivity.

JEL Classification: C21, L25, L53, O31, O33, O38, H25, H32, R11, R58.
Keywords: placed-based policy, R&D subsidy, EU funds, match-in-diff.
DOI: 10.32057/0.TD.2026.1526

* Bank of Italy, Regional Economic Research Division (Bari Branch). Email: antonio.veronico@bancaditalia.it

1. Introduction[1]

Innovation is the process of creating new ideas and bringing them to life, often through the use of technology. It is the driving force behind progress and growth in society, leading to advancements in many fields including medicine, transportation, communication, and energy. Innovation can take many forms, from the development of new products and services, to the creation of new processes and systems that make existing products and services more efficient and effective. The positive impact of innovation on the economic growth of a country is well documented in the literature (Solow 1956, Romer 1986). Nevertheless, most of the time firms do not devote sufficient resources to research and development (R&D henceforth). Indeed, innovation has the classical features of public goods, as it is both non-rivalrous and, to some extent, non-excludable. For this reason, if firms are not capable to fully internalize the positive externalises generated by research activities, innovative investments can be below the socially optimal level (Nelson 1959). Furthermore, imperfections in the credit markets can limit the amount of R&D projects undertaken by financially constrained firms (Hall and Rosenberg 2010). For all these reasons, there is scope for policy making to support and stimulate research activities of firms and private institutions. Given the public budget constraints, the evaluation of the effectiveness of these policies has become extremely important. In this paper, I provide novel empirical evidence on this topic by analysing a regional R&D subsidy promoted by Apulia, a lagging behind region in Southern Italy, in the aftermath of the Great Recession. The policy targeted small and medium firms operating in the region, and I evaluate its impact over the period 2009-2015. As a first step, I assess if the subsidy had some effects on the survival probability of recipients firms. Then, I test the additionality of the incentive. That is, I try to understand whether the public resources crowded-out or crowded-in private investments in R&D. Then, I analyze the potential effects of the subsidy on the productivity of firms and explore the channels from which they stem from.

To estimate the causal effects of interest, I gather balance sheets data on firms from 2004 to 2015, and exploit the assignment mechanism and payments data to identify treated and control firms. This mechanism implied the allocation of the subsidy to all the firms whose research project scored a minimum threshold until the budget run out. As result, there were some firms which were assigned the minimum score, but didn't received the subsidy because all the public budget was allocated to firms in a higher position in the final ranking. I use this set of "eligible" firms as control group. Then, to evaluate the impact of the policy I employ a Difference-in-Differences estimation technique by comparing the average outcome of subsidized and non-recipients but eligible firms. To further address the endogeneity concerns

[1]For very helpful comments, I thank Monica Andini, Maurizio Lozzi and Vincenzo Mariani. For the provision of the data, I would like to express my sincere thanks to Adriana Agrimi from *Regione Puglia* and Marco Di Ciano from *Innovapuglia*. The opinions expressed are those of the author and do not necessarily reflect the views of the Bank of Italy or the Eurosystem. All remaining errors are my own.

caused by the non-random assignment of the subsidy, I improve the comparability of treated and untreated firms by matching on their pre-treatment characteristics. First of all, I find that treated firms are 12% less likely to exit the sample with respect to control firms. Moreover, the policy was effective in stimulating additional investments in R&D, as treated firms increased their intangible assets by more than seven times. This effect is quite large, however when accounting for the presence of outliers in the distribution, the increase lower in magnitude but remains sizable. I do not find any crowding-in effect on tangible assets. Moreover, when looking at labor productivity (expressed as the ratio between the value of production and the number of employees), I do not find any difference between the two group of firms. This reflects mainly an increase in employment, that can be both a direct and indirect effect of the subsidy, and a rise in the value of production, although the latter coefficient is not statistically significant. The result on productivity is confirmed also by considering the cost of labor per employee and the TFP[2]. These results are robust to changes in the matching procedure, and in the estimator. A common concern in policy evaluation exercises is that the results may be biased by spillovers between treated and untreated units, which lead to violations of the stable unit treatment variable assumption (SUTVA; Rubin 1980). To address this issue I show that the estimates are robust also to changes in the control group by including only firms which are unrelated to the subsidy under study. By leveraging a unique administrative data source, I am able to distinguish between the firms that received only the subsidy under study ("one subsidy firms") and those that also received other regional subsidies ("multiple subsidies firms"). I find that the policy had a crowding-in effect on intangible assets for both group of firms, however the estimated impact for the latter group is stronger. The same holds for the estimated survival probability. The "multiple subsidy" group mainly drives the effect of the policy on employment, which turns out also positive but not statistically different from zero for the firms belonging to the "one-subsidy" group. There are no effects for the tangible assets and the productivity. It is important to note that this analysis focuses exclusively on the direct effects of the policy and therefore does not account for potential general equilibrium impacts[3].

This paper ties to the literature that studies the effectiveness of public support to innovation at the micro-level. In general, incentives for innovation come in the form of either tax credits or subsidies. The empirical literature that has assessed their effects is vast. In a review Becker (2015) remarks that studies on tax credit mostly display positive results, by stimulating additional investments in innovation especially for financially constrained firms. Other surveys (Zúñiga-Vicente et al. 2014; Cerqua and Pellegrini 2020; Bocci et al. 2023) points out that the effects of R&D subsidies are very heterogeneous across programs and papers. Although

[2]I cannot rule out the possibility of a positive effect on productivity beyond 2015. However, as I do not have information on other incentives paid after that year it becomes difficult to justify the comparability of treated and control firms.

[3]See Siegloch et al. (2025) for a discussion on the spillover effects of place-based policies.

results supporting the additionality hypothesis prevail, there are valuable contributions in favour of the substitution hypothesis and others that demonstrate a negligible effect. For instance, in the context of Italy Bronzini and Iachini (2014) studies the impact of a regional R&D subsidy on innovation input (i.e. private R&D spending) and finds a positive effect only for small businesses. Similarly, Hud and Hussinger (2015) and Fantino and Cannone (2013) confirm the additionality of these policies in the European context. Other researches focus instead on the impacts on innovation output (i.e. patents), finding either positive (Bronzini and Piselli 2016) or null effects (Accetturo 2022). Some studies analyse the indirect effect of public support to R&D on the performance of recipients firms. For instance, Bertamino et al. (2016) studies the effects of a placed based policy in Italy on firms performances and productivity and finds that the measure had a positive effects on valued added and ROA for small firms but no effects on productivity. Liberati et al. (2016) assess the impact of science and technological parks in Italy and shows that entering a science and technology park did not generally improve firms' business performance and their propensity to innovate compared with external counterparts. Chinetti (2023) in a study on the effectiveness of a regional program implemented in a lagging behind area in the South Italy, shows that the program has considerable indirect effects on medium-large low tech service firms' labour demand but not on overall firms' productivity. This paper relates also to the literature which studies the role of subsidies' accumulation by firms to understand if the persistence in receiving public money is based on sheer reputation or on firms' characteristics related to competence and commitment to innovative activities (Antonelli and Crespi 2013). On the topic, Santoleri and Russo (2023) studies the effects of different R&D subsidies to start-up firms in Italy, and shows that securing local subsidies increases start-ups' probability to obtain additional public incentives, but at the same time these programs had a null effect on both innovation input and output.

This paper improves on the existing literature by providing novel quasi-experimental evidences which leverage unique administrative data of payments made by the regional administration to firms. This data allows me to identify more precisely treated firms with respect to those gathered by publicly available final rankings. Moreover, with this data I can separate other regional policies which may confound the effect of the subsidy under study. The rest of the paper is organized as follows: section 2 describes in detail the policy object of study and the regional context; section 3 outlines the estimation strategy; section 4 describes the data, the outcomes of interest and provides some summary statistics about treated and untreated firms; section 5 presents the estimated results and the heterogeneity analysis; section 6 shows the robustness checks; section 7 discusses the cost-efficiency of the policy; section 8 concludes.

2. The policy and its context

The policy under study was part of a larger set of interventions in favor of less developed regions in Southern Italy with the objective of fostering R&D activities by private firms located in these areas (Lotti and Stefani 2014). In particular, in this paper I analyze a specific policy called *Aid to research investments of SMEs*. It was put in place by the Apulia region in 2008 during the financial crisis, with the objective of stimulating firms' innovation investments. Like many other similar placed-based incentives, this policy was co-financed by the European Regional Development Fund (ERDF), an instrument created to reduce the gap between the levels of development of Europe's regions and the extent to which the less-favoured ones are lagging behind. Indeed, between 2005 and 2007, the Apulian GDP per capita was approximately one third lower than the Italian average, and the Great Recession (2007-2013) caused a 10% drop of this indicator. The policy was directed to all small and medium firms[4] located within the Apulian borders, apart from those operating in the primary sector, some manufacturing activities (such as metallurgy) and tourism. The initial amount of resources allocated by the policy was 28 millions of euros. However, before the publication of the final ranking, the budget was raised by an additional 20 millions bringing the total allocation of resources to 48 millions of euros. This amount was roughly 32% of the total investments in intangible assets[5] made by Apulian small and medium limited companies. Indeed, these firms invested a total of 150 million of euros on average between 2005-2007, an amount below the Italian mean, but above its median value (Figure 1 left panel). In the same time period, the incidence of investments in intangible assets over the value of production was 0.8%, again below the Italian average (Figure 1 right panel).

To apply for the subsidy, firms had to submit a project proposal in either industrial research or in experimental development. Additionally, applying companies had to meet certain financial requirements: first, the ratio between own capital and the net cost of the proposed project had to be less than 50%; second the ratio between net financial costs and revenues had to be lower than 8%. The aid intensity was capped at 70% of eligible costs for small enterprises and 60% for medium-sized enterprises. The maximum fundable contribution was limited to 1 million euros for industrial research and 700 thousand euros for experimental development activities. Companies were required to cover at least 25% of the total project costs, either through their own resources or external financing. The costs admitted to subsidization included those for the hiring of high skill employees involved in the project, the costs of new equipment and machinery, the expenses related to the research acquired from third parties (such as Universities) and those for the purchase or registration of patents. After the submission an independent

[4] In Italy, small and medium firms are defined based on European Commission Recommendation 2003/361/EC and the Decree of 18 April 2005. An SME is a company with fewer than 250 employees, an annual turnover not exceeding €50 million, and/or an annual balance sheet total not exceeding €43 million.

[5] Intangible assets are a good proxy of firms' R&D activity as I will explain better in Section 4.

committee[6] was given the task of evaluating the proposal by assigning a score based on both the quality of the project and the previous experience of the firm in R&D activities.

To be eligible for the subsidy a firm was required to score at least 60 points. Then, in the final ranking each firm with a score above the minimum threshold got the subsidy only until the resources ended. The recipients were required to start the project within a fast and predetermined time-frame (30 days), proceed quickly with the implementation of the investments, and conclude them within a maximum of 24 months from the decision. The subsidy could be paid in a single balance solution, in two equal installments with the first one during the investment realization, or through a 30% advanced payment subject to the submission of a bank guarantee or insurance policy. The final ranking was composed by 264 firms out of which 243 scored 60 points (the minimum threshold) and, on the base of the total resources allocated only the first 149 should have received the subsidy. However, by analyzing the data on the payments made by the public administration, only 117 firms received the subsidy and 27% of the budget was left unused (Table 1). The choice of firms that were actually paid as the treatment group, instead of the full set of winning firms identified in the final ranking, could introduce bias into the estimates due to the self-selection of the paid firms. However, as shown in Table A.1, the two groups of firms do not differ substantially in terms of pre-treatment observable characteristics; therefore, the bias is likely negligible.

Figure 2 plots the distribution of the subsidy among the payed firms; the average subsidy amounted to 300 thousand euros, while the median value to 261 thousand. Firms in the top 5% of the subsidy distribution received more than 700 thousand euros. For what concerns, instead, the total costs declared by firms for the implementation of their research projects, they amounted to 64 millions, almost double the total subsidy paid. The average project's cost was of 548 thousand euros which was roughly half of the average subsidy, and 8% higher than the average investment of payed firms in 2007. Figure 3 plots the distribution of payed firms over NACE 1 digit sectors. Almost half of the firms operates in the manufacturing sector, while another sizable share are in advanced services such as information and communication or professional and scientific activities.

3. Estimation strategy

To address the endogeneity issue that arises by the non-random assignment of the incentive, I employ a Difference-in-Differences (DID) estimation strategy. This method allows me to control for unobservable firm specific features and also common shocks. In the recent years, the theoretical literature around DID estimators has made several steps forward: it highlights the flaws of the canonical two-way fixed effects estimator and proposes alternative estimation strategies in presence of heterogeneous treatment (see Roth et al. 2023 for a comprehensive

[6] It was composed by both economic experts employed in the regional administration and university professors.

review). To understand which estimator best fits this framework, it is useful to identify the treatment timing. As the awarded firms got payed in different time points between 2010 and 2013 (Table 1), it would seem reasonable to assume that a firm is treated only after having received the subsidy (or at least the first payment). This would naturally lead to a staggered time heterogeneous treatment framework. However, by analyzing more carefully the timeline of the administrative obligations envisaged by the policy (Figure 4), subsidized firms should have started their projects, and hopefully their investments, as soon as the final ranking got published in 2009. Moreover, the investment projects had to be completed by the end of 2011. If the treated group would have started investing already in 2009, the adoption of the subsidy payment year as treatment date would lead to a violation of the no-anticipation assumption. Given this, it seems reasonable to assume that all the awarded firms got treated in 2009. Moreover, to ensure the comparability of treated and control firms, I adopt a recent estimation technique which combines matching with DID, described in the section below[7].

