The Effects of R&D Tax Credits and Subsidies on Private R&D in Mexico

Emmanuel Chavez

PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

WORKING PAPER N° 2020 – 29

The Effects of R&D Tax Credits and Subsidies on Private R&D in Mexico

Emmanuel Chavez

JEL Codes: H25, O32, O38
Keywords: Research and Development, Innovation, Propensity Score Matching, Public Policy, Impact Evaluation

AGENCE NATIONALE DE LA RECHERCHE
ANR

Funded by a French government subsidy managed by the ANR under the framework of the Investissements d'avenir programme reference ANR-17-EURE-001

The Effects of R&D Tax Credits and Subsidies on Private R&D in Mexico*

Emmanuel Chávez[†]

May 24, 2020

Abstract

This research studies the effects of a R&D tax credit and a R&D subsidy in Mexico. The Mexican tax credit removed the usual market oriented traits that define most tax credits. It essentially acted as a "deferred" subsidy, as firms got a discount on their corporate tax at the end of the fiscal year. Whereas the subsidy granted the funds at the start of the R&D project. My estimates show that both policies had a positive impact on innovation personnel, but the subsidy's impact was larger. As for patents, the impacts are less clear but favor the subsidy over the tax credit. The subsidy appears to have allowed less profitable firms to take on their R&D projects. This might have driven the larger subsidy effects. The awarding procedure in both programs is similar. Firms submitted their R&D projects to a non tax collecting institution. The projects were evaluated according to detailed guidelines. The awarded projects were selected based on the evaluations. The guidelines allow to construct a set of conditioning variables in a matching estimation approach. In addition, I use the difference-in-difference matching method to purge time-invariant unobservables.

Keywords: Research and Development, Innovation, Propensity Score Matching, Public Policy, Impact Evaluation

JEL Codes: H25, O32, O38

*I am very grateful to Antoine Bozio, my PhD advisor, for his constant support. I also thank Jimmy Lopez and Loriane Py for their valuable comments. In addition, I thank personnel at CONACYT for gently providing the data that made this research possible. This work has been funded by a French government subsidy managed by the Agence Nationale de la Recherche under the framework of the Investissements d'avenir programme reference ANR-17-EURE-001.

[†]PhD candidate at Paris School of Economics.

1 Introduction

Direct subsidies and tax credits are the two most widely used policies to promote research and development (R&D) in private companies. In most countries, a key difference between both policies is the agency granted to the public versus the private sector. Tax credits are more market oriented as firms are free to decide throughout the year the R&D projects they carry out. At the end of the fiscal year they get a discount on their corporate taxes amounting to a *share of their total* R&D expenditures. R&D subsidies are more government directed as a public institution selects the private R&D projects to fund. The Mexican tax credit rules removed the market orientation feature. Under both the tax credit and the subsidy, firms had to go through a selection process where a non tax collecting institution selected the R&D projects to award. The real difference between both programs concerned the moment when firms are rewarded. For the R&D subsidy, firms were granted the funds at the beginning of the calendar year; when they were set to start the R&D project. For the tax credit, firms received a discount on their taxes at the end of the fiscal year. In this sense, the Mexican tax credit acted as a "deferred" subsidy.

This paper studies the effect of the policy change: allocating the public funds at the start of the R&D project rather than delaying them at long after the project starts. Literature analyzing the effects of R&D tax credits and subsidies on private R&D is extensive. The post-2000 papers mostly show that both policies have a crowding-in effect on private R&D spending, i.e. private firms spend on R&D more than they would have without the public funds. However, few papers compare the effects of both policies in the same study –some exceptions are Parisi and Sembenelli (2003) and Carboni (2011)–. Most comparisons are drawn by literature reviews like that of Becker (2015) or Cerulli (2010). In addition, most literature on R&D tax credits and subsidies focuses on high income countries. Outcomes on countries far from the technology frontier may be different. Companies in these countries may be less efficient in translating R&D awards into meaningful R&D activities for many reasons. For instance, personnel with technical or scientific education may be scarce, or proper research equipment could be difficult to obtain. So, this paper adds to previous literature, first by exploiting the unique characteristic of the Mexican tax credit and comparing the effects of the policy change. Second, by analyzing the policy effects in a country far from the technology frontier.

My estimates show that both the R&D tax credit and the R&D subsidy led firms to allocate more personnel to innovation activities. But the size of the effect is larger for the subsidy. As for patent registrations, there is some evidence that the R&D subsidy could have had a positive effect on patents but it is not conclusive in at least a 95 percent confidence level. On the other hand, it can be confidently discarded that the tax credit had an effect on patents. So, overall the positive effect on private R&D activities is larger under the subsidy. My estimations suggest that this result is driven by the R&D subsidy allowing solvency constrained firms to take R&D activities. These firms might have been excluded from the tax credit due to lack of resources to start their R&D projects. My results are different from those of Carboni (2011). He finds that the Italian R&D tax credits have bigger effects than the subsidies. However, Italian tax credits allow firms more decision power to decide which R&D projects to take on. This highlights the importance of studying the features of a policy setting where firms have less agency.

Specifically, I analyze two programs: 1) *Estímulo Fiscal a la Investigación y el Desarrollo de Tecnología* (EFIDT), the R&D tax credit; and 2) *Programa de Estímulos a la Innovación* (PEI), the subsidy. Both were granted by *Consejo Nacional de Ciencia y Tecnología* (CONACYT), the institution in charge of the federal government science and technology policies. The EFIDT program lasted from 2001 to 2008. In year 2009, it was replaced with the PEI program. As CONACYT switched from the tax credit to the subsidy, the granting rules remained quite similar. To get the public support, companies submitted at least one R&D

project to CONACYT. Submitted projects were assessed following detailed evaluation guidelines. Based on the evaluations, a CONACYT committee chose the projects to award.

As awards were not randomly allocated, supported and non-supported firms could differ on unobserved characteristics correlated with the outcome. For example, R&D intensive firms may have more expertise on applying to the R&D programs, and hence be more likely to be awarded. I deal with this endogeneity problem with two metodologies. First, I use the propensity score matching approach. The method assumes that there is a set of observable conditioning variables for which outcomes are independent of treatment conditional on those variables. Matching estimation depends on the ability to construct the set of conditioning variables. The policy design and the data I have in this research gives two advantages for this aim: 1) the evaluation guidelines provide a solid base to construct the set of conditioning variables, and 2) my data provides a rich set of variables allowing to actually construct the set. In my second estimation approach, I use the difference-in-difference (DID) matching methodology proposed by Heckman et al. (1997, 1998) to purge time-invariant unobserved variables that may have not been accounted for with the matching method.

This research relies on three sets of data. The first two sets are administrative data provided by CONACYT on the EFIDT tax credit and the PEI subsidy. These datasets are not publicly available. They have data on the firms and projects that were awarded with the EFIDT program in the 2004-2008 period and the PEI program in the 2009-2013 period. The third dataset is the Economic Census collected by *Instituto Nacional de Estadística y Geografía* (INEGI). The census provides information on the firm outcome and control variables.

The rest of the paper is structured as follows: in Section 2, I present a summary on previous research on the effects of tax credits and subsidies. In this section I also include a description of the tax credit design in several countries. In Section 3 I present an overview of the spending on R&D in Mexico since year 2000. In addition, I describe with detail the R&D tax credit and subsidy programs in the country. Section 4, presents the datasets I use in this research and gives a thorough description of the methodology. In Section 5 I show my results and comment their implications. Finally, Section 6 concludes.

2 Overview of R&D Literature and Policies

Research and development (R&D) activities are credited as an important tool to promote future growth due to their positive effect on productivity.[1] However, literature has noted that private spending on R&D may be lower than the socially desirable, mainly two for reasons: first, private firms do not appropriate completely the returns of their own R&D investment as some of the benefits spill to other agents (Griliches, 1979); second, R&D activities are risky, as returns from R&D spending may take long to be perceivable, if at all.[2] Public support of R&D might encourage private firms to invest more on these activities and move R&D spending closer to the social optimum. The two most popular policies to promote private R&D are tax credits and direct subsidies. Literature analyzing their effects on private R&D is vast. The earlier research was not conclusive on the direction of these effects.[3] However, these early studies did little to deal with the typical problem of sample selection bias of impact evaluation studies. If R&D intensive firms are more likely to apply for –and be granted with– R&D supporting programs, OLS estimates on the effects of public support on private R&D might be biased.

Post-2000's literature has done bigger efforts to deal with the endogeneity problem. This

[1]See research by Cameron et al. (2005) in the United Kingdom, O'Mahony and Vecchi (2009) in the US, UK, Japan, Germany and France, or Bravo-Ortega and García Marín (2011) global study.

[2]Research by Hyytinen and Toivanen (2005) or Czarnitzki (2006) indicates that investment in R&D activities is riskier than investment in physical assets.

[3]Hall and Reenen (2000) review pre-2000 literature on the effects of R&D tax credits on private R&D. In the studies they analyze, tax credits tend to have a positive effect on R&D expenditure, but with great variation. David et al. (2000) review early literature on R&D subsidies. Evidence on the effects of this policy was less conclusive, as many studies report crowding-out effects, i.e. firms decrease their private R&D expenditure when they receive the subsidies.

later research concludes more firmly that R&D tax credits and subsidies have positive effects on private R&D. In Table 1, I show a summary of the post-2000 literature. The table does not include post-2000 papers that only use the traditional OLS estimates. It also excludes papers that use cross-country data; only papers with firm level data are included. Most studies find positive effects of subsidies and tax credits on private R&D. Some find no effects, but none of the studies in Table 1 find negative effects. So, recent evidence indicates that public support does not crowd-out private R&D spending: a dollar of public support is matched with, at least, an additional dollar of private spending.[4]

Table 1 shows that most literature is limited to high income countries. Of the 38 papers included in the table, only five study middle income countries: Özçelik and Taymaz (2008) in Turkey, Jia and Ma (2017) and Chen et al. (2018) in China, and Calderón (2009) and Chávez (2019) in Mexico. More research is needed in countries far from the technology frontier. Policies that support private R&D may have different effects in these countries as their firms may be less capable to assimilate resources destined to R&D. It could be that, when faced to a problem that requires research and innovation, middle-income country firms prefer to adopt solutions already developed in technology-frontier countries, instead of carrying out R&D activities of their own. Limited access to R&D inputs may also play a role. In addition, Table 1 shows that most papers –specially those that analyze R&D subsidies– study just one outcome: private R&D spending. It is important to analyze additional outcomes, as the effects of R&D policies may differ depending on the outcome. For instance, research by Bozio et al. (2014) in France shows that R&D tax credits have a positive impact in private R&D spending, but have no effect on patents. This gives a broader perspective on the policy's effectiveness.

[4]Some studies find significant effects only for subsets of firms. For instance, Lach (2002), Bronzini and Iachini (2014) and Kobayashi (2014) find that the public R&D support in Israel, Italy and Japan, respectively, is only significant for small and medium sized firms. Görg and Strobl (2007) find that R&D subsidies in Ireland are effective for national firms, but not for foreign Ireland based firms. Jia and Ma (2017) show that R&D tax incentives in China have positive effects in R&D spending of private firms, but not on public firms. Hægeland and Møen (2007) find that the Norwegian tax credit is only effective for firms in low R&D sectors.

Finally, few studies analyse the impact of both policies in the same context. Only two papers in Table 1 study both R&D tax credits and subsidies: Parisi and Sembenelli (2003) and Carboni (2011) in Italy. They find positive and significant effects of both policies on private R&D spending. Yet, Carboni (2011) finds that the size of the effect is larger for tax credits. In addition, he finds that R&D subsidies –as opposed to tax credits– lead firms to obtain more external credit to finance their R&D projects. So, the general effects of the two policies point in the same direction. But there are differences on the size of the effects and the channels through which they operate.

In the case of Mexico it is specially important to analyze both policies simultaneously as the Mexican R&D tax credit has a rare characteristic: its awarding process is very similar to the regular R&D subsidy. To get the tax credits, firms had to submit a R&D project to a non-tax collecting authority. The projects went through and evaluation process, and the authority selected those that got the tax credit. In most countries, the R&D tax credit design does not require firms to go through a selective process. Firms just declare their R&D expenditures to the tax authority and deduct the amount from their future owed corporate income tax. In this sense, the usual tax credit design is market oriented: firms allocate spending to the R&D projects they choose and are compensated at the end of the fiscal year. R&D subsidies are more government oriented since a public authority chooses the projects to award with the grants. The Mexican R&D tax credit design removes the market component. In Mexico, the tax credit acts as a "deferred" subsidy: instead of getting funds at the start of the R&D project, tax credited firms got the funds at the end of the fiscal year.

Table 2 summarizes the R&D tax credit design across countries. Column 2 shows that, as mentioned above, the selective tax credit designs are rare. Only in Mexico and Norway, firms have to submit a R&D project to an authority different than the tax revenue administration.

This assignment criteria affects the base of R&D expenditures that firms can deduct from the owed corporate tax. In most countries, the deductible base is determined as a proportion of the firm's *total* R&D expenditures. In Mexico, the deductible base is a proportion of the R&D project approved by the non-tax authority (see Column 3). In other aspects, the Mexican design is similar to other countries. For instance, as shown in Column 5, countries usually allow companies to carry-forward the tax credit to future fiscal years in case the credited amount is larger than present year's owed corporate taxes.

3 R&D Spending and Policies in Mexico

3.1 Spending

Over the last two decades, Mexico's R&D spending has steadily increased, almost doubling its size in the 2004 to 2017 period. However, as shown in Figure 1, growth of the R&D/GDP ratio is has been more modest, going from around 0.40 percent of GDP in year 2004 to 0.48 percent in 2017. As for the sources that funded this rise in R&D spending, Figure 1 depicts clearly that it is the share of publicly financed R&D that increased in the period –although growth was somewhat reversed in the last years–. On the other hand, the share of private R&D decreased notably. Public R&D rose strongly, not only as a share of total R&D spending, but also in real terms. Compared to the initial level, the rise in absolute public R&D disbursments is mantained even if we take into account the large spending cuts of the last three years of the period. On the other hand, absolute private R&D spending has remained mostly flat. So, the increase in total R&D spending that we observe over the last two decades, mostly came from the increase in public R&D funding and not form the private sector.[5]

[5] In the last four years, there was a rise in R&D spending from non-for-profit organizations. However, it still remains small in absolute terms compared to public or private sector funding. More information on the R&D context in Mexico can be found in CONACYT (2004–2017).

The modest rise in the Mexican R&D/GDP ratio in the last decades keeps the country faraway from the levels observed in most OECD countries. Figure 2 shows the R&D/GDP ratio in selected countries. Average R&D spending in OECD countries stands at 2.35 percent of GDP. In Korea, R&D expenditure as a share of GDP stands at a whopping 4.24 percent (OECD, 2018). Mexico's investments on R&D are small compared to high income countries, but they are comparable to those of its Latin American counterparts. The average R&D/GDP ratio in the region stands at 0.55 percent. Note, however that the region has its front runners and Mexico is not one of them. For instance, Brazil, the region's leader, spends 1.28 percent of its GDP in R&D activities. In addition, Figure 2 shows that high R&D spending countries tend to have higher shares of funding coming from the private sector. In Korea, Japan and Germany, the countries with the highest R&D/GDP ratios shown in the figure, private shares of total R&D spending go from 65 to 80 percent. The contrast to Mexico's 20 percent is stark. Shares of private sector funding of R&D are also higher in countries with "medium" levels of the R&D/GDP ratio. For instance, in Brazil, about half of R&D spending comes from the private sector. Thus, it is likely that Mexico's private sector should invest higher sums so total R&D spending significantly increases.

3.2 Policies

The largest R&D policies in Mexico are managed by *Consejo Nacional de Ciencia y Tecnología* (CONACYT). In 2017, it spent 46 percent of all Federal Government's R&D expenditure. Over the last two decades, CONACYT's main policies directed to support private R&D were: 1) *Estímulo Fiscal a la Investigación y el Desarrollo de Tecnología* (EFIDT), a tax credit; and 2) *Programa de Estímulos a la Innovación* (PEI), a direct subsidy.[6] The EFIDT tax credit lasted from 2001 to 2008. Firms wishing to use the tax credit had to submit a R&D project to CONACYT. A CONACYT committee selected the projects to award. The discount on the corporate income tax amounted to 30 percent of the total

[6]Besides the EFIDT and PEI programs, CONACYT managed other funds and programs to support private R&D. For more information on these programs see Villavicencio (2010) and Villavicencio (2011).

expenditures of the approved R&D project. With carry-forward to 10 fiscal years in case the awarded sums were larger than the tax owed in the year when the tax credit was granted.

Authors such as Unger (2011) have criticized several aspects of the EFIDT program, among them: 1) it was awarded to relatively few firms. For instance, in year 2005, 15 firms concentrated 50 percent of the total amount awarded to all firms. 2) The awarded credits were highly concentrated by sector, as in year 2005, half of all credits were awarded to firms in the automotive industry. CONACYT may have taken note of these critiques, as in year 2007, EFIDT awarding rules provided some prioritizing of small and medium sized firms. Nonetheless, due to these or other concerns, in year 2009 CONACYT decided to suppress the EFIDT program altogether. It was replaced with the PEI subsidy, which lasted from 2009 to 2018.[7] I know of no public document that explains why CONACYT decided to switch its private R&D supporting approach from a tax credit to a direct subsidy. However, some reasons can be deduced. The first is resource availability: awarding a tax credit does not require an R&D supporting agency to hand out its own resources to the private firms. Instead, it is the federal government that forgoes a share of the corporate taxes it would have received from the tax credit awarded firms. So, in periods of less resource availability, a R&D supporting agency might decide to support private R&D via tax credits instead of direct subsidies. In fact, CONACYT budget grew steadily in the period of time up to the tax credit-subsidy switch.[8] The second reason is agency: subsidies concede a higher power to decide which R&D projects and firms get supported. However, as mentioned above, the

[7]The federal government that took office in December 1, 2018 decided to discontinue the PEI subsidy as well, alleging that scarce R&D funds should not be transferred to private firms. The administration argues that funds can be put to better use in other type institutions. See the press conference on June 27, 2019 by CONACYT's director Elena Álvarez-Bullya https://www.youtube.com/watch?v=i5cYvoXEzGM.

[8]Specifically, it went from 14,833 million MXN in 2002 (at 2018 constant prices) to 29,956 million MXN in 2009, and then to 40,341 million MXN in 2015, the year with the highest budget. More resource availability might have allowed CONACYT to decide directly which R&D projects to support via subsidies. Since year 2015, cuts have reduced CONACYT's budget to 26,925 million MXN in year 2018. In year 2017, the agency decided to reintroduce the EFIDT tax credit, although the awarded amounts were small compared to those of the 2000's decade. In year 2019 the PEI subsidy was completely abolished, but the tax credit remained. Diminished resource availability might have been behind this decision. Table 3 shows more data on CONACYT's budget.

Mexican R&D tax credit design made firms to go through a selection process very similar to the one of a typical R&D subsidy. So, this second reason does not seem of much relevance in the Mexican context.

As CONACYT switched from tax credits to direct subsidies, the granting decision remained similar.[9] To receive the awards, firms had to submit a R&D project that clearly defined the activities to be undertaken with the government support. All submitted projects were evaluated by individuals chosen from a CONACYT directory of evaluators. Evaluators were usually academics, scientists or researchers affiliated to universities or research centres. The evaluation was performed according to guidelines that indicated the project and firm characteristics to assess. The evaluations were then passed to a CONACYT committee, which chose the R&D projects to award based on them. The firm characteristics to assess remained similar when CONACYT switched from the EFIDT tax credit to the PEI subsidy.[10]

The PEI subsidy was paid in the first months of the calendar year, when the firms started their R&D projects. And it had to be spent during the calendar year. The EFIDT tax credit was announced at the start of the fiscal year. Awarded firms had to spend the tax credited amount on the awarded R&D project during the fiscal year. Then, they got a discount on their corporate income tax at the end of the fiscal year. So, in this setting the Mexican tax credit acted as a "deferred" subsidy.

[9] In both programs, a company could submit more than one R&D project. More than one project per company could be awarded in the same year. Projects already supported by another CONACYT program could not be submitted, effectively ruling out most other available public R&D programs in Mexico. For both programs, companies wishing to apply to the awards had to be legally constituted, registered with the tax authority and be up-to-date with their tax obligations. So, informal firms were not eligible.