3.1. The match-in-diff approach

To estimate the economic impact over time of the policy, I employ a recent evaluation technique proposed by Imai et al. (2023). This procedure combines matching and DID strategies to build counterfactual outcomes of treated units and then, computes credible estimates of treatment effects for the subsidized firms as:

$$\begin{aligned}\delta(F, L) = \mathbb{E}[Y_{i,t+F}(X_{it} = 1, X_{i,t-1} = 0, \{X_{i,t-l}\}_{l=2}^{L}) \\ -Y_{i,t+F}(X_{it} = 0, X_{i,t-1} = 0, \{X_{i,t-l}\}_{l=2}^{L}) | X_{it} = 1, X_{i,t-1} = 0]\end{aligned} \tag{1}$$

The first line of equation (1) is the potential outcome under treatment (i.e. $X_{it} = 1, X_{i,t-1} = 0$), while the second line is the potential outcome in the control status. $\delta(F, L)$ is the average causal effect of being treated on the outcome, F periods after the treatment, while assuming that the potential outcome only depends on the treatment history up to L years earlier. This type of non-parametric estimator has the advantage of not imposing a precise model for the data generation process, that if incorrectly specified may lead to biases. Matching methods improve the validity of causal inference by reducing model dependence and offering intuitive diagnostics (Imai and Kim 2021). Nevertheless, in the robustness section I provide the results for alternative estimators like the classical two-way fixed effects model and a doubly-robust version of it (Sant'Anna and Zhao 2020). The estimation procedure involves four different steps. In the first one, it requires the use of matching methods to build, for each treated unit, a matched set $M_{i,t}$ in which are collected all control firms that have the same treatment history (from time $t - L$ to $t - 1$) of treated unit i. Since in this case the treatment is time homogeneous, the matched set will be the same for all treated units. In the next step, each

[7] Another possible identification strategy would be to exploit the final ranking of firms in a regression discontinuity framework. In the appendix I explain why this option is not feasible.

$M_{i,t}$ is refined through a matching algorithm. The objective is to build a weight for each control in $M_{i,t}$ based on its similarity with the treated unit considered. Next, counterfactual outcomes for each treated observation are built using the weighted average of the control units in the refined matched set. Finally, it's possible to compute the DID estimates of the average treatment effect on the treated (ATT) for each observation and then, average it across all treated units. Formally, the ATT estimator is given by:

$$\hat{\delta}(F, L) = \frac{1}{\sum_{i=1}^{N} \sum_{t=L+1}^{T-F} D_{it}} \sum_{i=1}^{N} \sum_{t=L+1}^{T-F} D_{it} \left\{ (Y_{i,t+F} - Y_{i,t-1}) - \sum_{i' \in M_{it}} w_{it}^{i'} (Y_{i',t+F} - Y_{i',t-1}) \right\} \tag{2}$$

where D_{it} is the treatment indicator and $w_{it}^{i'}$ are the normalized non-negative weights which sum up to one. This non-parametric generalization of the DID estimators is based on two major assumptions: the no interference and the conditional parallel trend assumption. Both assumptions are discussed in depth and tested in section 6.

4. Data and outcome variables

In this paper I employ two different sources of data. The first one is the resolutions of the regional public administration, which contain information on the policy characteristics, the final ranking, the amount and timing of the subsidy payments and the list of firms that have obtained other R&D incentives in the years after the policy object of this study. This unique source of data allows the identification, among the firms which applied for this subsidy, of those that in the period 2009-2015 obtained other forms of financial support to R&D from the regional administration. This gives me the opportunity to isolate confounding policies that may induce biases in the estimates. However, I cannot rule out the possibility that the same firms could have obtained other public resources to support different types of activities. The second source is firms' balance sheets provided by the Cerved group. This information is then enriched by administrative data on employment (Infocamere and Inps data). This source gives access to almost all accounting variables, on the statement of assets and liabilities of limited companies.

Regarding the outcome variables, the natural candidates for the evaluation of the policy are the costs admitted for the subsidy such as the expenses for the purchase of new machinery, the acquisition of patents, the hiring of employees and those for the research activities made directly by firms. However, the available data does not contain precise measures of all those items. Indeed, firms' R&D activity can take many forms and the related expenses can be classified under different balance sheet voices. The best proxies are tangible and intangible assets[8]. I focus on stocks rather than flows (such as intangible investments) as the first measure is much less noisy than the second, and it also takes into account the depreciation of capital. Then, I evaluate the impact of the policy on the productivity of firms. Investments in research and development generate knowledge that can enhance efficiency by enabling the production of a desired output quantity using fewer resources. Most indicators of productivity are computed as the ratio of an output volume to an input volume. In this analysis, I employ the simplest measure of productivity which is the ratio between the value of production and employment. Moreover, I disentangle the single channels that drive the effects on productivity by looking at how the awarded firms adjusted the number of employees and the value of production in response to the shock caused by the subsidy. In particular, the impact on employment, in addition to reflecting a direct effect of the policy, could also reflect an indirect effect since new investments could generate new job opportunities. Among the robustness exercises, I also propose alternative measures to proxy for R&D activities and productivity (i.e. the cumulative intangible and tangible investments, the value-added per employee, the cost of labour per employee and the TFP). I express all variables in natural logs to interpret the coefficients as log-differences with respect the first pre-treatment period.

4.1. Sample selection and descriptive statistics

Upon removing the firms that failed to achieve the minimum score, the remaining dataset is an unbalanced panel consisting of 221 firms, including 109 treated entities, monitored over a 21-year span from 2001 to 2021. One of the main difficulties in policy evaluation is the possibility that the results can be biased by the presence of multiple policies displaying their effects during the period under analysis. As the data provided contains the detail solely regarding subsidies provided by the Apulian regional administration, I have dropped 26 firms with their primary headquarter located outside the administrative border of the region. This exclusion is based on the possibility that these firms could have also received subsidies from the administrations of other Italian regions. Moreover, I drop 16 untreated units who got other funds from the Apulian regional administration. The analysis stops at 2015 since I am unaware of the other incentive that firms could have received after that year. Additionally, to give some evidence of common trends, at least in the pre-treatment period, I take 2004 as the starting date. Figure 5 shows the treatment distribution after the selection procedure. The unbalanced panel is made of 179 firms (out of which 91 treated). The first evidence that emerges is that, in the post-treatment period, subsidized firms are more likely to be observed with respect to control units. By combining the data with the *Infocamere* dataset which

[8] This can be considered a good approximation of firms' R&D activity. According to the Italian Civil Code tangible assets include all physical capital held by a company for long-term use in its operations, such as land, buildings, machinery, and equipment. Instead, intangible assets include the costs sustained for research, patents, software, licenses, advertisement, and more in general the costs of intangible goods that have multiyear utility. Moreover, Bronzini and Iachini (2014) shows that R&D, patents, software, licences, trademark and ongoing intangible assets cover on average about 66% of the total intangible assets.

collects information on the causes of firms' closure, I find that most of the cases (where the cause can be identified) are episodes of bankruptcy and liquidation (Table 2). However, for almost one third of the cases, I was not able to retrieve the reason behind the unobservability of firms' balance sheets. The unequal pattern of missing data between treated and untreated firms motivates the first part of the analysis where I try to assess whether the policy increased the survival probability of the awarded firms.

Table 3 displays pre-treatment averages (2004-2008) of some firms' characteristics by treatment status in the unbalanced panel. Intangible, tangible and total assets, the value of production and productivity (given by the ratio between the value of production and the number of employees) are expressed in thousand of euros. Subsidized firms tend to be larger than control firms: the average number of total assets is 45% higher, and treated firms employ approximately 12 additional workers relative to control firms. Nevertheless, treated and control firms display, on average, the same weight of intangible assets over total assets. Moreover, by looking at the performances (i.e. productivity) control and treated firms appear to be comparable, and the same evidence holds also for the return on assets (computed as the ratio between the net income and total assets). The two groups display a similar degree of openness towards the credit market (i.e. the leverage), and sectoral distribution. The third column of Table 3 reports the standardized mean differences (SMDs) between the two groups, which summarize the unbalances of the sample. Several empirical studies have examined the appropriateness for using SMDs in balance assessment (Belitser et al. 2011 and Ali et al. 2014); in general, values less than 0.1 are sign of good balance between two groups. Figure 6 illustrates the temporal evolution of the outcomes for both treated and control firms. In the top left panel is plotted the probability of observing the balance sheet of the firms. The increase in the probability of observation in the pre-treatment period reflects the entry of new limited companies into the sample[9], while the reduction observed in the post-treatment period signifies the exit of firms from the sample. As showed before in Figure 5 there is a difference in the "survival" probability of treated and control firms, and this gap reaches a maximum of 26% in 2015.

The top right panel of the figure displays a significant rise in intangible assets of treated firms after 2008 compared to control firms. The left panel in the middle of the figure shows some differences in the level of log of tangible assets for treated and control units which decreases towards the end of the analyses period. Notable distinctions in log of productivity (right panel in the middle of the figure) are not observed in the initial post-policy years, but some disparities emerge after 2011 due to a decline in the productivity of control firms. The logarithm of the number of employees (bottom left panel) displays a slight rise for both

[9] The survival lines are not equal to one in first treatment period because for some firms I can observe the balance sheet only in the post-treatment period (see Figure 5). These firms are excluded from the regression analysis.

treated and control firms. The former group shows a similar trend also for the log of the value of production (bottom right panel). The figure underscores that, prior to treatment, untreated and treated firms demonstrate some differences that point towards the need of adopting matching procedures to build a more credible control group.

5. Results

5.1. Matching tuning

In this section, I explore the effects of the R&D subsidy on the outcomes described in the previous section. I adopt the match-in-diff strategy proposed by Imai et al. (2023) to address the selection bias caused by the non-random assignment of the subsidies. The matching procedure requires the choice of three parameters: the number of time lags (L), the variables on which performing the matching ($\mathbf{Z}$), and the algorithm. For what concerns the first choice, as this method requires a balanced panel for a chosen outcome, and given the pattern of missing values of the data (see Figure 5), I face the classic bias-variance trade-off. Indeed, by choosing a larger number of lags the synthetic counterfactual will better approximate the potential outcome of treated firms in the counterfactual scenario (in which they had not won the subsidy). At the same time, some observations are lost since not all firms are observed L periods before the policy. Therefore, in the main specification I choose a lag window of three periods (i.e. 2006 as the first year of observation) and in the robustness section I test the sensitivity of the estimates to different choices of L. As second step, I choose the firms' characteristics on which performing the matching. In the previous section, by comparing subsidized and control firms in the pre-treatment period, it came to light that the main difference between the two groups is size. For this reason, in the baseline specification I match only on the "pure economic" outcomes object of the estimation (i.e. the log of intangible assets, the log of tangible assets, the log number of employees and the log of productivity[10]), plus a dummy equal to one if the firm operates in the industry sector. The algorithm used for the refinement is the covariate balancing propensity score weighting which assigns a larger weight to the control observation that are more similar to a specific treated observation[11]. The matching procedure is repeated for each outcome and, since it requires a balanced panel, the sample size may vary depending on whether for each firm the outcome is always observed or not.

5.2. Survival analysis

Both Figure 5 and 6 highlighted that the balance sheets of treated firms are more likely to be observed than control firms. Given this evidence, it seems natural to assess whether it is a

[10] The value of production is excluded to make the data, on which the matching is performed, linearly independent.

[11] See Imai and Ratkovic (2014) for more details.

direct effect of the subsidy. However, by analysing the reasons behind the pattern of missing data, I was not able to identify the cause for almost one third of the cases. For this reason, what I will estimate in this section is the effect of the subsidy on the probability of observing the balance sheet of treated firms. With a little forcing let's call this outcome "survival probability". This is a natural question to ask since the period under analysis straddles two crisis (i.e. the financial and the sovereign debt crisis[12]). Several works have estimated the long-term survival of firms after receiving public support. For instance, Moller and Lotti (2013) studies the effect of an Italian incentive to boost female entrepreneurship, and finds that subsidized firms show higher survival rates than non-subsidized firms for a period of up to five years after incorporation. Similarly, Zhang and Xu (2019) reports that government supports to Chinese manufacturing firms significantly decrease the likelihood of firm exit.

To provide additional evidence on the topic I proceed in the following way. First, I define a new binary outcome ($survive_{it}$) equal to 1 if the balance sheet of firm i is observed in year t and 0 otherwise. Then, I keep in the sample only the firms that are always observable in the pre-treatment period[13] (i.e. 2006-2008). Finally, I perform the matching and estimating procedure described in the previous sections[14]. In Figure 7 I evaluate the goodness of the matching by comparing the pre-treatment dynamics of standardized mean differences (SMDs) of the variables in $\mathbf{Z}$, both before and after the refinement. In the left panel all outcomes show unbalances in favor of subsidized firms before the refinement, except for the log of productivity. The right panel shows the balance improvement after the matching. For all lags, the SMDs of the outcomes are more aligned towards the zero than before. The same information can be inferred by looking at Figure 8 which shows a scatter plot of the outcomes balance for each lag, both before (horizontal axis) and after (vertical axis) the refinement. Since all the points lie to the right of the 45-degree line, this indicates that the balance has improved for all variables.

The estimated coefficients are reported in the first column of Table 4. Panel A shows the treatment effects pooled across the entire lead window, while in panel B the event study estimates are reported. I find that the policy had a positive effect on the survival of the selected treated firms: they are 12% more likely to be observed with respect to control firms[15]. The event study displayed in the top left panel of Figure 9 shows that the effect increased over time, and reached the maximum of 21% five years after the treatment.

[12]See Arrighetti et al. (2018) for an assessment of Italian firms' survival during the two recessions.

[13]This step is required since the matching is performed on all the variables in $\mathbf{Z}$ and for all the chosen time periods lags. The firms dropped in the procedure amount to 46.

[14]It is worth noting that in this case the estimator is a diff-in-means rather than a diff-in-diffs one, since the pre-treatment values of the outcome are by construction always equal to 1.

[15]The coefficient estimated is likely a lower bound of the true causal effect, since I have excluded from the analysis the firms born after 2006. These younger firms may have been more affected by the two crises happened between 2007 and 2015.