[10] For the PEI subsidy, the granting process included a grade threshold on the evaluations. Projects graded below the threshold by the evaluators were not passed to the CONACYT committee, so they did not get the subsidy. The committee selected which projects got the PEI grants from those that were graded above the threshold. However, not all projects above the cut-off got the grants. Since my estimation strategy is based on comparing awarded firm outcomes to those of non-awarded firms (selected with a propensity score), I only need information on which firms were actually awarded with the EFIDT tax credit or the PEI subsidy. Thus, the grade threshold in the PEI granting scheme does not affect my estimates. I use PEI's grade threshold in Chávez (2019) to estimate the impact of the subsidy in a wider set of outcomes with a regression discontinuity design.

4 Data and Methodology

4.1 Data

This research relies mainly on three sets of data. Two sets are administrative data provided by CONACYT. The first of these two datasets provides information on 8,428 projects belonging to 1,316 firms that were *awarded* with the EFIDT tax credit in the 2004 to 2008 period. The second dataset contains information on 6,361 projects that were *submitted* to the PEI subsidy from 2009 to 2013 by 2,842 companies. The PEI and EFIDT datasets are not publicly available. The third main dataset I use in this research comes from the 2014, 2009 and 2004 Economic Censuses collected by *Instituto Nacional de Estadística y Geografía* (INEGI). The census gathers information on a wide set of firm economic activities including labour, expenditures, income, output, assets, research and development, among many others. From the censuses I get the control and outcome variables.

Data on the EFIDT and PEI datasets is provided yearly. The census is gathered every five years. The economic census released on year 2014 contains information that corresponds to activities performed on year 2013. The census released on 2009 contains information for 2008. And the 2004 census contains information for 2003. So, the 2014 Economic Census provides post-treatment data for the firms in the PEI dataset –awarded projects in the 2009-2013 period–, and the pre-treatment data is obtained from the 2009 census. Similarly, the 2009 census provides post-treatment data for the EFIDT awarded firms –going from 2004 to 2008–, and I get the pre-treatment data from the 2004 census. The CONACYT and INEGI datasets do not count with a common code that uniquely identifies the firms across datasets. Hence, to merge the firms in the CONACYT data to the censuses, I use identifying variables present across datasets, namely: the firm's name, address and sector. Merging firms this way bears some complications. Among others, spelling mistakes are common. Regular merging commands cannot handle this type of merging, thus, I follow a special algorithm to merge the firms.[11] With this algorithm, I successfully merge around 70 percent of firms in the EFIDT dataset and 50 percent of firms in the PEI dataset (see Table 4). The reason I can merge a higher rate of firms in the EFIDT dataset is the presence of relatively more large firms. These tend to have higher survival rates and fewer registration mistakes in both the CONACYT and INEGI datasets.

4.2 Methodology

This research aims at analyzing the impacts of the EFIDT R&D tax credit and the PEI R&D subsidy on outcomes related to private firm R&D activities. Let us start by describing the following basic econometric model:

$$Y_i = \alpha + \beta T_i + \epsilon_i \tag{1}$$

where Y_i is the outcome for firm i, T_i is a variable equal to one if firm i is awarded with an R&D supporting program and zero otherwise, and ϵ_i is a random error. In equation (1), β represents the effect of receiving the R&D award on the outcome. However, as awards are not randomly allocated, if we estimate (1), parameter β is likely to be biased as awarded and non-awarded firms can differ on unobserved characteristics correlated with outcome Y_i. I can deal with this problem if I find a set of observable conditioning variables X_i for which outcomes Y_i are independent of treatment T_i conditional on X_i. I follow this path with the propensity score matching approach introduced by Rosenbaum and Rubin (1983). The technique is based on pairing program participants with non-participants. The pairs are chosen based on the similarity in the probabilities to receive the R&D tax credits or subsidises. The impact of the program is estimated as the mean difference in the outcomes of the matched pairs. Deriving causality from the estimates obtained with the matching technique can be problematic as it depends on finding an appropiate set of conditioning variables X_i.[12] The

[11] A detailed description of the algorithm is provided in Appendix A.

[12] Research by Smith and Todd (2005) finds that matching estimators are highly sensitive to the choice

more information I have on the awarding criteria, the best I can replicate the probability to be awarded, and the less bias I can expect on my matching estimators.

The policies and data I use in this paper have important advantages for this aim. First, in both the EFIDT and PEI programs, evaluators assessed the submitted projects according to guidelines that clearly defined the firm characteristics to be measured. These evaluations were used later to determine which projects got the awards. I can follow the guidelines to define the conditioning variables X_i that resemble the criteria used to decide which firms got the awards. The economic census data provides a rich set of variables allowing to actually construct the set X_i based on the awarding criteria. A second advantage is the large number of firms available in the census databases. This offers a large pool of possible matches.[13] Indeed, all treated firms lie in the region of common support, i.e. all awarded firms have propensity scores lower than the maximum or higher than the minimum scores of the non-awarded firms. In addition, to guide the estimations, I take into account recommendations in previous research to construct the set of conditioning variables X_i.

Let us now describe the matching technique with more precision. First, define a propensity score as $P = Pr(T = 1|X)$. The propensity score allows to reduce the dimensionality of the set of conditioning variables X_i to a single score P_i that determines the probability to be awarded with a R&D program. A matching estimator takes the form:

$$\beta = \frac{1}{N_1} \sum_{i \in T_1 \cap C_p} [Y_{1i} - \hat{E}(Y_{0i}|T_i = 1, P_i)] \tag{2}$$

of different subsamples and to the set X_i. In addition, Heckman et al. (1996, 1997) and Lechner (2002) find that, for matching estimators to have low bias, datasets must include a rich set of variables related to program participation and outcomes to be able to construct a good set X_i. These variables should be measured in the same way for the treated and non-treated sample.

[13] Heckman et al. (1996, 1997, 1998) find that a source of bias in matching estimation may come from the absence of possible matches for all program participants. If there are treated firms for which no matches can be found, then the treatment effect has to be defined as the impact of the R&D programs for companies in the *common support region*.

Where

$$\hat{E}(Y_{0i}|T_i = 1, P_i) = \sum_{j \in T_0} W(i,j) Y_{0j}$$

T_1 denotes awarded firms and T_0 denotes non-awarded firms. C_p is the region of common support, and N_1 is the number of firms in the set $T_1 \cap C_p$. The match for each firm awarded with a R&D program $i \in T_1 \cap C_p$, is a weighted average over the outcomes of non-awarded firms $\sum_{j \in S_0} W(i,j) Y_{0j}$. Weights $W(i,j)$ depend on the distance between P_i and P_j. So the matching estimator β is simply a weighted difference between the outcomes of awarded firms and non-awarded firms. To select the non-awarded matched firms, define a neighbourhood $B(P_i)$ for each awarded firm i. Neighbours for i are non-awarded firms $j \in T_0$ for whom $P_j \in B(P_i)$. Matching estimation methods differ in the way the neighbourhood is defined. I follow three ways to define the neighbourhood $B(P_i)$: *one-to-one matching*, *k-nearest neighbours matching* and *kernel matching*.[14]

As for the variables that should be included in the set of conditioning variables X, the only variables that influence the estimates are those that determine selection into the awards T and the outcomes Y.[15] The firm characteristics assessed in the EFIDT and PEI evaluation guidelines can be grouped into the following categories: 1) firm experience in the market,

[14] In the *one-to-one matching*, the neighbourhood is $B(P_i) = min||P_i - P_j||, j \in I_0$. That is, the non-awarded firm with value P_j that is closest to P_i is selected as the match for i. I match without replacement so each $T = 0$ firm serves as match for at most one $T = 1$ firm. Since the census has a large number of firms, I do not need to reuse firms to preserve the quality of the matches. On the other hand, I increase the number of non-awarded firms to construct the conterfactual outcome, thus reducing the estimator variance. In the *k-nearest neighbours matching*, the awarded firm i with value P_i is matched to the $k = 10$ non-awarded firms for whom the value P_j is closest to P_i. Each of the $k = 10$ nearest neighbours receive the same weight when I construct the conterfactual mean outcome. Using $k > 1$ neighbours as matches reduces the variance, as more information is used to get a conterfactual on each awarded firm, but it may increase bias by using matches with more distant scores. In *kernel matching*, the match for each awarded firm i is constructed using a kernel-weighted average over multiple non-awarded firms. The weights depend on the distance between each non-awarded firm and the awarded firm for which the contractual outcome is being constructed. I use the epanechnikov kernel to construct my estimates. More information on this statistical method can be found in Silverman (1986).

[15] Variables that do not influence the treatment do not influence choice into the R&D programs, thus do not influence selection bias. The variables that influence treatment but do not influence the outcomes, create selection but do not have an impact on the distribution of the outcome, so they have no impact on possible selection bias.

which I control for with the *firm age*; 2) firm profitability, which I include with the firm *profit margin*, *interest costs over total costs* and a dummy variable that indicates availability of *own resources to fund innovation* projects; 3) region and sector preferences, which are controlled with six *region dummies* and six *sector dummies*; and 4) previous experience in R&D projects, which is controlled for with a dummy variable that indicates presence of an *innovation department*, a dummy variable that indicates *collaboration with research centres* and a dummy that indicates past *positive research spending.* In addition, I include other variables in set X based on previous research that analyses R&D tax credits and subsidies with matching techniques:[16] *total employees* (expressed in logarithm) to take into account firm size, as it has been noted that bigger firms tend to carry out more R&D activities; *fixed assets per employee* as research shows that more capital intensive firms might be involved in more R&D; *exports over total sales* and a dummy for *foreign capital participation*, to take into account the differences in R&D support that might come from more exposure to the international markets; a dummy for firms that receive other *government subsidies*, to control for the past expertise of firms on applying for and receiving public funds. Finally, note that for estimate β not to be biased, the set of controlling variables X must not depend on treatment T. Thus, all the variables in the conditioning set X are taken from the pre-treatment census year. For the EFIDT tax credit, I take control variables from the 2004 Economic Census. For the PEI subsidy, controls are taken from the 2009 Economic Census.

The matching estimator described in equation (2) assumes that, after conditioning on a set of firm characteristics X, outcomes are conditionally independent of treatment T. This is a strong assumption, β estimates from equation (2) may be biased by unobserved variables. So, I use an additional estimator to overcome this assumption. A differences-in-differences (DID) matching approach, as defined in Heckman et al. (1997, 1998) allows to

[16] See Almus and Czarnitzki (2003), Duguet (2004, 2012), Czarnitzki and Licht (2006), Aerts and Schmidt (2008), Özçelik and Taymaz (2008), Corchuelo and Martínez-Ros (2010), Czarnitzki et al. (2011), Carboni (2011), Cerulli and Poti (2012) and Kobayashi (2014).

compare the change in the outcome for the awarded firms, and the change in the outcome for the non-awarded firms. The DID matching estimator allows to purge all time-invariant unobservables in the matched firms. With this estimator I can recover the true impact of the R&D programs, as long as the non-observables in the treatment and control groups follow common time trends. The DID matching estimator is similar to the regular diff-in-diff estimator, but it does not impose a linear functional form and the control group is drawn via propensity score matching. In particular, it is given by:

$$\begin{aligned}\beta = & \frac{1}{N_{1t}} \sum_{i \in T_{1t} \cap C_p} \{(Y_{1ti} - \sum_{j \in T_{0t} \cap C_p} W(i,j) Y_{0tj}\} \\ & - \frac{1}{N_{1t'}} \sum_{i \in T_{1t'} \cap C_p} \{(Y_{1t'i} - \sum_{j \in T_{0t'} \cap C_p} W(i,j) Y_{0t'j}\}\end{aligned} \tag{3}$$

Where $T_{1t}, T_{1t'}, T_{0t}, T_{0t'}$ denote awarded and non-awarded firms in periods t –after the award took place–, and t' –before the award took place–. I show results for the one-to-one, k-nearest neighbours and kernel matching estimates, both in "cross-section" matching (eq. 2) and DID matching (eq. 3).

I evaluate two outcomes: 1) personnel working on innovation activities, and 2) patent registrations.[17] These outcomes are associated to a different "intensity" on the undertaking of R&D. In the economic censuses, innovation is defined as the introduction of new or significantly improved products (goods or services) or processes (including production methods) into the market. Innovation may come from well defined projects, or from routine improvements, spontaneous ideas or other *non-systematic* factors. So, the outcome on personnel working on innovation is my weak measure of R&D, as increases in innovation activities may be driven by well-defined projects or from non-systematic work. On the other hand,

[17] Both the EFIDT tax credit and the PEI subsidy granted money for the following factors associated to the submitted R&D project: wages, designs and prototypes, patents and intellectual property, operation costs, equipment for the research laboratory and laboratory improvements.

the patent registration outcome is my strong measure of R&D since it gives a more certain indication of the "success" of increased R&D undertaking.

5 Results

Let us first introduce the estimates of the EFIDT program. Table 5 shows the results of the propensity score model and Table 6 shows the "cross-section" matching results of β from equation (2). We see in Table 6 that, for the innovation personnel outcome, all estimates are positive and statistically significant. This means that the EFIDT program increases the probability of having personnel working on innovation by around 0.15 points in a zero to one scale, according to the one-to-one matching estimate. Concerning patenting activities, the estimates from equation (2) show a positive and statistically significant impact as well, but the effect is smaller. According to these estimates, the EFIDT program increased the probability of registering patents by around 0.08 points.

However, after conditioning on X, there might still be unobserved variables biasing β. If these follow common trends in the treatment and control groups, I can purge them with the DID matching estimator of equation (3). These results are shown in Table 7. For the outcome on innovation personnel, we see that the positive and statistically significant β estimates are maintained in all matching estimators. The size of the effect is reduced by half in the kernel estimates, but remains similar for the one-to-one and 10-nearest neighbor estimates. On the other hand, for the outcome on patent registration, we see that β estimates are no longer statistically significant. Further, they are negative in two of the matching estimators. The one-to-one estimates of β in cross-section and DID matching are depicted in Figure 3. For the innovation outcome, we clearly see estimates significantly different from zero at a 95 percent confidence level in the post-treatment period (eq. 2) and in the DID estimate (eq. 3). The mean difference between awarded and non-awarded matched firms (β)

in the pre-treatment period is not statistically different from zero. This means that firms in the treatment and control groups allocated personnel to innovation activities similarly before treatment. For the patent outcome, we see a β different to zero in the pre-treatment and post-treatment periods. So, patenting differences between treated and non-treated firms appear not be due to the tax credit. This results in a DID estimate not statistically different from zero. Thus, the evidence suggests that the EFIDT program led to an increase in the probability of having personnel allocated to innovation activities. But not to an increase in the probability of patent registration.

Now let us review the effects of the PEI program. Table 8 shows the results of the propensity score model and Table 9 shows the β estimates from equation (2). Table 9 shows that the subsidy has a positive and statistically significant impact on both innovation personnel and patent registration. For the innovation personnel outcome, the size of the effect is larger than the one of the EFIDT program, ranging from $\approx$0.30 to $\approx$0.40 point increase in a zero to one scale. The β estimates are also positive and significant for the patent registration outcome. And orders of magnitude are, as well, larger than those of the EFIDT program. Table 10 shows the DID matching estimates of equation (3). Concerning the innovation outcome, the estimates remain positive and statistically significant. With respect to the patents outcome, all DID matching coefficients are positive. Two of them are statistically different from zero at a 90 percent confidence level, but none at conventional 95 percent levels. Figure 4 depicts the one-to-one matching estimates of β in cross-section (eq. 2) and diff-in-diff (eq. 3) for the PEI subsidy. For the innovation outcome, mean differences between treated and non-treated firms are positive and significant in the post-treatment period and in the DID estimate. For the patent outcome, we see a DID estimate whose confidence interval is mostly positive, but I cannot accept a β different from zero by a small margin at the 95 percent confidence level.

Let us do a brief comparison of the effects of the EFIDT program versus the PEI program. I focus in my preferred estimator: the differences-in-differences matching. Concerning the innovation personnel, the point estimate of β, is larger for the PEI program by around 0.10 points in a zero to one scale. As for the patent registration outcome, the β point estimates for the PEI program are positive in all cases, although with low levels of significance. Concerning the EFIDT patenting estimates, all are close to zero, sometimes negative and below 90 percent level of significance in all cases. So, overall the evidence suggests that the PEI subsidy had a larger effect on private R&D activities compared to the EFIDT tax credit.

As mentioned above, in the Mexican policy setting, the R&D tax credit is the equivalent to a "deferred" subsidy. So, the different effects of the EFIDT tax credit versus the PEI subsidy should be considered, all other things remaining equal, as the result of a change in the moment when the funds are delivered, not as the result of a complete switch from a regular tax credit to a subsidy. With this in mind, the results of the propensity scores may shed light on the factors that drive the distinct effects of the EFIDT versus the PEI program. Tables 5 and 8 show the variables that determine the probability to be awarded with R&D support under each program. For both EFIDT and PEI, the probability to be awarded is increasing in: firm experience in the market, controlled with *firm age*; firm experience with R&D activities, controlled with presence of *innovation department*, among other variables; firm size, controlled with *employees*; and capital intensity, controlled with *fixed assets per employee*. However, solvency affects the probability to be awarded in different ways under the EFIDT program versus PEI. The *profit margin* increases the probability under the EFIDT program, whereas the effect is not significant for the PEI subsidy. This suggests that profitable firms could access more easily the tax credit compared to the subsidy. In addition, the availability of *own resources to fund innovation projects* has a negative sign in the PEI propensity score. This indicates that the PEI program might have favored firms that did not have access to other R&D resources. The switch from the tax credit to the subsidy basically changed the moment when the funds were disbursed to firms. So, granting access to funds at the start of the research project, might have included firms that were previously left out due to solvency or financial constraints. This in turn, could be a the driver of the larger effects on R&D activities that we observe under the PEI subsidy program.

6 Conclusion

This paper compares the effects of granting R&D funds at the start of the R&D project rather than long after the project began. My results suggest that this policy change is beneficial as the effects from the R&D subsidy are larger. This paper is in line with Chávez (2019). That research focuses on the R&D subsidy with a quasi-experimental methodology. It shows that the subsidy led to an increase in the probability of hiring innovation personnel, with no effects on other outcomes. In addition, this paper is complementary to research by Calderón (2009). He studies the R&D tax credit with a two step model, and finds that it had a positive effect on private R&D spending. My research complements the picture with other outcomes. Further analysis on patenting activities is needed, as this research did not give conclusive answers. Patent registration is a "stronger" R&D outcome as it is a more reliable measure of R&D success. So, the policy designs that may improve patenting activities in private firms are of great importance.

References

Aerts, K. and Schmidt, T. (2008). Two for the price of one?: Additionality effects of r&d subsidies: A comparison between flanders and germany. *Research Policy*, 37(5):806 – 822.

Almus, M. and Czarnitzki, D. (2003). The effects of public r&d subsidies on firms' innovation activities: The case of eastern germany. *Journal of Business & Economic Statistics*, 21(2):226–236.

Baghana, R. and Mohnen, P. (2009). Effectiveness of r&d tax incentives in small and large enterprises in québec. *Small Business Economics*, 33(1):91–107.

Becker, B. (2015). Public r&d policies and private r&d investment: A survey of the empirical evidence. *Journal of Economic Surveys*, 29(5):917–942.

Bloch, C. and Graversen, E. (2012). Additionality of public r&d funding for business r&d, a dynamic panel data analysis. *World Review of Science, Technology and Sustainable Development*, 9(2-4):204–220.

Bozio, A., Irac, D., and Py, L. (2014). Impact of research tax credit on r&d and innovation: Evidence from the 2008 french reform. Banque de France Working Paper 532, Banque de France.

Bravo-Ortega, C. and García Marín, A. (2011). R&d and productivity: A two way avenue? *World Development*, 39(7):1090 – 1107.

Bronzini, R. and Iachini, E. (2014). Are incentives for r&d effective? evidence from a regression discontinuity approach. *American Economic Journal: Economic Policy*, 6(4):100–134.

Busso, M., Fentanes, O., and Levy, S. (2018). The longitudinal linkage of mexico's economic census 1999-2014. IBD Technical Note 1477, Inter-American Development Bank.

Calderón, A. (2009). Evaluación del programa de estímulos fiscales al gasto en investigación y desarrollo de tecnología de las empresas privadas en méxico (efidt). Technical report, Foro Consultivo Científico y Tecnológico.