5.3. Effects on balance sheet outcomes

Columns 2 to 6 of Table 4 present the estimated coefficients on the balance sheet items directly targeted by the policy. In the second column, I find that treated firms increased their investments in R&D relative to control firms: the amount of intangible assets increased by approximately 2.2 log points in the period from 2009 to 2015, indicating a more than seven-fold increase[16] (i.e. an increase of about €1,800,000 relative to the pre-treatment mean of awarded firms[17]). Moreover, the coefficient is strongly significant. This result highlights that the policy had a crowding-in effect: without the subsidy, treated firms wouldn't have invested the same amount of money. The top-right panel of Figure 9 plots the dynamic over time of the estimated ATTs. The effect is bell-shaped: during the first years, the growth rate of subsidized firms' intangible assets rose relative to the one of control firms, and then slowed down after 2011, the year by which treated firms should have completed their project. By looking at the third column of Table 4 the subsidy had no impact on tangible assets. The estimated coefficients are small and never statically different from zero. A possible explanation for this result is that only the depreciation quotas of newly purchased machinery, directly attributable to the research project, were eligible for the subsidy. This probably did not constitute a strong enough incentive to stimulate additional investments in tangible assets. Moreover, panel C of Figure 9 does not show a significant heterogeneity of the effect over time. The fourth column of Table 4 reports the impact on productivity. The coefficient is close to zero and not statistically significant. Furthermore, by looking at panel D of Figure 9, the result holds also for each year observed after the treatment. These dynamics reflect the different responses of the two variables underlying the measure of productivity: the number of employees and the value of production. Columns 5 and 6 of Table 4 report the estimated coefficients for the two outcomes respectively. By looking first at employment, the pooled ATT is positive and slightly significant: the number of workers employed by treated firms grew approximately of 0.3 log points (35%, about 11.7 additional workers compared to the pre-treatment mean of awarded firms[18]). The effect grew with time, but the estimated coefficients loose precision four years after the intervention (panel E of Figure 9). The effect on employment might reflect both a direct and an indirect impact of the subsidy. In fact, the awarded firms could have used the funds also for the hire of new qualified personnel. The effects of the subsidy on the value of production are positive but not significant (last column of Table 4).

Overall, in a period of six years after the policy, the awarded firms became larger than those who were eligible for the subsidy but didn't get one. However, the increase in dimension did not translate into a better performance in terms of productivity. Since the analysis is censored at 2015, I cannot exclude that the investments in R&D produced additional effects after that year. The estimated coefficients should be interpreted with caution. Indeed, on one hand,

[16]The increase is computed as $(e^{2.173} - 1) * 100$.

[17]In the balanced panel this corresponds to approximately 230.000 euros.

[18]In the balanced panel this corresponds to approximately 33 workers.

the estimates are biased downward because they are conditioned on the survival of the firms throughout the period of time studied. On the other hand, the presence in the sample of firms that have received multiple incentives might confound the true effect of the policy. In the next subsection, I address the last point by exploiting an exclusive administrative data source which allows me to explore the heterogeneous effect of the subsidy depending on the number of R&D subsidies received by the firms between 2009 and 2015.

5.4. Heterogeneous effects: One vs. Multiple subsidies firms

The results presented in the previous section do not necessarily capture the true effect of the subsidy under study, as the estimated coefficients may be influenced by other regional policies implemented during the same period. In particular, among the firms that applied for the incentive, 45 also received additional R&D subsidies promoted by the Apulian administration between 2009 and 2015. On average, these firms obtained €225,000 in additional support—about 25% less than the amount associated with the subsidy analyzed in this paper. Of these firms, 36 eventually won and received the payment related to the subsidy under study. The purpose of this section is to examine how the estimated treatment effects vary when accounting for heterogeneity in the number of subsidies received. I distinguish between two groups of treated firms: "one-subsidy" firms, which to the best of my knowledge received only the subsidy analyzed in this study, and "multiple-subsidies" firms, which were also beneficiaries of other regional R&D incentives during the analysis period. It should be stressed that I cannot exclude the possibility that firms in either group may have received additional regional or national support. This distinction is relevant in light of recent evidence on the allocation and persistence of public subsidies in Italy. Santoleri and Russo (2023) shows that receiving an initial subsidy increases the likelihood of obtaining further public funds, although this does not necessarily translate into additional innovation. Similarly, Antonelli and Crespi (2013) documents a "picking-the-winner" strategy in Italian R&D support schemes, whereby public agencies tend to select firms with a higher probability of successfully completing the proposed projects. In what follows, I contribute to this literature by comparing the estimated effects for "one-subsidy" and "multiple-subsidies" firms, using the same pool of control units for both groups.

Before turning to the causal estimates, it is useful to compare the pre-treatment characteristics of the two treated groups with respect to control firms. Table 5 reports the standardized mean differences for selected covariates for "one-subsidy" firms (column 2), "multiple-subsidies" firms (column 3), and the entire treated sample (column 1). As previously observed for the full sample, firms in both groups tend to be larger in terms of employment. However, some notable differences emerge. "One-subsidy" firms display lower levels and lower incidence of intangible assets relative to control firms, while productivity and ROA appear fairly balanced.

Conversely, "multiple-subsidies" firms exhibit a higher share of intangible assets over total assets but lower productivity compared to controls. Moreover, the latter group is more likely to operate in the industry sector than control firms. Figure 10 illustrates the dynamics of average outcomes. "Multiple-subsidies" firms show a higher probability of survival relative to both control units and "one-subsidy" firms. They also experience larger increases in intangible assets and employment, although their productivity levels remain lower.

To ensure comparability, I rely on the same unit-specific synthetic controls constructed through the matching procedure described earlier. Table 6 reports the pooled ATTs for six outcomes, with separate estimates for the full treated sample (panel A), "one-subsidy" firms (panel B), and "multiple-subsidies" firms (panel C). The first column reports the effect on the probability of observing the balance sheet of treated firms. The impact is positive for both groups, but stronger for "multiple-subsidies" firms. A similar pattern emerges for intangible assets (column 2). "Multiple-subsidies" firms increased their intangible assets by approximately 2.7 log points relative to controls, compared to 1.7 log points for "one-subsidy" firms. The event study in the top-right panel of Figure 11 shows a bell-shaped response for both groups, although the post-2011 decline is slower among "multiple-subsidies" firms, suggesting either lower capital depreciation or faster replacement. Column 3 indicates that the subsidy had no detectable effect on tangible assets for either group. The estimated effects on productivity (column 4) differ in sign—positive for "one-subsidy" firms and negative for "multiple-subsidies"—but are statistically not different from zero in both cases. The event studies (panel D of Figure 11) show that productivity remains close to zero for "one-subsidy" firms during peak investment years, while it dips below zero for "multiple-subsidies" firms before converging again. Regarding employment (column 5), the positive and significant effect observed in the baseline specification appears to be almost entirely driven by "multiple-subsidies" firms, whose workforce expanded by 0.39 log points (roughly 47%). This may partly reflect specific hiring requirements or incentives associated with the additional subsidies received[19]. The effect for "one-subsidy" firms is positive but not statistically significant. Finally, the effects on the log value of production (column 6) are not statistically different from zero for both groups.
This heterogeneity analysis reveals that the positive and statistically significant results identified in the baseline specifications are mainly driven by "multiple-subsidies" firms. While the subsidy also increased intangible assets among "one-subsidy" firms, the effects on productivity and employment are not statistically significant for this group.

[19] For instance, under the policy *Aiuti alla diffusione delle tecnologie dell'informazione e comunicazione nelle PMI*, firms received bonus points for each new hire, subject to a two-year no-dismissal rule.

6. Robustness checks

In this section I provide several robustness checks to corroborate the results reported in sections 5.2 and 5.3 where I presented the estimated coefficients for the whole sample. In particular, I carry out several robustness exercises: I use a control group of firms completely unrelated to the subsidy under study with the objective of both increasing the sample size and also check if the no-interference assumption is reasonable. In another robustness check, instead, I test different specifications of the matching procedure such as using an alternative algorithm, or add more variables in $\mathbf{Z}$, and altering the number of lags on which the balancing is performed. Additionally, I provide the results on alternative measures of innovation input and productivity. Then, I show the sensitivity of the results to outliers and changes in the estimator. In the appendix, I explore the possibility of an alternative identification strategy and then explain why I chose to discard it.

6.1. Alternative control group

The first robustness exercise uses an alternative control group which is unrelated to the subsidy under study. The rationale behind this check is to assess whether the results are driven by spillovers or interactions between subsidized and eligible but unsubsidized firms (see Cerqua and Pellegrini 2013). This would result in a violation of the no-interference assumption, which is closely related to the stable unit treatment value assumption. To build this alternative control group I select all the Apulian firms that in 2008 had the requirements to apply for the subsidy: they must be small-medium firms, run their business on the Apulian territory, not operate in some specific NACE sectors (i.e. fishery, ship building, coal manufacturing, iron and steel industry, synthetic fibers and truisms), and at the moment of application the ratio between financial costs and revenues must be below 8%. In this way, I can identify more than 10.000 control firms that could have applied for the subsidy (more than 5.000 in the balanced panel). Then, I run the same matching procedure described in section 5, and present the pooled ATTs in the second column of Table 7. The top panel of the table reports the estimated coefficients for the survival probability, while the bottom panel shows the results for the balance sheet outcomes. Overall the estimated coefficients are bigger with respect to the one reported for the baseline sample (column 1 of Table 7). Only minor differences arises for the log employment and the log value of production. The effect on productivity remains statistically not different from zero. The impact on tangible assets turns out positive but only slightly statically significant. The conducted exercise allows me to exclude the hypothesis that there could have been negative spillover effects from treated firms to untreated ones. These would have arisen if the estimated coefficients had been systematically lower than those reported in the baseline analysis.

6.2. Alternative matching settings and placebo test

These robustness checks involve alternative specifications of the matching procedure. In particular, I first use the mahalanobis distance instead of the covariate balance propensity score weighing as a different matching algorithm (Table 7, column 3), then I add two covariates (leverage and ROA) in the vector $\mathbf{Z}$ on which the balancing is performed (column 4). Additionally, I modify the matching specification by both increasing by one (column 5) or two (column 6) the number of pre-treatment periods used to balance the characteristics of treated and control firms.

The effect on survival probability remains positive and statistically significant when I modify the matching algorithm and add additional covariates in $\mathbf{Z}$. Unsurprisingly, the coefficient becomes statistically not different from zero when I match on a larger number of lags (columns 5 and 6). In fact, by imposing an earlier starting date, I am comparing firms that are older and perhaps more resilient with respect to the firms used in the baseline sample. Hence, the relative difference between treated and control firms becomes small and statistically not different from zero. The result on the log of intangible assets is confirmed across all the robustness checks performed, as well as those for the log of tangible assets and the log of productivity. The estimated coefficients on the logged employment and value of production are all of the same sign and are estimated more or less precisely depending on the matching specification. Finally, to further support the absence of violations of the parallel trend assumption, column 7 reports estimates obtained by matching only on lags $t-5$ and $t-4$, while Figure 12 displays the results of a placebo test for lags from $t-5$ to $t-2$[20]. The figure shows that for all outcomes, despite matching being performed on only two lags, the coefficients estimated for the remaining pre-treatment periods ($t-3$ and $t-2$) are also not statistically different from zero. Moreover, the estimated effects reported in column 7 are consistent with those from the most comparable specification presented in column 6.

6.3. Sensitivity to outliers

In this section, I assess the sensitivity of the estimated effects on balance-sheet variables to the presence of outliers. Specifically, for each balance-sheet outcome considered, I apply winsorization using Tukey's fences as the threshold[21]. The results of this exercise are reported in the last column of Table 7. Among all outcomes, the one that appears most sensitive to the presence of extreme values in the distribution is the logarithm of intangible assets. Indeed, the estimated coefficient is almost half the size of that obtained in the baseline specification. The effect—equal to 1.2 log points—remains, however, large and statistically significant, implying a threefold increase amounting to roughly 600,000 euros. For the other outcomes, the results

[20] $t-1$ is set as the reference period.

[21] The lower bound is defined as $Q1 - 1.5 \times IQR$, and the upper bound as $Q3 + 1.5 \times IQR$, where $Q1$ is the first quartile,$Q3$ is the third quartile, and IQR is the interquartile range.

are broadly consistent with those of the baseline specification.

6.4. Additional outcomes

As an additional robustness check, I perform the estimates on alternative proxies of private R&D spending and the productivity. In particular, for the first outcome I evaluate the effect on the cumulative investments in intangible assets and tangible assets, which do not take into account the depreciation of capital. Regarding productivity, in addition to labor productivity I use value added per worker and labor cost per worker as alternative measures. The reason to include the latter variable is that in general more productive firms pay higher wages. Moreover, to adopt a more general measure of firm productivity, I use the total factor productivity (TFP) as the outcome variable[22]. All outcomes are expressed in logs except for the value-added per employee and the TFP. I estimate the impact of the subsidy using the baseline specification as well as the others described in the previous sections. The results are presented in Table 8. The first row shows the pooled ATTs for the log cumulative intangible investments: the coefficient is positive and statistically different from zero in all specifications, but is significantly lower in magnitude with respect to what was estimated for the stock of intangible assets. This could mean that subsidized firms invested in assets with a longer life span than those of untreated firms. In other words, the stock of control firms' intangible assets depreciate at a faster rate than the one of treated firms. For what concerns the alternative measure of tangible investments and productivity, I do not find any coefficient statistically different from zero.

6.5. Alternative estimators

In this subsection, I provide the results obtained by employing alternative estimators to the match-in-diff used in the main analysis. Although this type of estimator has the advantages of not imposing a precise model to the data generation process and building a credible control group for each treated unit, it can generate biased estimates when the matching occurs on pre-treatment outcomes: by forcing those of treated and untreated unit to be equal it might undermine the second difference in the diff-in-diff estimator (Ham and Miratrix 2024). For this reason, I redo the policy evaluation exercise by using alternative estimators. The first one is a naive two-way fixed effect model in which each outcome is regressed against a time-varying treatment indicator and unit and time specific indicators. This model assumes that the parallel trend assumption holds unconditionally, which hardly holds in this application. Therefore, in another model, I augment the naive TWFE by inserting controls for the pre-treatment variables contained in $\mathbf{Z}$[23]. In this model, I am imposing a conditional parallel

[22]TFP is estimated using the methodology described in Wooldridge 2009.

[23]In particular, I control in the regression for the interaction between a linear trend and each of pre-treatment values of the variables in $\mathbf{Z}$, i.e. the log of intangible and tangible assets, the log of productivity and the log of employment.

trend assumption which poses the additional requirement that the evolution of the outcome over time is modeled correctly. To relax this condition, in the last model, I employ a doubly-robust version of the TWFE. By combining the inverse probability weighting model with the outcome regression model, this estimator remains consistent for the ATT when at least one model of the two is correctly specified (Sant'Anna and Zhao 2020).