Cameron, G., Proudman, J., and Redding, S. (2005). Technological convergence, r&d, trade and productivity growth. *European Economic Review*, 49(3):775 – 807.

Cappelen, Å., Raknerud, A., and Rybalka, M. (2012). The effects of r&d tax credits on patenting and innovations. *Research Policy*, 41(2):334 – 345.

Carboni, O. (2011). R&d subsidies and private r&d expenditures: evidence from italian manufacturing data. *International Review of Applied Economics*, 25(4):419–439.

Cerulli, G. (2010). Modelling and measuring the effect of public subsidies on business r&d: A critical review of the econometric literature. *Economic Record*, 86(274):421–449.

Cerulli, G. and Poti, B. (2012). Evaluating the robustness of the effect of public subsidies on firms' r&d: an application to italy. *Journal of Applied Economics*, 15(2):287 – 320.

Chávez, E. (2019). The effects of public r&d subsidies on private r&d activities in mexico. PSE Working Papers 2019-73, Paris School of Economics.

Chen, Z., Lui, Z., Suárez Serrato, J. C., and Yi Xu, D. (2018). Notching r&d investment with corporate income tax cuts in china. NBER Working Paper 24749, National Bureau of Economic Research.

CONACYT (2004–2017). Informe general de la ciencia, la tecnología y la innovación. Technical report, Consejo Nacional de Ciencia y Tecnología (CONACYT).

Corchuelo, M. B. and Martínez-Ros, E. (2010). Who benefits from r&d tax policy? *Cuadernos de Economía y Dirección de la Empresa*, 13(45):145 – 170.

Czarnitzki, D. (2006). Research and development in small and medium-sized enterprises: The role of financial constraints and public funding. *Scottish Journal of Political Economy*, 53(3):335–357.

Czarnitzki, D., Hanel, P., and Rosa, J. M. (2011). Evaluating the impact of r&d tax credits on innovation: A microeconometric study on canadian firms. *Research Policy*, 40(2):217 – 229.

Czarnitzki, D. and Licht, G. (2006). Additionality of public r&d grants in a transition economy. *Economics of Transition*, 14(1):101–131.

Czarnitzki, D. and Lopes Bento, C. (2012). Evaluation of public r&d policies: a cross-country comparison. *World Review of Science, Technology and Sustainable Development*, 9(2/3/4):255 – 282.

David, P. A., Hall, B. H., and Toole, A. A. (2000). Is public r&d a complement or substitute for private r&d? a review of the econometric evidence. *Research Policy*, 29(4):497 – 529.

Dechezleprêtre, A., Eniö, E., Martin, R., Nguyen, K.-T., and Reenen, J. V. (2016). Do tax incentives for research increase firm innovation? an rd design for r&d. NBER Working Papers 22405, National Bureau of Economic Research.

Duguet, E. (2004). Are r&d subsidies a substitute or a complement to privately funded r&d? an econometric analysis at the firm level. *Revue d'economie politique*, 114(2):245–274.

Duguet, E. (2012). The effect of the incremental r&d tax credit on the private funding of r&d an econometric evaluation on french firm level data. *Revue d'economie politique*, 122(3):405–435.

Gonzalez, X., Jaumandreu, J., and Pazo, C. (2005). Barriers to innovation and subsidy effectiveness. *The RAND Journal of Economics*, 36(4):930–950.

Gonzalez, X. and Pazo, C. (2008). Do public subsidies stimulate private r&d spending? *Research Policy*, 37(3):371 – 389.

Görg, H. and Strobl, E. (2007). The effect of r&d subsidies on private r&d. *Economica*, 74(294):215–234.

Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. *The Bell Journal of Economics*, 10(1):92–116.

Hægeland, T. and Møen, J. (2007). Input additionality in the norwegian r&d tax credit scheme. Technical report, Statistics Norway.

Hall, B. and Reenen, J. V. (2000). How effective are fiscal incentives for r&d? a review of the evidence. *Research Policy*, 29(4):449 – 469.

Harris, R., Li, Q. C., and Trainor, M. (2009). Is a higher rate of r&d tax credit a panacea for low levels of r&d in disadvantaged regions? *Research Policy*, 38(1):192 – 205.

Heckman, J. J., Ichimura, H., Smith, J., and Todd, P. (1998). Characterizing selection bias using experimental data. *Econometrica*, 66(5):1017–1098.

Heckman, J. J., Ichimura, H., Smith, J., and Todd, P. E. (1996). Sources of selection bias in evaluating social programs: An interpretation of conventional measures and evidence on the effectiveness of matching as a program evaluation method. *Proceedings of the National Academy of Sciences*, 93(23):1341613420.

Heckman, J. J., Ichimura, H., and Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, 64(4):605–654.

Henningsen, M. S., Hægeland, T., and Møen, J. (2012). Estimating the additionality of r&d subsidies using proposal evaluation data to control for firms' r&d intentions. Discussion Papers 729, Statistics Norway, Research Department.

Ho, Y. (2006). *Evaluating the Effectiveness of State R&D Tax Credits*. PhD thesis, University of Pittsburgh.

Hussinger, K. (2008). R&d and subsidies at the firm level: an application of parametric and semiparametric two-step selection models. *Journal of Applied Econometrics*, 23(6):729–747.

Hyytinen, A. and Toivanen, O. (2005). Do financial constraints hold back innovation and growth?: Evidence on the role of public policy. *Research Policy*, 34(9):1385 – 1403.

Jia, J. and Ma, G. (2017). Do r&d tax incentives work? firm-level evidence from china. *China Economic Review*, 46:50 – 66.

Klette, T. and Møen, J. (2012). R&d investment responses to r&d subsidies: a theoretical analysis and a microeconometric study. *World Review of Science, Technology and Sustainable Development*, 9(2-4):169–203.

Kobayashi, Y. (2014). Effect of r&d tax credits for smes in japan: a microeconometric analysis focused on liquidity constraints. *Small Business Economics*, 42(2):311–327.

Lach, S. (2002). Do r&d subsidies stimulate or displace private r&d? evidence from israel. *The Journal of Industrial Economics*, 50(4):369–390.

Lechner, M. (2002). Some practical issues in the evaluation of heterogeneous labour market programmes by matching methods. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 165(1):59–82.

Lokshin, B. and Mohnen, P. (2012). How effective are level-based r&d tax credits? evidence from the netherlands. *Applied Economics*, 44(12):1527–1538.

Mulkay, B. and Mairesse, J. (2013). The R&D tax credit in France: assessment and ex ante evaluation of the 2008 reform. *Oxford Economic Papers*, 65(3):746–766.

OECD (2018). Main science and technology indicators 2018. Technical report, Organisation of Economic Co-operation and Development (OECD).

OECD (2019). Oecd compendium of information on r&d tax incentives. Technical report, Organisation of Economic Co-operation and Development (OECD).

O'Mahony, M. and Vecchi, M. (2009). R&d, knowledge spillovers and company productivity performance. *Research Policy*, 38(1):35 – 44.

Özçelik, E. and Taymaz, E. (2008). R&d support programs in developing countries: The turkish experience. *Research Policy*, 37(2):258 – 275.

Paff, L. A. (2005). State-level r&d tax credits: A firm-level analysis. *The B.E. Journal of Economic Analysis & Policy*, 5(1):1–25.

Parisi, M. L. and Sembenelli, A. (2003). Is private r&d spending sensitive to its price? empirical evidence on panel data for italy. *Empirica*, 30(4):357 – 377.

Rao, N. (2016). Do tax credits stimulate r&d spending? the effect of the r&d tax credit in its first decade. *Journal of Public Economics*, 140:1 – 12.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC Monographs on Statistics and Applied Probability. Chapman and Hall/CRC.

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

Unger, K. (2011). La política de estímulos fiscales a id en méxico. *El Trimestre Económico,*, 78(309):49–85.

Villavicencio, D. (2010). Recent changes in science and technology policy in mexico: Innovation incentives. In Martínez-Piva, J., editor, *Knowdelge Generation and Protection*, chapter 10, pages 263–288. Springer-Verlag, New York.

Villavicencio, D. (2011). Retos para el diseño de políticas en méxico en el marco de la innovación abierta. In Bracamonte Sierra, A. and Contreras Montellano, O., editors, *Ciencia, Tenología e Innovación Para el Desarollo Económico*, chapter 2, pages 73–102. El Colegio de Sonora; Consejo Estatal de Ciencia y Tecnología de Sonora, Hermosillo.

Wasi, N. and Flaaen, A. (2015). Record linkage using stata: Preprocessing, linking, and reviewing utilities. *Stata Journal*, 15(3):672–697(26).

Yang, C.-H., Huang, C.-H., and Hou, T. C.-T. (2012). Tax incentives and r&d activity: Firm-level evidence from taiwan. *Research Policy*, 41(9):1578 – 1588.