Table 9 illustrates the results. The first column reports the baseline estimates of the match-in-diff model for comparison. The second column shows the result for the naive TWFE model, and the last two columns for the conditional TWFE and the doubly-robust inverse probability weighting model (DRIPW). The top panel of the table shows the effect of the policy on the survival probability of treated firms. The estimated coefficient is very much similar across all the models. This is not a surprise: the underlying estimator, in this case, is a difference-in-means rather than a diff-in-diffs because the second difference is by construction equal to zero. The bottom panel reports the estimated coefficients on the subsidized items and the other balance-sheet outcomes. The effect of the incentive on logged intangible assets turns out positive but not statistically diffrent from zero when estimated with the unconditional TWFE. This estimate may be biased because treated and control firms are not fully comparable. Indeed, in the other specifications by controlling for the firm's characteristics the coefficient grows in magnitude and becomes statistically different from zero. The result is confirmed also by employing the doubly robust estimator. For what concerns the log tangible assets and the log productivity the estimated effect is consistent with the baseline coefficient. The positive effect found on the log employment is confirmed in the DRIPW specification in the DRIPW estimator. Finally, the reported estimate for the the log value of production is only marginally different from zero in the DRIPW model.

7. Cost-effectiveness of the policy

This section evaluates the cost-effectiveness of the policy through a back-of-the-envelope calculation, with results presented in Table 10. Based on the baseline estimate, the average cost of the policy per unit of intangible asset over the 2009–2015 period is approximately €0.17, corresponding to an investment multiplier of 6.1. Under the most conservative robustness check, which takes into account the presence of outliers in the distribution of intangible assets, the cost rises to €0.5, reducing the multiplier to 1.8. Conversely, the least conservative estimate suggests a cost of €0.08 and a multiplier of 12. The most conservative estimate is in line with those reported in previous studies. The closest comparison is Chinetti (2023), which analyzes a regional R&D subsidy introduced in another lagging region of Southern Italy in 2009 and finds an investment multiplier of 1.5 for intangible assets. Other studies focusing on more developed regions in Northern Italy (e.g., Bronzini and Iachini 2014 and Manaresi et al. 2025) report even lower multipliers, close to 1. Variations in the cost-effectiveness between studies may stem from differences in program design, such as lower aid intensity or a stronger emphasis on innovation adoption rather than industrial research.

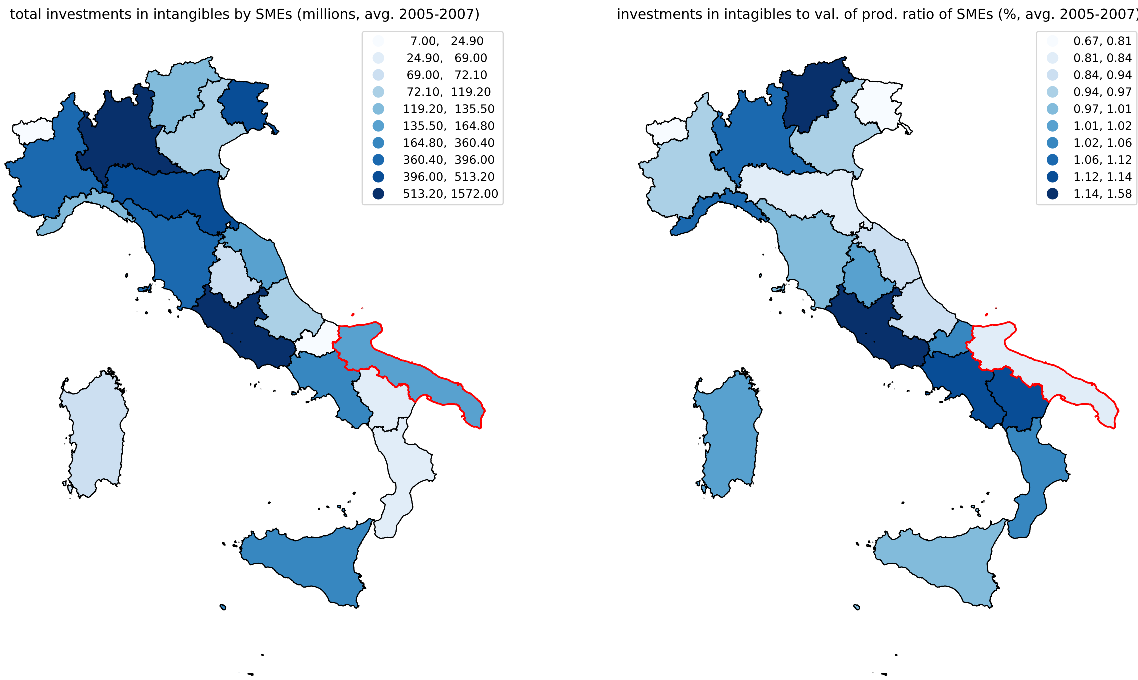
Regarding employment effects, although the estimates are only marginally significant, the cost per new job is found to range between €23,600 and €34,900. These values are lower than those reported by Cingano et al. (2022), who examines a large-scale investment program in Italy and estimates a cost of approximately €225,000 per new job in Southern Italy. However, they align more closely with other place-based policies—for instance, Siegloch et al. (2025) reports a cost of around €30,000 per new job in Germany. When accounting for heterogeneity between firms receiving a single subsidy and those receiving multiple subsidies, the analysis reveals a lower multiplier for intangible assets (3.8) and a higher cost per new job (€28,000) among the former group.

8. Conclusions

This paper provides novel empirical evidence of the effectiveness of public support aimed at stimulating private innovation activity in a lagging behind region in the aftermath of the Great Recession. In particular, I study a regional subsidy program implemented in Apulia, in South of Italy, between 2008-2009 which targeted small and medium enterprises. I improve on the existing literature by exploiting a unique administrative data source on the payments made by the regional administration to firms to precisely pin down treated firms and to separate other confounding regional policies. By comparing subsidized and eligible firms with a recent match-in-diff estimation strategy, I show that the policy was indeed effective in stimulating additional investments in innovation, as treated firms increased their intangible assets relative to control firms. In particular, the effect is bell-shaped: stronger during the project implementation years, and weaker towards the end. Although investments in tangible assets were among the eligible expenditure items for the subsidy, I find no effect of the policy on this outcome. Nevertheless, I find that the subsidy had a positive effect on employment which can be thought as both a direct and indirect effect of the policy. Moreover, treated firms did not show any productivity premium over untreated firms. That is, the rise in the value of production has been compensated by the growth of employment. The result is also confirmed for other proxies of productivity such as the cost of labor per employee and TFP. By leveraging a unique administrative dataset that collects information about various Apulian incentives for private investments implemented between 2010 and 2015, I divided the treated firms into two groups: those that received one subsidy and those that received more than one subsidy during the period under analysis. I show that although the additionality of the subsidy is confirmed, the estimated impact on employment turns out to be not statistically significant for the ”one subsidy” firms. The null effect of the policy on productivity is confirmed for both groups. This analysis highlights the importance of taking into account confounding factors when dealing a policy evaluation exercise. More importantly, from a policy perspective, these findings indicate that this type of innovation incentive did not lead to increased productivity compared to unsubsidized firms, even for those that received multiple incentives. However, several important questions remain open for future research. These include the long-term impact of the grants on firm performance and the general equilibrium effects. While this is a common limitation in the program evaluation literature, it is possible that the policy generated spillover effects impacting firms outside the sample. These indirect effects, particularly knowledge spillovers, could play a crucial role in amplifying the program's overall impact. Investigating these aspects would provide a more comprehensive understanding of the policy's effectiveness, although doing so poses significant empirical challenges.

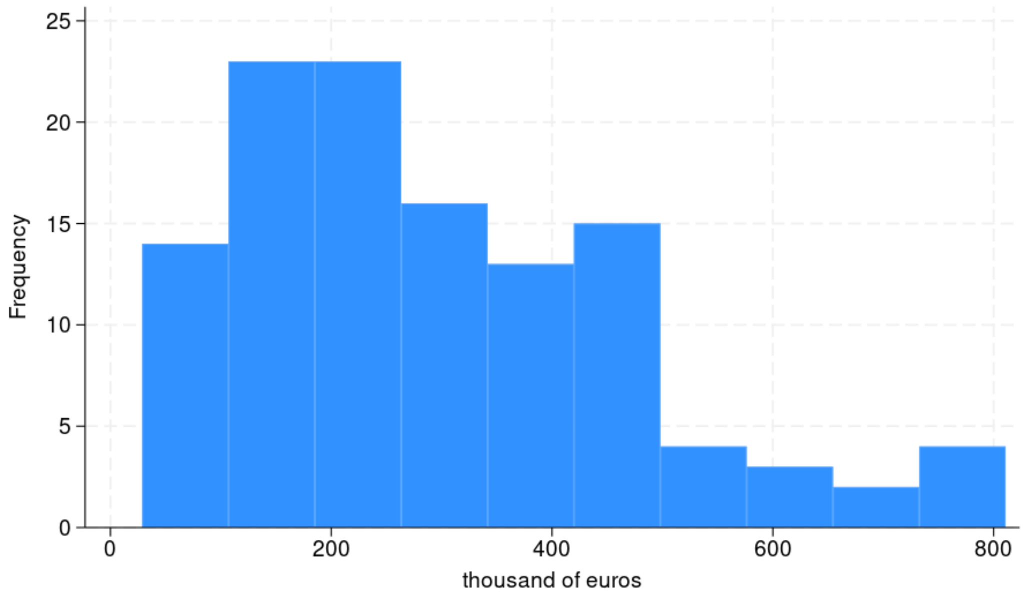
Figures

Figure 1: Investments in intangible assets in Italy



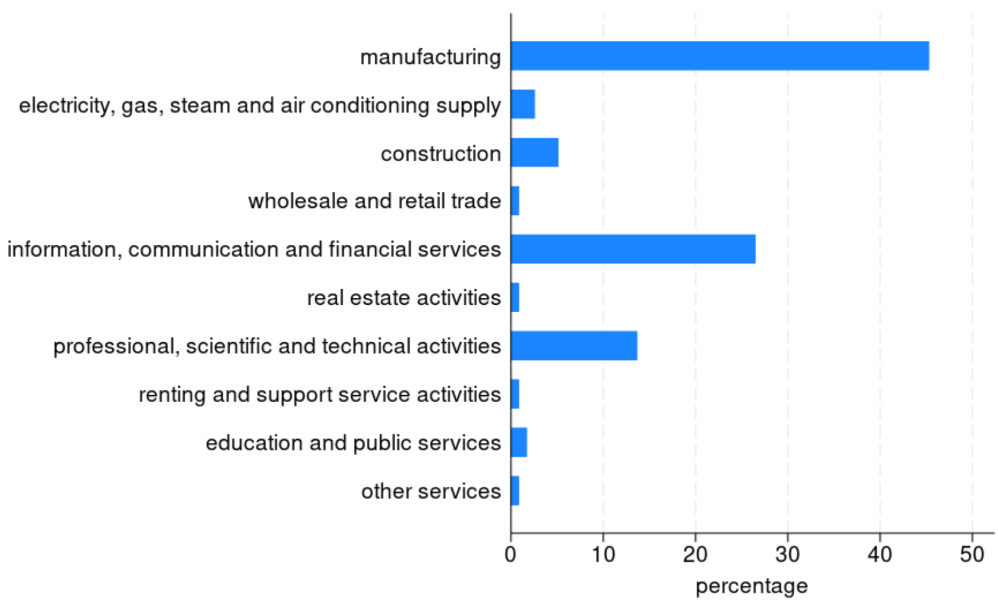
Note: the left panel shows the regional distribution of the average, over the period 2005-2007, total investments in intangible assets by small and medium limited companies. The right panel shows the regional distribution of the average, over the period 2005-2007, incidence of investments in intangible assets over the value of production of small and medium enterprises. Apulia region is highlighted by the red borders. Source: Cerved data.

Figure 2: Subsidy distribution



Note: the figure shows the distribution of the subsidies effectively payed to firms.

Figure 3: Sectorial distribution of firms



Note: the figure shows the distribution of payed firms over nace-1dgt sectors.