Tables

Table 1: Literature Review

Paper	Country	Technique	Effect on Outcomes: Private R&D Spending	Effect on Outcomes: Other
Subsidy				
Hussinger (2008)	DEU	Two Step Model	Positive	Researchers: positive
Bloch and Graversen (2012)	DNK	Panel Data Techniques	Positive	
Henningsen et al. (2012)	NOR	Panel Data Techniques	Positive	
Gonzalez et al. (2005)	ESP	Structural Models	No effect	
Klette and Møen (2012)	NOR	Structural Models	No effect	
Almus and Czarnitzki (2003)	DEU	Matching	Positive	
Czarnitzki and Licht (2006)	DEU	Matching	Positive[/1]	New products: positive
Czarnitzki and Lopes Bento (2012)	BEL, DEU, LUX, ESP	Matching	Positive	
Duguet (2004)	FRA	Matching	Positive	
Gonzalez and Pazo (2008)	ESP	Matching	No effect	
Özçelik and Taymaz (2008)	TUR	Matching	Positive	
Lach (2002)	ISR	Diff-in-Diff	Positive[/2]	
Cerulli and Poti (2012)	ITA	Diff-in-Diff, Matching, Two Step Model	No effect	
Görg and Strobl (2007)	IRL	Diff-in-Diff Matching	Positive[/3]	
Aerts and Schmidt (2008)	DEU, NLD	Diff-in-Diff Matching, Matching	Positive	
Chávez (2019)	MEX	Regression Discontinuity	No effect	Researchers: positive. Patents: no effect
Bronzini and Iachini (2014)	ITA	Regression Discontinuity		Investment: positive[/4]
Tax Incentives				
Cappelen et al. (2012)	NOR	Panel Data Techniques		New products: no effect Patents: no effect

Table 1 Continued: Literature Review

Jia and Ma (2017)	CHN	Panel Data Techniques	Positive/[5]	
Mulkay and Mairesse (2013)	FRA	Panel Data Techniques	Positive	
Calderón (2009)	MEX	Panel Data Techniques, Two Step Model	Positive	
Baghana and Mohnen (2009)	CAN	Structural Models	Positive	
Harris et al. (2009)	GBR	Structural Models, Panel Data Techniques		Output: positive
Rao (2016)	USA	Instrumental Variables	Positive	Researchers: positive
Lokshin and Mohnen (2012)	NLD	Instrumental Variables, Panel Data Techniques	No effect	
Czarnitzki et al. (2011)	CAN	Matching		New products: positive
Duguet (2012)	FRA	Matching	Positive	Researchers: positive
Kobayashi (2014)	JPN	Matching	Positive/[6]	
Corchuelo and Martínez-Ros (2010)	ESP	Matching, Two Step Model	Positive/[7]	
Yang et al. (2012)	TWN	Matching, Instrumental Variables	Positive	
Hægeland and Møen (2007)	NOR	Diff-in-Diff	Positive/[8]	
Paff (2005)	USA	Diff-in-Diff	Positive	
Bozio et al. (2014)	FRA	Diff-in-Diff, Matching	Positive	Patents: no effect
Ho (2006)	USA	Diff-in-Diff, Matching	Positive	Researchers: positive/[9]
Chen et al. (2018)	CHN	Bunching Estimators		Productivity: positive
Dechezleprêtre et al. (2016)	GBR	Regression Discontinuity	Positive	Patents: positive
Subsidy and tax incentive				
Parisi and Sembenelli (2003)	ITA	Panel Data Techniques	Positive	
Carboni (2011)	ITA	Matching	Positive/[10]	

Note: This table shows a summary of previous literature on the effects of R&D subsidies and R&D tax credits on private R&D activities. /[1]The effect is larger in poorer regions. /[2]Significant for small and medium sized firms. /[3]Significant for firms owned by nationals. /[4]Significant for small firms. /[5]Significant for private firms. /[6]Significant for small and medium sized firms. /[7]Significant for large firms. /[8]Significant for firms in low R&D sectors. /[9]Significant for high-tech firms. /[10]The effect of tax credits is larger.

Table 2: Summary of Tax Credit Designs

Country	R&D project approval by non-tax authority	Deductible R&D Expenditures	Cap to deductible R&D R&D Expenditures	In case tax credit is larger than owed tax
Canada	No	15%-35% of firm R&D exp.	No	Carry-forward 20 years
China	No/[1]	50% of firm R&D exp.	For fixed R&D assets	Carry-forward 10 years
France	No	30% of firm R&D exp.	No	Remainder paid as grant
Italy	No	25% of avrg. firm R&D exp. in last three years	10 m. euros	Remainder paid as grant or carry-foward unlimited
Japan	No	12% of firm of R&D exp.	10%-30% of corp. inc. tax	No carry-froward or grant paid
Mexico	Yes/[2]	30% of approved project	Not specified	Carry-forward 10 years
Netherlands	No	16%-32% of firm R&D exp.	No	Carry-forward 1 year
Norway	Yes/[3]	18%-20% of approved project	half m. euros	Remainder paid as grant
Spain	No	8%-40% of R&D exp.	25% of corp. inc. tax., 60% of payroll tax	Remainder paid as grant or carry-forward 20 years
Taiwan	No	35% of firm R&D exp.	Not specified	Remainder paid as grant or carry-forward 5 years
United Kingdom	No	32%-200% of firm R&D exp.	7.5 m. pounds for SMEs	Remainder paid as grant with cap of 16.5% of exp., or carry-foward unlimited
United States	No	(R&D/sales) $\times$ (sales)	No	Carry-forward 15 years

Note: This table shows a summary of the main features of the R&D tax credits in selected countries.
/[1]The firm has to be approved by the Ministry of Science. /[2]By the Council of Science and Technology. /[3]By the Council of Research.
Sources: OECD (2019).

Table 3: CONACYT, EFIDT and PEI budgets

(millon 2018 Mexican Pesos)		
Year	**CONACT budget**	**EFIDT and PEI spending [/1]**
2002	14,833	958
2003	15,813	923
2004	15,566	1,764
2005	15,530	5,090
2006	16,833	6,549
2007	17,266	7,086
2008	20,893	6,741
2009	24,069	2,366
2010	25,956	3,218
2011	27,364	3,071
2012	28,696	2,478
2013	33,621	3,594
2014	39,546	4,551
2015	40,341	4,055
2016	39,211	4,585
2017	31,472	1,826

Note: This table shows CONACYT total budget and spending under the EFIDT and PEI programs.

[/1]From 2002 to 2008, the column shows the total amount of tax credits awarded under the EFIDT program. From 2009 to 2017, the column shows the total amoun of subsidies awarded under the PEI program.

Sources: CONACYT (2004–2017).

Table 4: Number of Observations

Number of observations		
	Projects	**Firms**
Subsidy		
In CONACYT database	6,361	2,842
Matched to Census	3,658	1,407
	57.5%	*49.5%*
Tax Credit		
In CONACYT database	8,428	1,316
Matched to Census	6,453	932
	76.6%	*70.8%*

Note: This table shows the number of observations in the EFIDT and PEI datasets and those firms successfully matched to the Economic Census.
Sources: EFIDT database, PEI database and 2014, 2009 Economic Census.

Table 5: Propensity score model on the EFIDT tax credit

Dependent variable: dummy= 1 for firms receiving the EFIDT tax credit	**Coef.**	**Std. Err.**	**P. Val.**
Firm age	0.079***	0.030	0.008
Profit Margin	0.176**	0.078	0.025
Paid Interests / Total Expenditures	0.164	0.103	0.112
Dummy for innovation department	0.355***	0.050	0.000
Dummy for personnel trainning of new technologies	0.223***	0.062	0.000
Dummy for positive research spending	0.467***	0.061	0.000
Dummy for gov subsidy reception	-0.044	0.171	0.798
Log of total employees	0.285***	0.021	0.000
Dummy for foreign capital participation	-0.036	0.066	0.585
Exports / Total Sales	-0.125	0.090	0.163
Fixed assets per employee	0.000***	0.000	0.000
Dummy for agriculture and extraction activities	-0.136	0.246	0.582
Dummy for construction activities	-0.448	0.291	0.124
Dummy for low tech industries	0.608***	0.094	0.000
Dummy for high tech industries	0.981***	0.094	0.000
Dummy for commerce activities	0.462***	0.099	0.000
Dummy for service sector	0.000	(omitted)	
Dummy for region west	0.210***	0.062	0.001
Dummy for region northwest	-0.150*	0.079	0.059
Dummy for region northeast	0.002	0.065	0.970
Dummy for region southeast	0.080	0.099	0.424
Dummy for region south	0.049	0.127	0.702
Dummy for region center	0.000	(omitted)	
Intercept	-4.574***	0.155	0.000

Note: This table shows the propensity score model on the firms that received the EFIDT tax credit from 2005 to 2008.

*** Statistically significant difference at the 1 percent confidence level.

** Statistically significant difference at the 5 percent confidence level.

Sources: EFIDT tax credit database and 2009, 2004 Economic Census.

Table 6: Cross-section matching results for the EFIDT tax credit

Average treatment effects	Mean diff.	Std. Err.	T-test	P. Val.
Outcome				
Was there personnel working in innovation activities in 2008? (Yes=1, No=0)				
One to one matching (no replacement)	0.151***	0.027	5.674	0.000
Nearest neighbor matching (10)	0.152***	0.020	7.520	0.000
Kernel matching	0.244***	0.018	13.334	0.000
Did the firm register patents in 2008? (Yes=1, No=0)				
One to one matching (no replacement)	0.087***	0.020	4.380	0.000
Nearest neighbor matching (10)	0.080***	0.017	4.625	0.000
Kernel matching	0.087***	0.017	5.265	0.000
Untreated observations	40,904			
Treated observations	503			
Total observations	41,407			

Note: This table shows β estimates from equation (2) for the EFIDT tax credit.

*** Statistically significant difference at the 1 percent confidence level.

** Statistically significant difference at the 5 percent confidence level.

* Statistically significant difference at the 10 percent confidence level.

Sources: EFIDT tax credit database and 2009, 2004 Economic Census.

Table 7: Diff-in-diff matching results for the EFIDT tax credit

	Average treatment effects			
	Mean diff.	**Std. Err.**	**T-test**	**P. Val.**
Outcome				
Was there personnel working in innovation activities in 2008? (Yes=1, No=0)				
One to one matching (no replacement)	0.163***	0.037	4.352	0.000
Nearest neighbor matching (10)	0.132***	0.030	4.434	0.000
Kernel matching	0.086***	0.027	3.225	0.001
Did the firm register patents in 2008? (Yes=1, No=0)				
One to one matching (no replacement)	-0.014	0.032	0.440	0.660
Nearest neighbor matching (10)	0.001	0.027	0.044	0.965
Kernel matching	-0.012	0.025	0.462	0.644
Untreated observations	40,904			
Treated observations	503			
Total observations	41,407			

Note: This table shows β estimates from equation (3) for the EFIDT tax credit.

*** Statistically significant difference at the 1 percent confidence level.

** Statistically significant difference at the 5 percent confidence level.

* Statistically significant difference at the 10 percent confidence level.

Sources: EFIDT tax credit database and 2009, 2004 Economic Census.

Table 8: Propensity score model on the PEI subsidy

Dependent variable: dummy= 1 for firms receiving the PEI subsidy			
	Coef.	**Std. Err.**	**P. Val.**
Firm age	0.089**	0.037	0.017
Profit Margin	0.296	0.171	0.083
Paid Interests / Total Expenditures	0.720	1.271	0.571
Dummy for innovation department	0.413***	0.065	0.000
Dummy for colaboration with research centers	0.652***	0.070	0.000
Dummy for gov subsidy reception	0.046	0.180	0.798
Log of total employees	0.167***	0.024	0.000
Exports / Total Sales	-0.214**	0.107	0.046
Fixed assets per employee	0.000**	0.000	0.042
Dummy for lack of innovation funding	-0.172**	0.081	0.034
Dummy for agriculture and extraction activities	-0.165	0.301	0.584
Dummy for construction activities	0.150	0.199	0.450
Dummy for low tech industries	0.861***	0.147	0.000
Dummy for high tech industries	1.274***	0.145	0.000
Dummy for commerce activities	0.609***	0.158	0.000
Dummy for service sector	0.000	(omitted)	
Dummy for region west	0.085	0.089	0.337
Dummy for region northwest	0.205**	0.092	0.027
Dummy for region northeast	0.099	0.085	0.247
Dummy for region southeast	-0.240	0.194	0.216
Dummy for region south	0.343***	0.133	0.010
Dummy for region center	0.000	(omitted)	
Intercept	-4.450***	0.206	0.000

Note: This table shows the propensity score model on the firms that received the PEI subsidy from 2009 to 2013.

*** Statistically significant difference at the 1 percent confidence level.

** Statistically significant difference at the 5 percent confidence level.

Sources: PEI subsidy database and 2014, 2009 Economic Census.

Table 9: Cross-section matching results for the PEI subsidy

	Average treatment effects			
	Mean diff.	**Std. Err.**	**T-test**	**P. Val.**
Outcome				
Was there personnel working in innovation activities in 2013? (Yes=1, No=0)				
One to one matching (no replacement)	0.307***	0.043	7.134	0.000
Nearest neighbor matching (10)	0.360***	0.035	10.392	0.000
Kernel matching	0.404***	0.032	12.605	0.000
Did the firm register patents in 2013? (Yes=1, No=0)				
One to one matching (no replacement)	0.150***	0.037	4.050	0.000
Nearest neighbor matching (10)	0.160***	0.031	5.104	0.000
Kernel matching	0.193***	0.030	6.416	0.000
Untreated observations	36,711			
Treated observations	241			
Total observations	36,952			

Note: This table shows β estimates from equation (2) for the PEI subsidy.

*** Statistically significant difference at the 1 percent confidence level.

** Statistically significant difference at the 5 percent confidence level.

* Statistically significant difference at the 10 percent confidence level.

Sources: PEI subsidy database and 2014, 2009 Economic Census.

Table 10: Diff-in-diff matching results for the PEI subsidy

	Average treatment effects			
	Mean diff.	**Std. Err.**	**T-test**	**P. Val.**
Outcome				
Was there personnel working in innovation activities in 2013? (Yes=1, No=0)				
One to one matching (no replacement)	0.220***	0.055	3.970	0.000
Nearest neighbor matching (10)	0.277***	0.044	6.336	0.000
Kernel matching	0.131***	0.040	3.299	0.001
Did the firm register patents in 2013? (Yes=1, No=0)				
One to one matching (no replacement)	0.114*	0.064	1.763	0.079
Nearest neighbor matching (10)	0.100*	0.053	1.881	0.061
Kernel matching	0.069	0.043	1.593	0.113
Untreated observations	36,711			
Treated observations	241			
Total observations	36,952			

Note: This table shows β estimates from equation (3) for the PEI subsidy.

*** Statistically significant difference at the 1 percent confidence level.

** Statistically significant difference at the 5 percent confidence level.

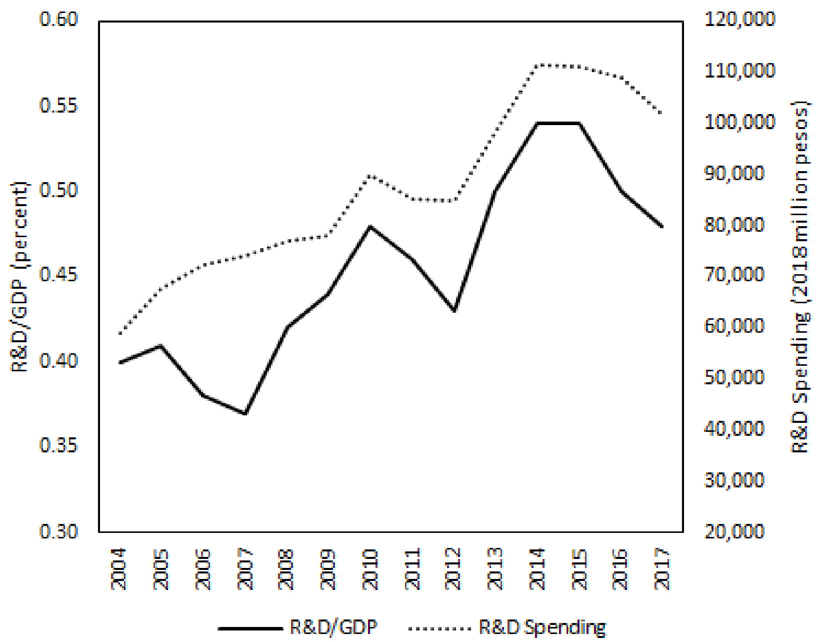
* Statistically significant difference at the 10 percent confidence level.

Sources: PEI subsidy database and 2014, 2009 Economic Census.

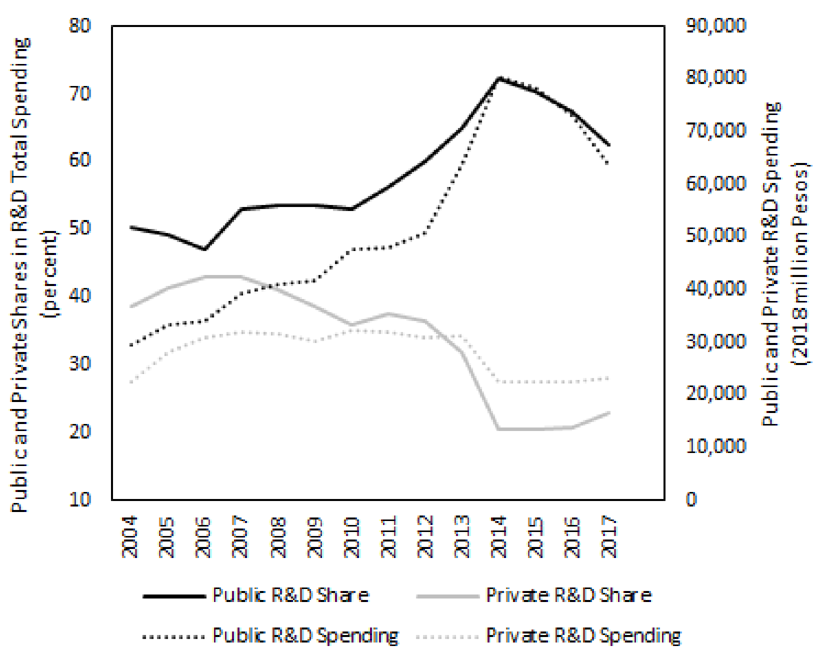
Figures

Figure 1: R&D spending

(a) Total R&D

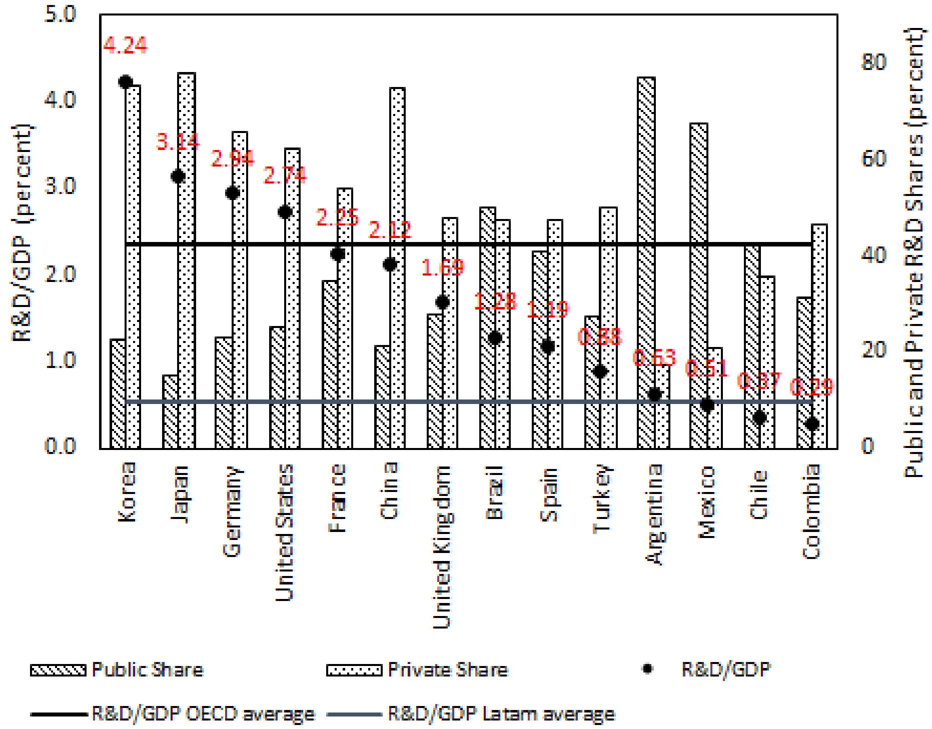


(b) Public and Private R&D



Note: This figure shows R&D spending in Mexico in the 2004 to 2017 period.
Sources: CONACYT (2017).