Figure 4: Timeline of the policy

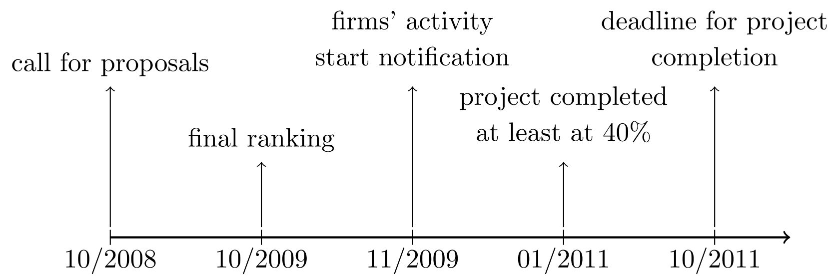
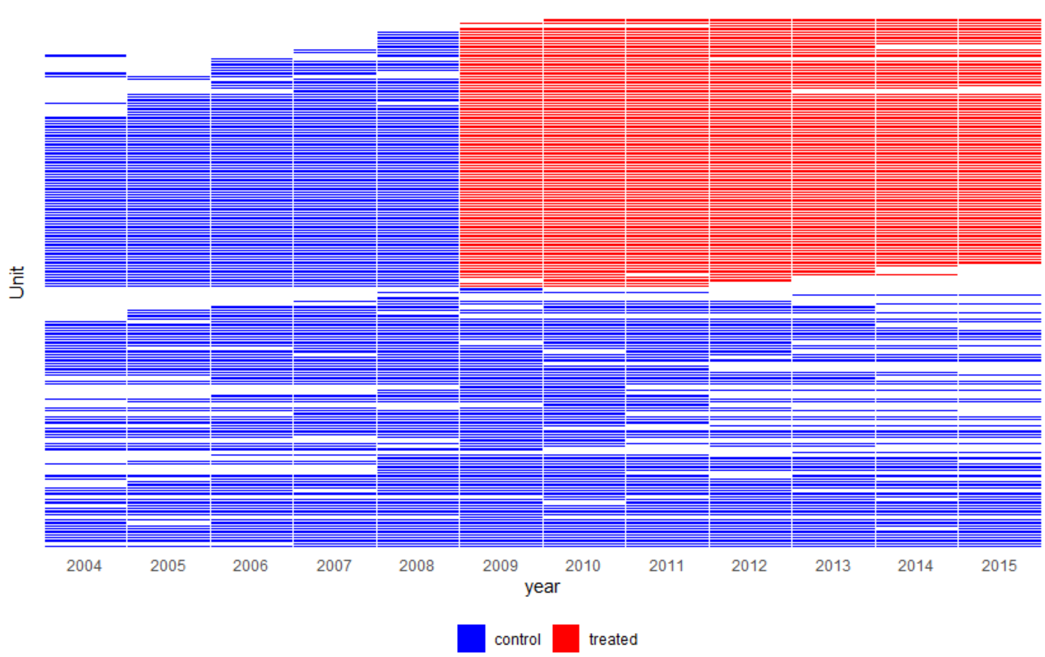
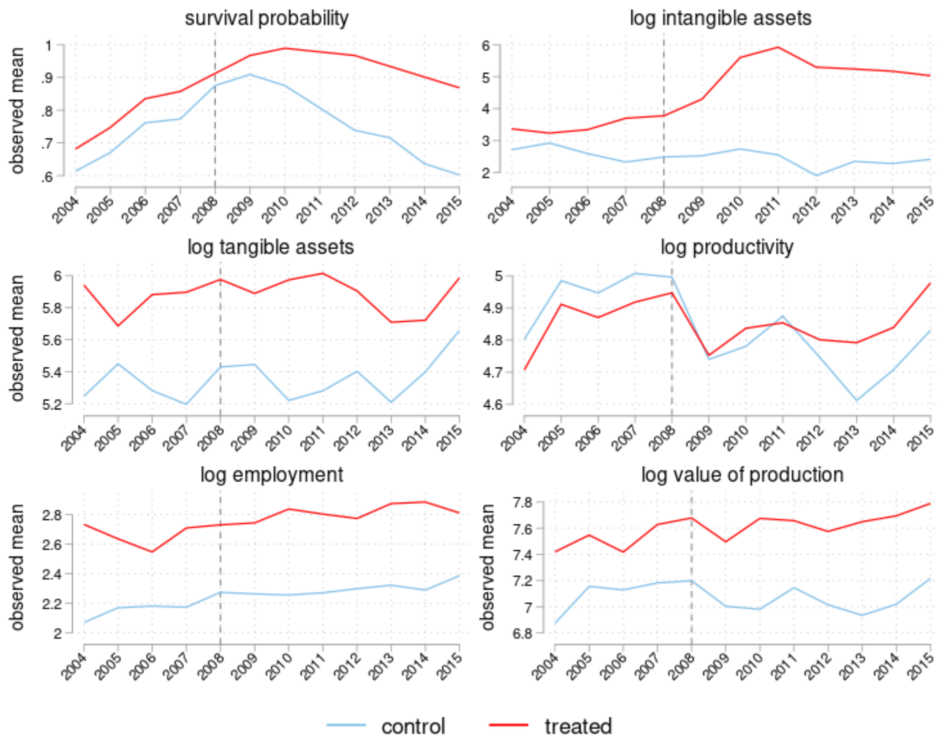


Figure 5: Treatment distribution



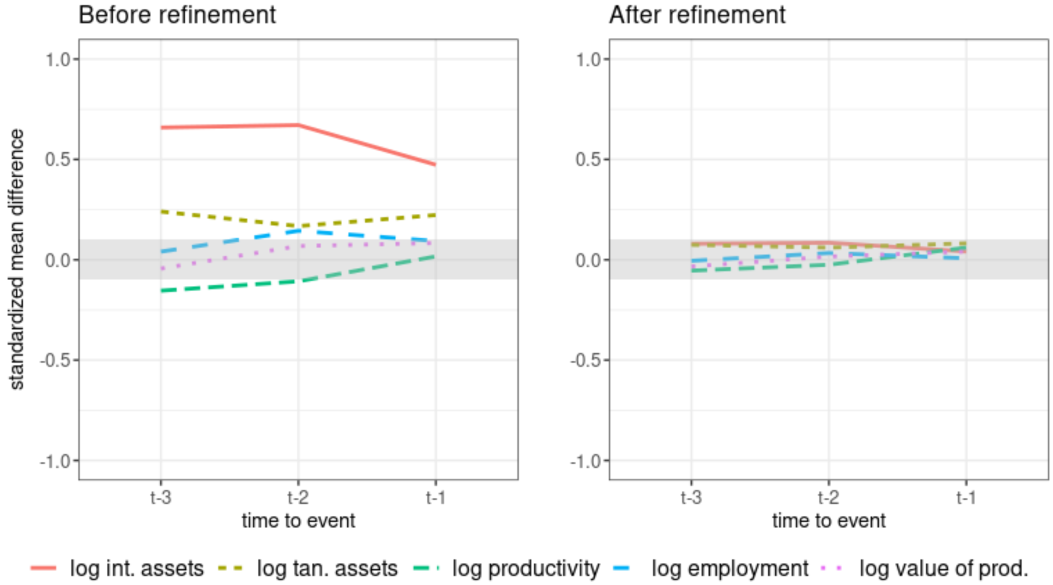
Note: the figure shows the balance sheets data observed in the sample for each firm between 2004 and 2015 by treatment status. The treatment date is set to 2009. Blank spaces correspond to missing balance sheets. The total number of firms, which are depicted on the vertical axis, is 179 out of which 91 treated.

Figure 6: Outcomes' dynamics by treatment group



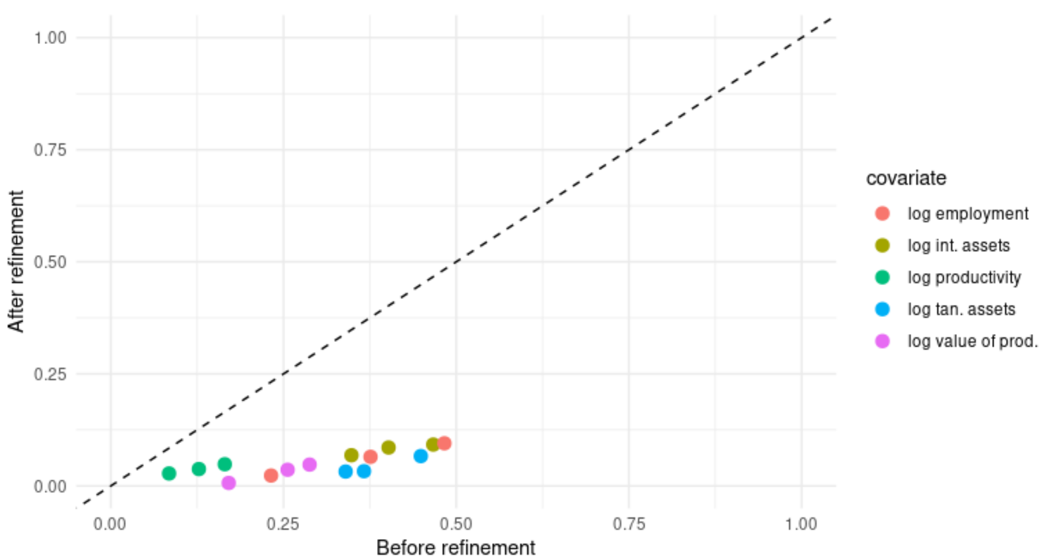
Note: the figure shows the evolution over time of the average outcomes by treatment group in the unbalanced panel. The six outcomes are the probability of observing the balance sheet of the firms, the log of intangible assets, the lof of tangible assets, the log of productivity expressed as the ratio between the value of production and employment, the log of employment, and the log of the value of production.

Figure 7: Outcomes' balance before and after refinement



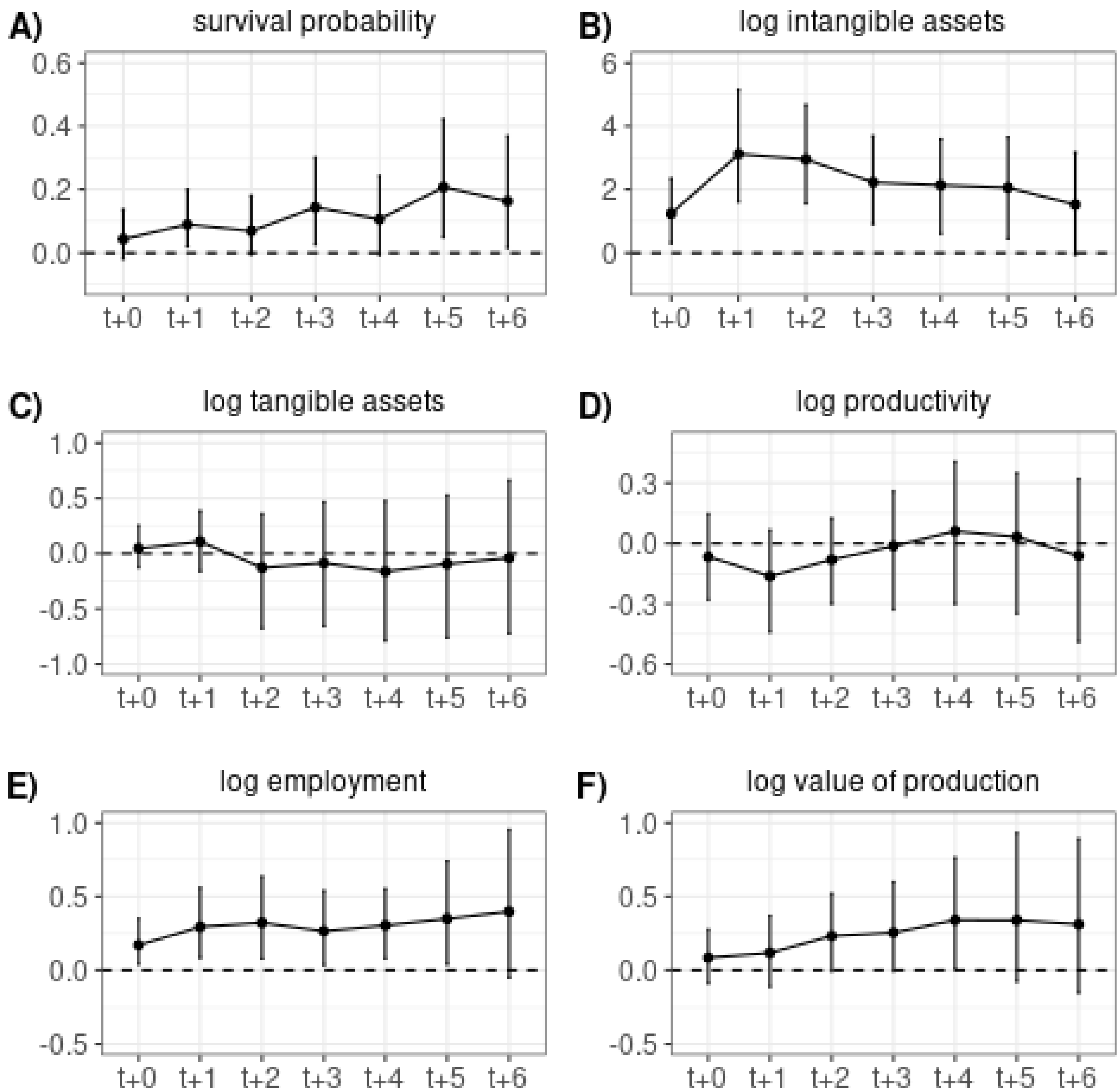
Note: the figure shows on the left panel the SMDs in the pre-treatment period for the four outcomes (log intangible assets, log employment, log productivity and log value of production) before matching. The right panel shows the same indicator after the matching procedure. The grey area marks the 0.1 circle inside which an outcome can be considered fairly balanced between the two groups.

Figure 8: Outcomes' balance - scatter plot



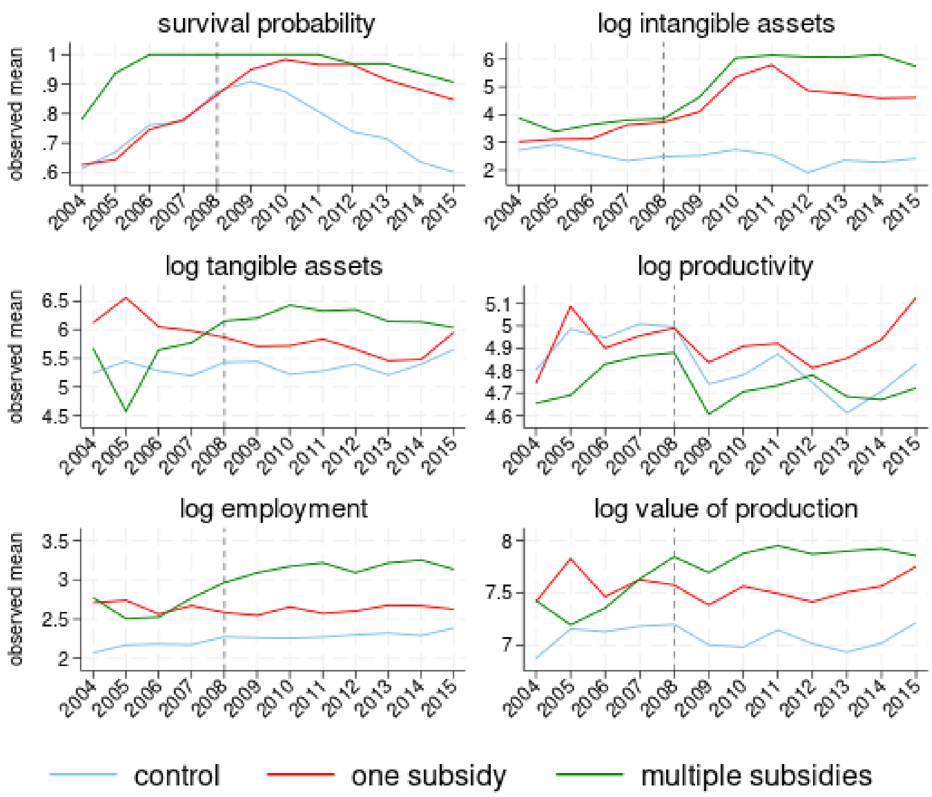
Note: the figure shows a scatter plot in which each point is the SMD of an outcome in a pre-treatment year before (on the horizontal axis) and after (on the vertical axis) the refinement procedure. Points below the 45° line are the one for which the balance improved after performing the matching.

Figure 9: Event study estimates for the baseline sample



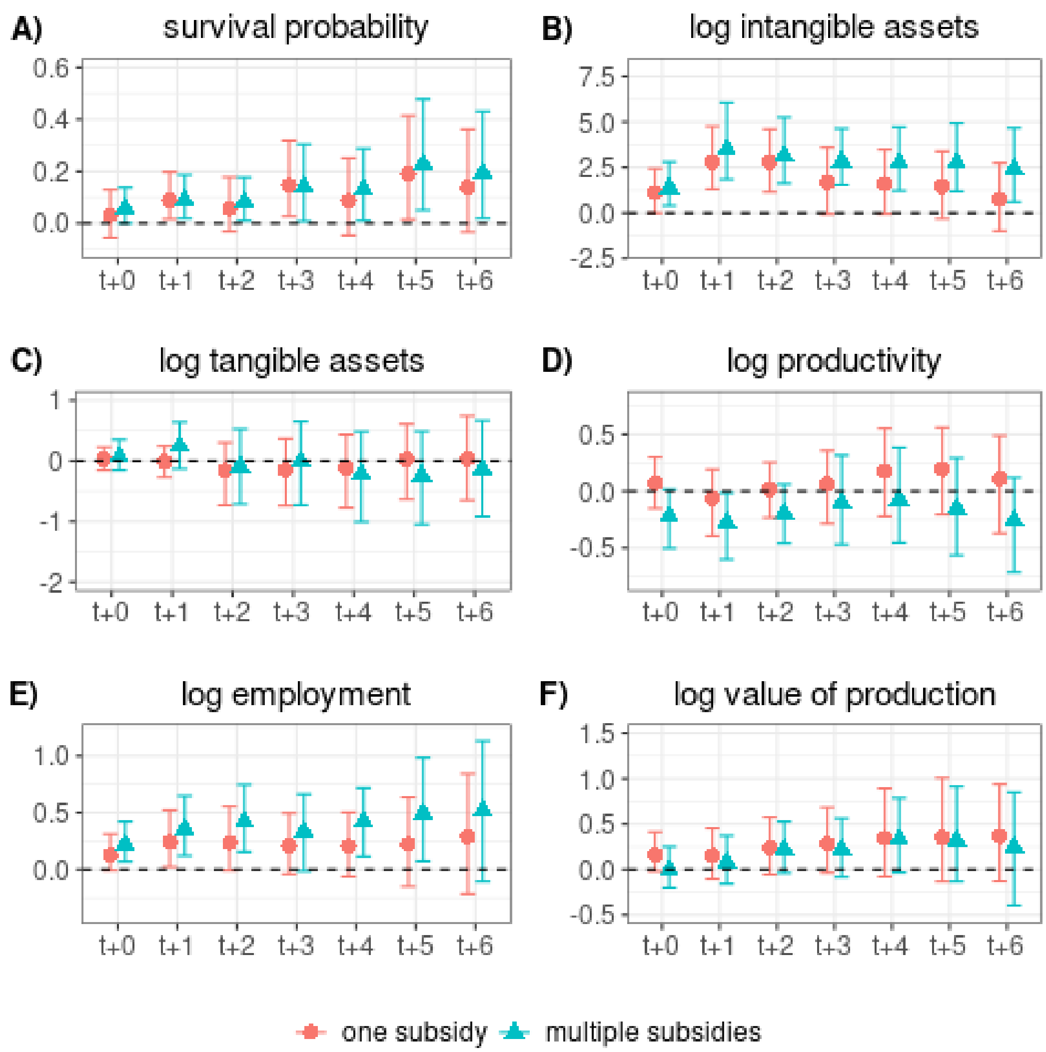
Note: the figure shows the event study estimates over the period 2009-2015 for the following outcomes: the survival probability, the log of intangible assets, the log of tangible assets, the log of productivity, the log of employment, and the log value of production. The vertical lines are the bootstrapped 95% confidence interval. The sample is composed by a balanced panel of 95 firms (62 treated). The estimates has been produced with the R's statistical software package PanelMatch (see Imai et al. 2023).

Figure 10: Outcomes' dynamics by treatment group - One subsidy vs Multiple subsides



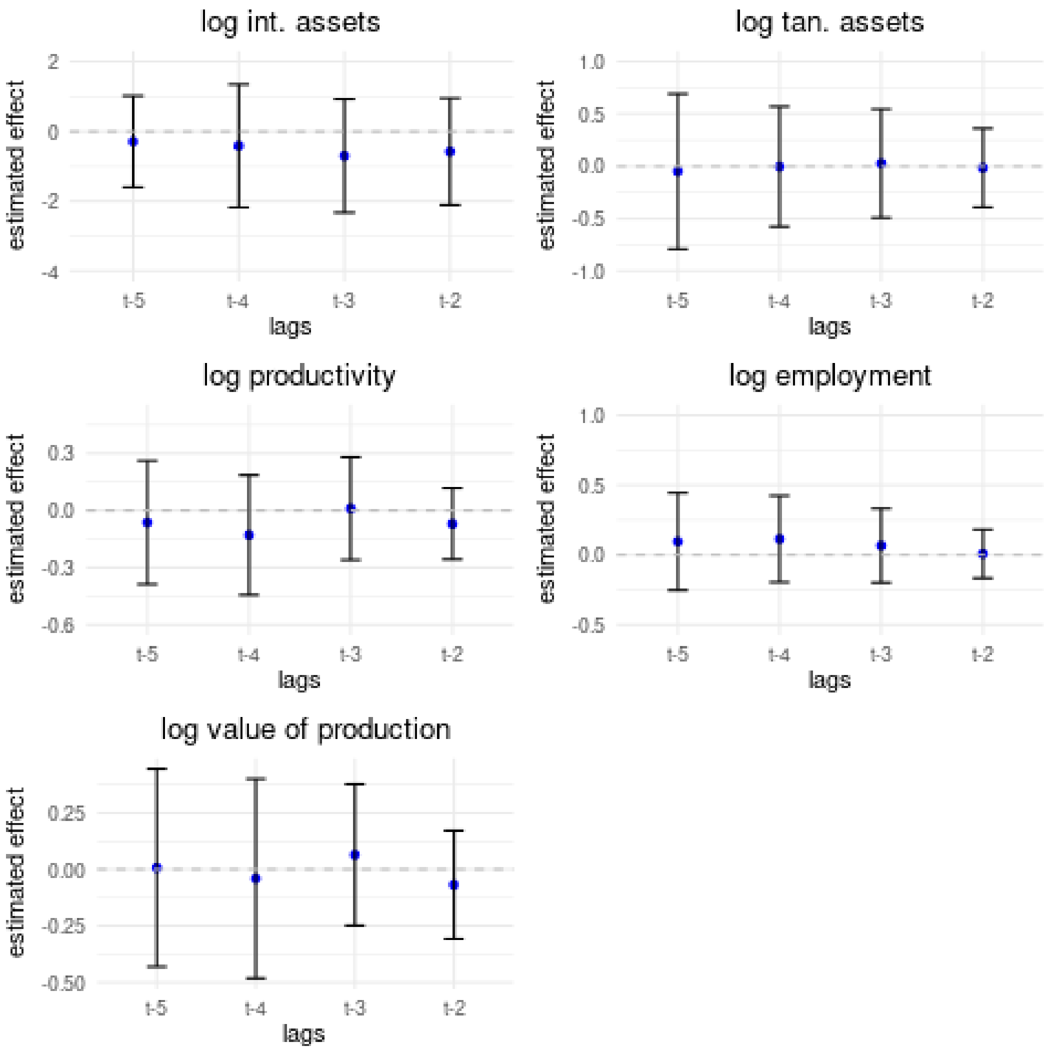
Note: the figure shows the evolution over time of the average outcomes by type of treatment. The "one subsidy" group refers to all firms which in the period under analysis obtained only one R&D subsidy from the Apulian regional administration. The "multiple subsidies" group refers instead to those firms which also got other R&D subsides in the same period. The six outcomes are the survival probability, the log of intangible assets, the log of tangible assets, the log of productivity, the log of employment, and the log value of production. log intangible assets, log productivity, log employment and log value of production.

Figure 11: Event study estimates - One subsidy vs Multiple subsidies



Note: the figure shows the event study estimates over the period 2009-2015 for the two groups of "one subsidy" and "multiple subsidies" firms. The "one subsidy" group refers to all firms which in the period under analysis obtained only one R&D subsidy from the Apulian regional administration. The "multiple subsidies" group refers instead to those firms which also got other R&D subsides in the same period. The outcomes analysed are: the survival probability, the log of intangible assets, the log of tangible assets, the log of productivity, the log of employment, and the log value of production. log intangible assets, log productivity, log employment and log value of production. The vertical lines are the bootstrapped 95% confidence interval. The estimates has been produced with the R's statistical software package PanelMatch (see Imai et al. 2023).

Figure 12: Event study estimates - placebo test



Note: the figure shows the placebo estimates for the pre-treatment periods. The matching procedure has been carried out only on lags $t-5$ and $t-4$. The outcomes analysed are: the log of intangible assets, the log of tangible assets, the log of productivity, the log of employment, and the log value of production. The vertical lines are the bootstrapped 95% confidence interval. The estimates has been produced with the R's statistical software package PanelMatch (see Imai et al. 2023).

Tables

Table 1: Subsidy payments by year and type

Year	*Advance payment*	*Project at 50%*	*Second half*	*Total*	*Cumulative*	*Budget %*
2010	4,148	1,895	-	6,043	6,043	13
2011	257	8,538	1,556	10,351	16,394	34
2012	-	1,947	10,213	12,160	28,554	59
2013	-	-	6,440	6,440	34,994	73
Firms' total declared costs						64,199
Average project's costs						548
Average subsidy						299

The table reports information about the payments made by the regional administration in relation to the subsidy by year and type. The amounts are expressed in thousand of euros. Firms had the possibility to get an anticipation of payment, or choose to get the subsidy in two tranches or in a single solution on the basis of the sustained costs. The total budget assigned to the policy was 48 millions of euros.

Table 2: Causes of balance sheets missingness

	Control	Treated	**Row total**
Administrative cancellation	4	0	4
Bankruptcy	2	1	3
Merger or split	4	2	6
Liquidation	11	5	16
Transfer of registered office	2	2	4
Other causes	1	0	1
Unknown	11	2	13
Column total	**35**	**12**	**47**

The table shows the number of firms by cause of balance sheet absence. The row "unknown" refers to all the cases where the author was not able to retrieve the reason for the missingness of firms' balance sheets. The information has been collected via the Infocamere dataset.

Table 3: Pre-treatment summary statistics - Unbalanced Panel

	Control Mean(SD)	Treated Mean(SD)	Std. mean difference
Intangible assets	251.01	212.90	-0.08
	(542.25)	(366.83)	
Tangible assets	1008.72	1467.84	0.21
	(2,166.28)	(2,196.85)	
Total assets	1457.27	1931.90	0.18
	(2,760.27)	(2,625.50)	
% of int./tot. assets	0.20	0.20	0.02
	(0.24)	(0.24)	
Num. of employees	18.20	30.94	0.34
	(26.78)	(46.05)	
Value of production	3437.06	5420.09	0.21
	(5,575.23)	(12,320.29)	
Productivity	261.40	202.70	-0.15
	(477.41)	(265.55)	
Return on assets	4.84	5.60	-0.09
	(10.01)	(7.38)	
Leverage	40.72	42.00	0.04
	(32.93)	(31.08)	
Industry	0.55	0.57	0.05
	(0.50)	(0.50)	

The table shows the pre-treatment averages (2004-2008) of firms' characteristics in an unbalanced panel of 179 firms (91 treated). The third column reports the standardized mean differences which are computed as: $\frac{\mu_t - \mu_c}{\sqrt{(std_t^2 + std_c^2)/2}}$. A SMD less than 0.1 is a sign of good balance between the groups. Intangible assets, total assets, the value of production and productivity are expressed in thousand of euros.

Table 4: Estimation results

Outcomes:	Survival probability	Log int. assets	Log tan. assets	Log productivity	Log employment	Log value of prod.
	(1)	(2)	(3)	(4)	(5)	(6)
			(a) Pooled			
ATT	0.117	2.173	-0.051	-0.042	0.301	0.241
	[0.029;0.242]	[1.005;3.580]	[-0.541;0.349]	[-0.303;0.190]	[0.067;0.542]	[-0.033;0.553]
			(b) Event study			
T+0	0.042	1.236	0.047	-0.066	0.170	0.086
	[-0.021;0.136]	[0.297;2.337]	[-0.127;0.253]	[-0.280;0.146]	[0.039;0.350]	[-0.088;0.277]
T+1	0.089	3.110	0.107	-0.163	0.295	0.119
	[0.020;0.199]	[1.601;5.147]	[-0.160;0.383]	[-0.439;0.065]	[0.084;0.558]	[-0.110;0.368]
T+2	0.068	2.948	-0.128	-0.081	0.323	0.233
	[-0.006;0.180]	[1.551;4.667]	[-0.674;0.357]	[-0.301;0.126]	[0.080;0.636]	[-0.006;0.520]
T+3	0.143	2.215	-0.087	-0.013	0.266	0.255
	[0.026;0.300]	[0.901;3.681]	[-0.656;0.462]	[-0.329;0.259]	[0.035;0.539]	[-0.005;0.596]
T+4	0.106	2.128	-0.162	0.060	0.305	0.340
	[-0.008;0.243]	[0.588;3.581]	[-0.788;0.476]	[-0.302;0.406]	[0.081;0.551]	[0.006;0.762]
T+5	0.207	2.052	-0.095	0.032	0.349	0.339
	[0.049;0.422]	[0.434;3.647]	[-0.763;0.525]	[-0.350;0.351]	[0.039;0.738]	[-0.074;0.929]
T+6	0.162	1.517	-0.042	-0.061	0.397	0.312
	[0.015;0.368]	[-0.067;3.160]	[-0.721;0.660]	[-0.490;0.320]	[-0.049;0.951]	[-0.155;0.891]
Firms	133	98	98	98	98	98
Obs.	1,330	980	980	980	980	980

Bootstrapped 95% confidence interval in parentheses. Panel (a) reports the estimated coefficients pooled over the entire lead window. Panel (b) reports the estimated coefficients for the event study from 2009 to 2015. The estimates have been produced with the R's statistical software package PanelMatch (see Imai et al. 2023).