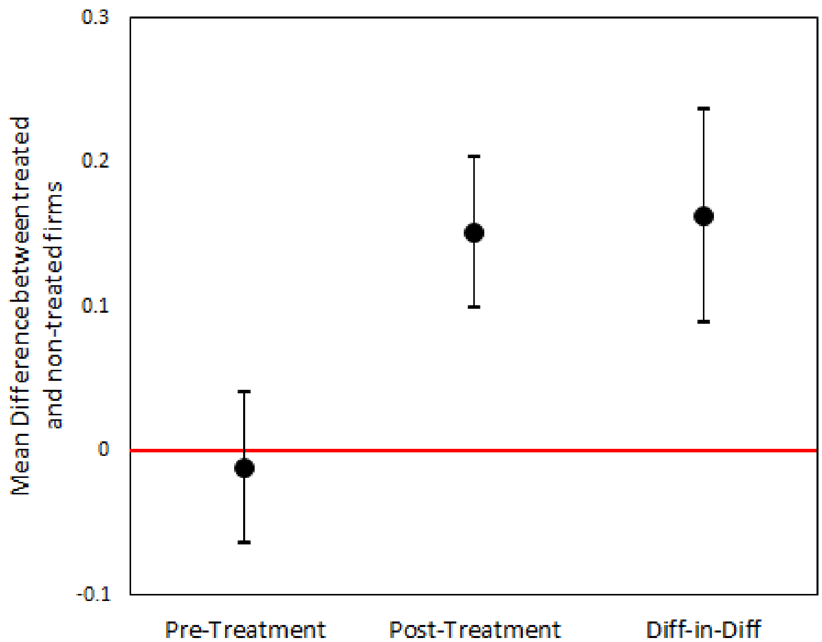
Figure 2: R&D international comparison



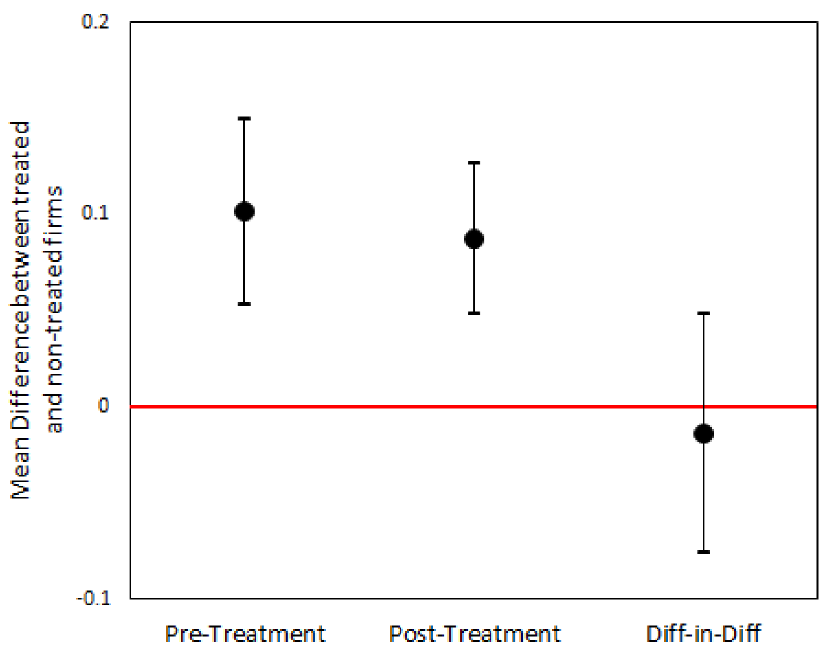
Note: This figure shows the R&D/GDP ratio in selected countries, as well as the public and private shares of R&D funding within countries.
Sources: CONACYT (2017) and OECD (2018).

Figure 3: EFIDT tax credit estimates

(a) Outcome: personnel working in innovation activities



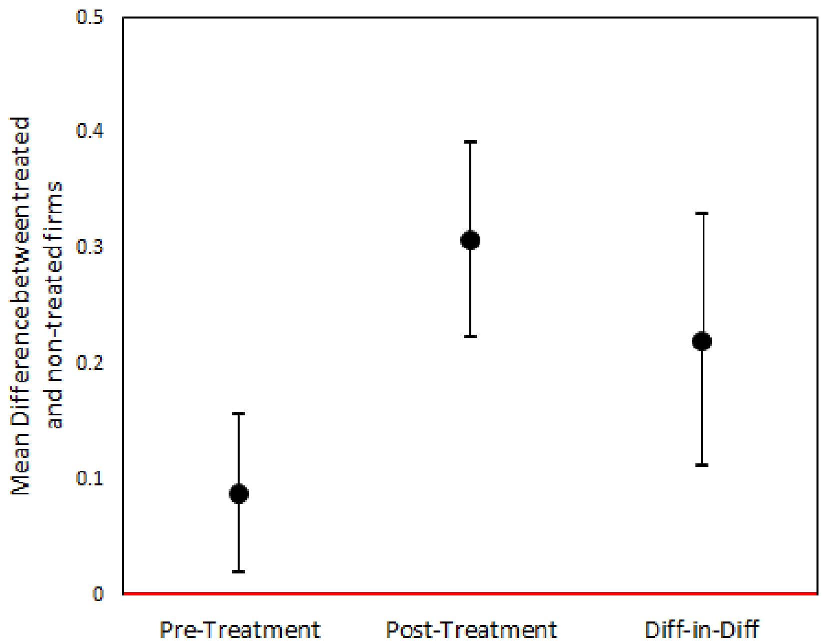
(b) Outcome: patent registration



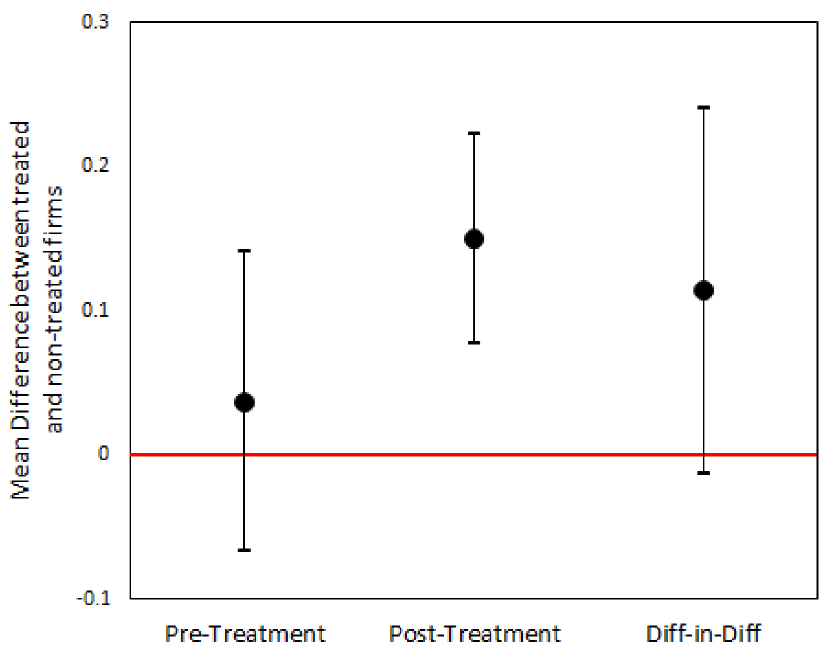
Note: This figure shows the point estimate of the mean differences between the EFIDT tax credit awarded and non-awarded firms in the pre-treatment and post-treatment periods. Differences were obtained with the one-to-one matching estimate of equation (2). In addition, the figure shows the point estimate from the DID estimates of equation (3). Lines around the point estimates show the 95 percent confidence intervals.
Sources: EFIDT tax credit database and 2009, 2004 Economic Censuses.

Figure 4: PEI subsidy estimates

(a) Outcome: personnel working in innovation activities



(b) Outcome: patent registration



Note: This figure shows the point estimate of the mean differences between the PEI subsidy awarded and non-awarded firms in the pre-treatment and post-treatment periods. Differences were obtained with the one-to-one matching estimate of equation (2). In addition, the figure shows the point estimate from the DID estimates of equation (3). Lines around the point estimates show the 95 percent confidence intervals.

Sources: PEI subsidy database and 2014, 2009 Economic Censuses.

A Merging Datasets

I merge two CONACYT datasets to the 2014, 2009 and 2004 Economic Censuses. The CONACYT datasets contain: 1) data on the firms that were awarded with the EFIDT tax credit in the 2004 to 2008 period, and 2) data on the companies that submitted projects to the PEI dataset in the 2009 to 2013 period. CONACYT provided information on four firm identifying variables: name, address, economic sector and size. I use these variables to merge firms to the economic census datasets.

I start by standardizing firm names in both datasets. First, I substitute all letters with an accent, such as "á" with the same letter without the accent ("a"). I also substitute "ñ" by "n". Then, I drop all special characters such as "&", "%", "-", etc. In addition, I drop all dots, commas and punctuation marks. I also drop all blank spaces that are not between words and convert all letters to upper case. Finally, I drop all acronyms that denote the firm legal status, such as "SA DE CV", "SC DE RL", "AC", among many others.

Once names in both datasets are standardized, I proceed to the actual merging. First I merge firms by formal name. If I do not find an exact match with the formal name, I use the name shown at the establishment. If I do not find a match with the formal and establishment name within the state, I move to the other states until I find a match with either the formal or establishment name. The Economic Census dataset contains a registry of physical establishments. Since one firm can own many establishments and have them registered under the same name, matching by name can bring many establishment matches for the same firm in the CONCACYT datasets. When this happens, I keep the establishment with the highest number of employees, given that the sector and address are also matched.

There are many instances in which I cannot get an exact match using names. Company names might be misspelled in the datasets. For instance, instead of a correct company name such as "EMPRESAS INNOVADORAS MEXICANAS", we could have "EMPRESA INNOVADORAS MEXICANAS", "EMPRESAS INNOV MEXICANAS", "EMPRESAS INNOV MEX", among many other variants. To face this matter, I use the *recklink2* Stata command.[18] The command uses probabilistic matching and throws a set of likely matches ranked by a score. I use this probabilistic matching for the firms that I could not get an exact name match. I follow the same process described above, i.e. merging by name, address and sector. This process brings many probable matches for one firm. For each firm, I check its probable

[18]For detailed information on this Stata command see Wasi and Flaaen (2015).

matches one by one and select the one I think is the best based on the name, address sector and firm size.

I merge the firms in the EFIDT dataset to the 2009 Economic Census, and those of the PEI dataset to the 2014 Economic Census. These censuses give the post-treatment data for each of the EFIDT and PEI programs. However, I must find the merged EFIDT and PEI firms in the previous census datasets to get the pre-treatment information. The 2014 and 2009 censuses count with a variable that identifies firms in both censuses. I use this variable to find the PEI firms that I merged to the 2014 Economic Census in the 2009 census. The 2004 census does not have a variable that identifies firms across different census datasets. To find the EFIDT firms that I merged to the 2009 Economic Census in the 2004 census, I use the the identifying variables created by Busso et al. (2018).