Table 5: SMDs by number of subsidies

	Baseline (1)	One Subsidy (2)	Multiple Sub. (3)
Intangible assets	-0.08	-0.19	0.04
Tangible assets	0.21	0.27	0.11
Total assets	0.18	0.24	0.08
% of int./tot. assets	0.02	-0.17	0.26
Num. of employees	0.34	0.33	0.38
Value of production	0.21	0.29	-0.00
Productivity	-0.15	-0.09	-0.25
Return on assets	0.09	0.07	0.12
Leverage	0.04	-0.05	0.17
Industry	0.05	0.13	-0.10

The table shows the SMDs of firms' pre-treatment characteristics, averaged over the period 2004-2008, with respect to a control group of never treated firms. The first column reports the SMDs for the baseline sample, while the second column for the "one subsidy" group and the third column for the "multiple subsidies" one. The "one subsidy" group refers to all firms which in the period under analysis obtained only one R&D subsidy from the Apulian regional administration. The "multiple subsidies" group refers instead to those firms which also got other R&D subsides in the period (2009-2015). The standardized mean differences are computed as: $\frac{\mu_t - \mu_c}{\sqrt{(std_t^2 + std_c^2)/2}}$. A SMD less than 0.1 is a sign of good balance between the groups. Intangible assets, total assets, the value of production and productivity are expressed in thousand of euros.

Table 6: Estimation results - One subsidiy vs Multiple subsidies

Outcomes:	Survival probability	Log int. assets	Log tan. assets	Log productivity	Log employment	Log value of prod.
	(1)	(2)	(3)	(4)	(5)	(6)
(a) Baseline						
ATT	0.117	2.173	-0.051	-0.042	0.301	0.241
	[0.029;0.242]	[1.005;3.580]	[-0.541;0.349]	[-0.303;0.190]	[0.067;0.542]	[-0.033;0.553]
Firms	133	98	98	98	98	98
Obs.	1330	980	980	980	980	980
(b) One subsidy						
ATT	0.105	1.743	-0.050	0.081	0.218	0.273
	[0.009;0.246]	[0.361;3.289]	[-0.522;0.350]	[-0.199;0.375]	[-0.028;0.510]	[-0.012;0.639]
Firms	101	69	69	69	69	69
Obs.	1010	690	690	690	690	690
(c) Multiple subsidies						
ATT	0.131	2.676	-0.053	-0.182	0.394	0.202
	[0.035;0.279]	[1.364;4.236]	[-0.564;0.454]	[-0.501;0.108]	[0.132;0.731]	[-0.069;0.541]
Firms	93	64	64	64	64	64
Obs.	930	640	640	640	640	640

Bootstrapped 95% confidence interval in parentheses. The table reports the pooled coefficients over the entire lead window (2009-2015). Panel (a) reports the effects for the baseline sample, panel (b) for the ”one subsidy” group and panel (c) for the ”multiple subsidies” one. The ”one subsidy” group refers to all firms which in the period under analysis obtained only one R&D subsidy from the Apulian regional administration. The ”multiple subsidies” group refers instead to those firms which also got other R&D subsides in the period (2009-2015). The estimates have been produced with the R’s statistical software package PanelMatch (see Imai et al. (2023)).

Table 7: Robustness checks - Different matching models and control group

	Baseline (1)	Alternative control group (2)	Mahalanobis distance (3)	Additional covariates (4)	L = 4 (5)	L = 5 (6)	Match on t-5 and t-4 (7)	Winsorization (8)
				panel (a)				
Survival prob.	0.117	0.176	0.097	0.121	0.104	0.079	0.060	-
	[0.029;0.242]	[0.121;0.250]	[0.023;0.196]	[0.024;0.242]	[0.006;0.221]	[-0.020;0.210]	[-0.019;0.164]	-
Average Balance	0.05	0.00	0.05	0.04	0.05	0.04	0.10	-
% Improvement	83.23	99.99	85.35	86.43	84.11	87.59	69.42	-
Firms	133	10,690	133	133	118	103	103	-
Observations	1,330	106,900	1,330	1,330	1,298	1,236	1,236	-
				panel (b)				
Log int. assets	2.173	2.798	1.781	2.309	2.189	2.494	2.579	1.217
	[1.021;3.619]	[2.107;3.635]	[0.650;3.030]	[0.930;4.202]	[0.749;3.964]	[0.655;5.204]	[0.855;5.218]	[0.512;1.916]
Log tan. assets	-0.051	0.204	0.059	0.017	-0.077	-0.073	-0.137	-0.055
	[-0.497;0.350]	[-0.001;0.417]	[-0.267;0.339]	[-0.391;0.397]	[-0.477;0.293]	[-0.608;0.380]	[-0.630;0.305]	[-0.556;0.338]
Log productivity	-0.042	-0.013	0.019	-0.043	0.068	0.068	0.101	-0.025
	[-0.287;0.188]	[-0.159;0.125]	[-0.152;0.206]	[-0.309;0.187]	[-0.117;0.274]	[-0.143;0.278]	[-0.108;0.343]	[-0.191;0.162]
Log employment	0.301	0.232	0.230	0.276	0.161	0.139	0.109	0.324
	[0.070;0.548]	[0.107;0.371]	[0.023;0.463]	[0.055;0.517]	[-0.122;0.426]	[-0.095;0.422]	[-0.149;0.374]	[0.121;0.565]
Log value of prod.	0.241	0.219	0.239	0.222	0.203	0.172	0.201	0.206
	[-0.011;0.529]	[0.086;0.372]	[0.004;0.494]	[-0.001;0.484]	[-0.119;0.541]	[-0.138;0.547]	[-0.094;0.577]	[-0.081;0.502]
Average Balance	0.05	0.00	0.05	0.06	0.06	0.08	0.07	0.06
% Improvement	77.68	99.99	77.52	75.62	75.33	68.14	70.78	72.4
Firms	98	5,452	98	98	90	78	78	98
Observations	980	54,520	980	980	990	936	936	980

Bootstrapped 95% confidence interval in parentheses. The table reports the pooled estimates over 2009-2015 for each model. The first column reports the estimates for the baseline sample. Column 2 reports the estimates for a model in which the control group of firms unrelated to the policy under study. Column 3 uses as matching algorithm the minimization of the mahalanobis distance. Column 4 adds in **Z** the ROA and the leverage. In column 5 the matching is performed on four lags (2004-2008), while in column 6 five lags are used (2006-2008). Column 7 uses again five lags but the matching procedure is performend only on lags $t-5$ and $t-4$. The last column reports the coefficient for the winsorized outcomes. The average balance rows report the average standardized mean difference of **Z** after the matching procedure. The improvement rows show the percentage improvement in the average standardized mean difference over the pre-matching value. The estimates have been produced with the R's statistical software package PanelMatch (see Imai et al. 2023).

Table 8: Robustness checks - Additional outcomes

	Baseline	Alternative control group	Mahalanobis distance	Additional covariates	L = 4	L = 5	Match on t-5 and t-4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log cumulative	1.092	1.130	0.959	1.056	0.951	0.988	0.936
int. investments	[0.633;1.484]	[0.809;1.439]	[0.425;1.399]	[0.516;1.497]	[0.228;1.469]	[0.269;1.525]	[0.416;1.445]
Log cumulative	0.146	0.113	0.174	0.204	0.094	0.092	0.007
tan. investments	[-0.294;0.481]	[-0.118;0.309]	[-0.237;0.479]	[-0.199;0.508]	[-0.274;0.384]	[-0.399;0.394]	[-0.504;0.349]
Value added	-0.741	1.617	0.144	-1.241	1.888	3.747	2.506
per employee	[-10.989;8.380]	[-4.889;8.033]	[-7.837;7.079]	[-11.811;6.705]	[-6.982;9.896]	[-4.232;10.744]	[-5.142;9.285]
Log cost of labour	-0.001	0.027	0.015	-0.010	-0.030	0.015	0.021
per employee	[-0.087;0.068]	[-0.012;0.063]	[-0.067;0.087]	[-0.094;0.065]	[-0.140;0.048]	[-0.064;0.075]	[-0.050;0.080]
TFP	-0.037	-0.058	0.002	0.013	0.030	-0.045	0.003
	[-0.202;0.118]	[-0.172;0.049]	[-0.132;0.160]	[-0.188;0.199]	[-0.124;0.224]	[-0.223;0.116]	[-0.210;0.208]
Average Balance	0.07	0.03	0.07	0.07	0.07	0.13	0.11
% Improvement	72.32	97.12	72.65	68.98	75.97	57.13	63.47
Firms	98	5,452	98	98	90	78	78
Observations	980	54,520	980	980	990	936	936

Bootstrapped 95% confidence interval in parentheses. The table reports the pooled estimates over 2009-2015 for each model. In the baseline model **Z** includes the log of cumulative int. investments, the value-added per employee and the log of employment. The first column reports the estimates for the baseline sample. Column 2 reports the estimates for a model in which the control group of firms unrelated to the policy under study. Column 3 uses as matching algorithm the minimization of the mahalanobis distance. Column 4 adds in **Z** the ROA and the leverage. In column 5 the matching is performed on four lags (2004-2008), while in column 6 five lags are used (2006-2008). Column 7 uses again five lags but the matching procedure is performed only on lags $t-5$ and $t-4$. The average balance rows report the average standardized mean difference of **Z** after the matching procedure. The improvement rows show the percentage improvement in the average standardized mean difference over the pre-matching value. The average balance rows report the average standardized mean difference of **Z** after the matching procedure. The improvement rows show the percentage improvement in the average standardized mean difference over the pre-matching value. The estimates have been produced with the R's statistical software package PanelMatch (see Imai et al. 2023).

Table 9: Robustness checks - Alternative estimators

	Baseline (1)	TWFE (2)	Cond. TWFE (3)	DR IPW (4)
		panel (a)		
Survival prob.	0.127	0.128	0.135	0.127
	[0.024;0.260]	[0.086;0.170]	[0.056;0.214]	[0.032;0.222]
Firms	133	133	133	133
Observations	1.330	1.330	1.330	1.330
		panel (b)		
Log int. assets	2.086	1.108	2.168	1.915
	[0.904;3.489]	[-0.453;2.668]	[0.832;3.504]	[0.870;2.960]
Log tan. assets	0.035	0.079	-0.001	0.191
	[-0.351;0.396]	[-0.324;0.481]	[-0.410;0.409]	[-0.285;0.666]
Log productivity	-0.046	0.062	0.048	-0.110
	[-0.300;0.189]	[-0.189;0.313]	[-0.158;0.254]	[-0.362;0.143]
Log employment	0.291	0.114	0.195	0.337
	[0.074;0.555]	[-0.223;0.451]	[-0.035;0.425]	[0.095;0.579]
Log value of prod.	0.228	0.173	0.240	0.222
	[-0.026;0.508]	[-0.183;0.529]	[-0.017;0.496]	[-0.002;0.446]
Firms	98	98	98	98
Observations	980	980	980	980

Bootstrapped 95% confidence interval in parentheses. The table reports the pooled estimates over the period 2009-2015. Column 1 reports the estimated coefficient from the model of Imai et al. (2023). Column 2 reports the estimated coefficient from a regression where each outcome is regressed against a time varying treatment indicator, firm and year fixed effects. In column 3 the previous specification is augmented by controlling for the interaction of the variables in **Z** with a firm specific linear trend. Column 4 reports the estimates for the doubly robust inverse probability weighting estimator (see Sant'Anna and Zhao 2020) by controlling for the interaction of the variables in **Z** with a firm specific linear trend. **Z** contains all the lags between 2006 and 2008 of the following variables: log of intangible assets, log of tangible assets, log of productivity, log of employment and log of value of production.

Table 10: Cost-effectiveness of the policy

	Cost per unit of intangible asset (1)	Cost per new employee (2)
	average over the period 2009-15	
Baseline estimate	0.165	25.7k
Most conservative estimate	0.541	34.9k
Least conservative estimate	0.083	23.6k
Heterogeneity:		
- one subsidy firms	0.261	28.0k
- multiple subsidies firms	0.099	26.0k

The table presents the cost per unit of intangible assets and the cost per new employee associated with the policy under analysis. All values are expressed in euros and refer to the average treated firm over the 2009–2015 period. The increases in intangible assets and employment are calculated relative to their respective average levels during the pre-treatment period (2006–2009). The row labeled "Most conservative estimate" reports values based on the smallest statistically significant coefficient obtained from the robustness checks, while the "Least conservative estimate" row reflects those derived from the largest statistically significant coefficient.

References

ACCETTURO, A. (2022): "Subsidies for innovative start-ups and firm entry," *Industrial and Corporate Change*, 31, 1202–1222.

ALI, M. S., R. H. H. GROENWOLD, W. R. PESTMAN, S. V. BELITSER, K. C. B. ROES, A. W. HOES, A. DE BOER, AND O. H. KLUNGEL (2014): "Propensity score balance measures in pharmacoepidemiology: a simulation study," *Pharmacoepidemiology and Drug Safety*, 23, 802–811.

ANTONELLI, C. AND F. CRESPI (2013): "The "Matthew effect" in R&D public subsidies: The Italian evidence," *Technological Forecasting and Social Change*, 80, 1523–1534.

ARRIGHETTI, A., F. LANDINI, AND E. BARTOLONI (2018): "Firm survival during economic downturns: is selecton based on cleansing or skill accumulation?" .

BECKER, B. (2015): "Public R&D policies and private R&D investment: a survey of the empirical evidence," *Journal of Economic Surveys*, 29, 917–942.

BELITSER, S. V., E. P. MARTENS, W. R. PESTMAN, R. H. GROENWOLD, A. DE BOER, AND O. H. KLUNGEL (2011): "Measuring balance and model selection in propensity score methods," *Pharmacoepidemiology and Drug Safety*, 20, 1115–1129.

BERTAMINO, F., R. BRONZINI, M. D. MAGGIO, AND D. REVELLI (2016): "Local policies for innovation: the case of technology districts in Italy," *Questioni di Economia e Finanza (Occasional Papers)*.

BOCCI, C., A. CALOFFI, AND M. M. ET AL. (2023): "Evaluating Public Support to the Investment Activities of Business Firms: A Multilevel Meta-Regression Analysis of Italian Studies," *Ital Econ J*, 1–34.

BRANSON, Z. AND F. MEALLI (2019): "The Local Randomization Framework for Regression Discontinuity Designs: A Review and Some Extensions," .

BRONZINI, R. AND E. IACHINI (2014): "Are Incentives for R&D Effective? Evidence from a Regression Discontinuity Approach," *American Economic Journal: Economic Policy*, 6, 100–134.

BRONZINI, R. AND P. PISELLI (2016): "The impact of R&D subsidies on firm innovation," *Research Policy*, 45, 442–457.

CALONICO, CATTANEO, F. AND TITIUNIK (2017): "rdrobust: Software for Regression Discontinuity Designs." *Stata Journal 17*, 372–404.

CATTANEO, M. D. AND R. TITIUNIK (2022): "Regression Discontinuity Designs," *Annual Review of Economics*, 14, 821–851.

Cerqua, A. and G. Pellegrini (2013): "Beyond the SUTVA: How Industrial Policy Evaluations Change When We Allow for Interactions Among Firms," *Sapienza University of Rome School of Economics Working Paper No. 15.*

——— (2020): "Evaluation of the Effectiveness of Firm Subsidies in Lagging-Behind Areas: The Italian Job," *Scienze Regionali, Italian Journal of Regional Science*, 477–500.

Chinetti, S. (2023): "Investment Subsidies Effectiveness: Evidence from a Regional Program," *Italian Economic Journal*, 9, 723–759.

Cingano, F., F. Palomba, P. Pinotti, and E. Rettore (2022): "Make subsidy work: rules vs." *Temi di discussione (Economic working papers).*

Fantino, D. and G. Cannone (2013): "Evaluationg the efficacy of European regional funds for R&D," *Temi di discussione (economic working papers)*, 902.

Grembi, V., T. Nannicini, and U. Troiano (2016): "Do Fiscal Rules Matter?" *American Economic Journal: Applied Economics*, 8, 1–30.

Hall, B. H. and N. Rosenberg (2010): "Chapter 14 - The Financing of R&D and Innovation," in *Handbook of The Economics of Innovation, Vol. 1*, North-Holland, vol. 1 of *Handbook of the Economics of Innovation*, 609–639.

Ham, D. W. and L. Miratrix (2024): "Benefits and costs of matching prior to a Difference in Differences analysis when parallel trends does not hold," .

Hud, M. and K. Hussinger (2015): "The impact of R&D subsidies during the crisis," *Research Policy*, 44, 1844–1855.

Imai, K. and I. S. Kim (2021): "On the Use of Two-Way Fixed Effects Regression Models for Causal Inference with Panel Data," *Political Analysis*, 29, 405–415.

Imai, K., I. S. Kim, and E. H. Wang (2023): "Matching Methods for Causal Inference with Time-Series Cross-Sectional Data," *American Journal of Political Science*, 67, 587–605.

Imai, K. and M. Ratkovic (2014): "Covariate balancing propensity score," *Journal of the Royal Statistical Society*, 243–263.

Liberati, D., M. Marinucci, and G. M. Tanzi (2016): "Science and technology parks in Italy: main features and analysis of their effects on the firms hosted," *The Journal of Technology Transfer*, 41.

Lotti, F. and M. L. Stefani (2014): "Le iniziative regionali per favorire l'innovazione d'impresa," *Questioni di Economia e Finanza (Occasional Papers).*

MANARESI, F., E. RUSSO, AND P. SANTOLERI (2025): "Some causal effect of a place-based innovation policy," *forthcoming*.

MOLLER, L. AND F. LOTTI (2013): "Female Entrepreneurship and Government Policy: Evaluating the Impact of Subsidies on Firms' Survival," .

NELSON, R. R. (1959): "The Simple Economics of Basic Scientific Research," *Journal of Political Economy*, 67, 297–306.

ROMER, P. M. (1986): "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, 94, 1002–1037.

ROTH, J., P. H. SANT'ANNA, A. BILINSKI, AND J. POE (2023): "What's trending in difference-in-differences? A synthesis of the recent econometrics literature," *Journal of Econometrics*.

RUBIN, D. B. (1980): "Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment," *Journal of the American Statistical Association*, 75, 591–593.

SANTOLERI, P. AND E. RUSSO (2023): "Spurring Subsidy Entrepreneurs," *European Commission, JRC135716.*

SANT'ANNA, P. H. AND J. ZHAO (2020): "Doubly robust difference-in-differences estimators," *Journal of Econometrics*, 219, 101–122.

SIEGLOCH, S., N. WEHRHÖFER, AND T. ETZEL (2025): "Spillover, Efficiency, and Equity Effects of Regional Firm Subsidies," *American Economic Journal: Economic Policy*, 17, 144–80.

SOLOW, R. M. (1956): "A Contribution to the Theory of Economic Growth," *The Quarterly Journal of Economics*, 70, 65–94.

STUART, E. A., B. K. LEE, AND F. P. LEACY (2013): "Prognostic score–based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research," *Journal of Clinical Epidemiology*, 66, S84–S90.e1.

WOOLDRIDGE, J. M. (2009): "On estimating firm-level production functions using proxy variables to control for unobservables," *Economics Letters*, 104, 112–114.

ZHANG, D. AND G. XU (2019): "Does Government Subsidy Affect Firm Survival? Evidence from Chinese Manufacturing Firms," *Emerging Markets Finance and Trade*, 55, 2628–2651.

ZÚÑIGA-VICENTE, J. , C. ALONSO-BORREGO, F. J. FORCADELL, AND J. I. GALÁN (2014): "Assessing the effect of public subsidies on firm r&d investment: a survey," *Journal of Economic Surveys*, 28, 36–67.

A. Appendix

Alternative identification strategy

Another possible strategy to identify the causal impact of the policy under study is to exploit the score obtained by each firm in the final ranking in a regression discontinuity (RDD) framework. This quasi-experimental technique capitalizes on the fact that the treatment assignment is a function of a score such that all (or some) units that places themselves above a certain threshold receive the treatment. If units are unable to "sort" around this threshold, the discontinuous change in the treatment probability can be used to learn about the local causal effect of interest, because units with scores barely below the cutoff can be used as a comparison group for units with scores barely above it (see Cattaneo and Titiunik 2022 for a good methodological exposition). Some recent papers have combined the exogenous variation stemming from discontinuities in the treatment assignment with the one coming from the repeated observation of units over time, and then called this methodological framework "difference-in-discontinuities" (Grembi et al. 2016). In this analysis, I could apply a similar strategy. First of all is useful to assess the discontinuity in the treatment assignment. Figure A.1 shows two polynomials fit (on the left-side and on the right-side of the score threshold normalized to zero) of the probability of being treated. In the figure, all firms on the left of the threshold are untreated since the estimated probability is always zero. However, on the right-hand side the probability of being treated is different from one, as there are some firms which scored above the threshold but didn't received the subsidy[24] . This evidence points in the direction of fuzzy design in which the placement of a firm above or below the threshold provides an exogenous variation which can be used to tackle the endogeinty problem caused by the non-random assignment of the treatment. Moreover, the left panel of Figure A.2 plots the number of treated and untreated firms by different bandwidths of the score. For the maximum bandwidth [-58; +58] I observe at total of 95 firms (62 treated and 33 controls), and this number sharply drops as I narrow the window. Due to the limited number of observations around the threshold a local randomization framework would suits better this application. Indeed, the local randomization approach has the advantage of requiring minimal extrapolation and avoiding the use of smoothing methods. It views the running variable as random, which introduces a stochasticity to the assignment mechanism in an RDD. Under this perspective, it is assumed that exists a window around the cutoff such that the treatment assignment for units is as-if randomized. Hence, the next step is to identify such window. To do so, I propose a test similar in spirit to the one suggested by Branson and Mealli (2019). In particular, I choose a starting window h and compute within that window the average standardized mean difference, between treated and control firms, of pre-treatment averages

[24] This could be the result of some firms violating one of the many legal obligations on which the subsidy was conditioned on.

of the outcomes.[25]. Then, I test whether this statistic is greater or less than 0.1[26]. As next step, I increase the size of the window to include more observations and repeat the test. I state that the treatment is as-if randomly assigned in the largest window h such that the average standardized mean difference of the pre-treatment averages of firms' characteristics is less than 0.1. The right panel of Figure A.2 shows the result of this exercise by plotting the aforementioned statistic as a function of the bandwidth h. Unfortunately, there isn't a window for which the test is passed[27]. That is, in all windows, treated and control firms substantially differ in terms of pre-treatment outcomes, and thus it is implausible to assume the random assignment of the treatment. The difference-in-differences technique remains the best way to identify the causal parameters of interest.

[25]In particular, I compute the averages over the period 2006-2008 for the intangible and tangible assets, the productivity, the employment, and the value of production, all in logs.

[26]This is the commonly agreed threshold in the literature (Stuart et al. 2013).

[27]I find similar results by using the standard approach of computing t statistics for the mean difference of the outcomes.

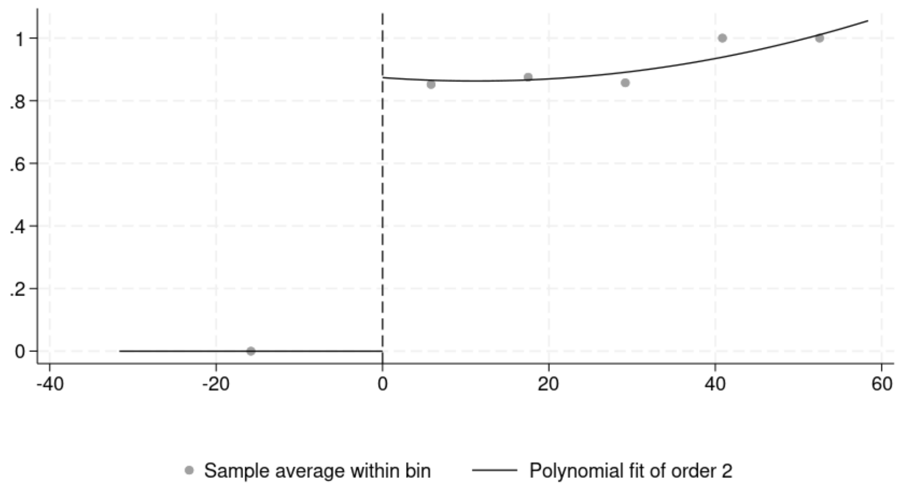
Additional tables and figures

Table A.1: Pre-treatment summary statistics - Winners vs. Paid firms

	Winners Mean(SD)	Paid Mean(SD)	Std. mean difference
Intangible assets	251.50	216.04	0.08
	(477.18)	(375.86)	
Tangible assets	1540.91	1539.29	0.00
	(2747.38)	(2377.18)	
Total assets	2033.78	2030.39	0.00
	(3245.06)	(2829.93)	
% of int./tot. assets	0.22	0.20	0.06
	(0.25)	(0.24)	
Num. of employees	27.29	29.98	-0.06
	(42.58)	(46.66)	
Value of production	5019.09	5426.89	-0.03
	(11225.54)	(12379.63)	
Productivity	212.70	214.20	-0.01
	(277.06)	(295.54)	
Return on assets	5.91	6.39	-0.06
	(7.72)	(7.90)	
Industry	0.58	0.57	0.01
	(0.50)	(0.50)	

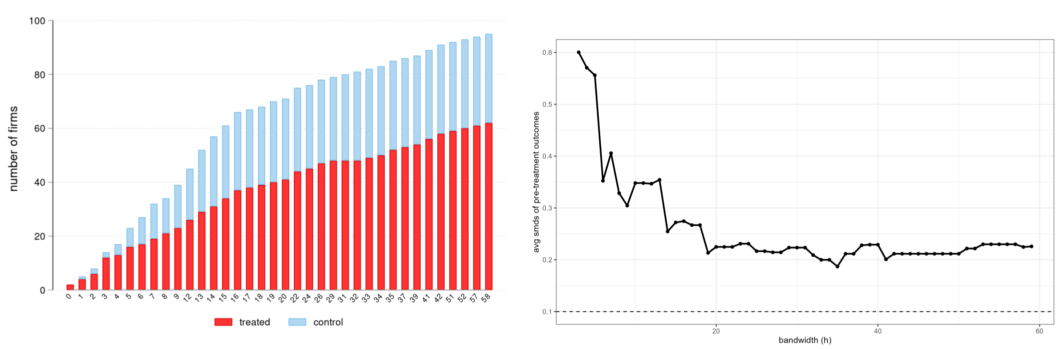
The table shows the pre-treatment averages (2004-2008) of firms' characteristics. The "Winners" group refers to winning firms identified in the final ranking. The "Paid" group refers to firms who received the subsidy payment. The third column reports the standardized mean differences which are computed as: $\frac{\mu_t - \mu_c}{\sqrt{(std_t^2 + std_c^2)/2}}$. A SMD less than 0.1 is a sign of good balance between the groups. Intangible assets, total assets, the value of production and productivity are expressed in thousand of euros.

Figure A.1: Treatment discontinuity around the score threshold



Note: the figure shows two polynomial fits to the left and to the right of the score threshold normalized to zero. The figure has been produced with the STATA's package RDROBUST (see Calonico and Titiunik 2017).

Figure A.2: Number of firms and average SMDs by score bandwidth



Note: the left figure shows the number of treated and control firms observed in the sample by score's bandwidth. The right figure shows the SMDs averaged across four variables (log intangible assets, log employment, log productivity and log value of production) by score's bandwidth.