



**UNITED NATIONS
UNIVERSITY**

UNU-MERIT

Working Paper Series

#2010-055

**Public R&D subsidies and productivity:
Evidence from firm-level data in Quebec**

Rufin Baghana

Public R&D Subsidies and Productivity: Evidence from firm-level data in Quebec

Rufin Baghana*

(Ministère des Finances, Québec)

August, 2010

ABSTRACT

This paper analyses empirically the impacts of public R&D grants on private R&D investments and on the productivity growth of the manufacturing firms in a context where fiscal incentives are present. Using the conditional semiparametric difference-in-differences estimator on longitudinal data from Quebec we show that firms that use public grants for R&D in conjunction with tax credits for R&D perform better in terms of R&D input additionality than firms that use only tax credits for R&D. We then use a production function to assess the effectiveness of public R&D grants in the productivity growth of firms. We find that for each additional dollar of public R&D grant, output increases by 0.134 dollars. We conclude that the additional return of direct subsidies is positive but lower than the return on the R&D financed by own funds or R&D tax credits.

Keywords

R&D, Public subsidies, Quebec, Productivity, Difference-in-differences

JEL: H25 en O32.

I am grateful to Pierre Mohnen for helpful suggestions, comments and comprehensive reviews on earlier drafts of this paper. I thank the ISQ Science, Technology and Innovation team, particularly Brigitte Poussart for her help in setting up the dataset. All mistakes remain mine.

* Corresponding author: rufin.baghana@finances.gouv.qc.ca

UNU-MERIT Working Papers
ISSN 1871-9872

**Maastricht Economic and social Research and training centre on Innovation and
Technology, UNU-MERIT**

*UNU-MERIT Working Papers intend to disseminate preliminary results of research
carried out at the Centre to stimulate discussion on the issues raised.*

Disclaimer: The views expressed in this article are the sole responsibility of the author and do not necessarily reflect those of the Ministry of Finance of Quebec

INTRODUCTION

Most countries now recognize the need for supporting and strengthening innovation activities in order to ensure productivity growth and hence long-term prosperity and competitiveness in the global economy. Neo-classical theory explains the need for government involvement in technological advance and innovation activities by the Arrow-Nelson rationale, which is the key argument of the neo-classical approach of market failure (Nelson (1959), Arrow (1962)). Essentially, the market failure approach points to the underinvestment from private firms in R&D leading to a supply of knowledge in the economy that is less than what would be socially desirable. In an effort to fill this gap, several governments implemented various incentive programs targeted to private firms such as fiscal-based frameworks or more direct public schemes such as grants, loans, procurements, venture capital investments or other instruments.

In Canada, as in most OECD countries, the orientation taken in the last decades by governments shows a clear tendency towards fiscal incentives such as tax credits. This choice has been motivated mainly by the distinctive feature of tax credits of being neutral with respect to the choice of the industry or of projects. However, this neutrality also has the unattractive feature that companies tend to choose systematically the projects with the highest rates of return leaving so aside the least profitable projects. For this reason, it is not surprising that many governments spend every year, in addition to their fiscal incentive programmes, several millions of dollars in specific direct subsidies programmes to firms in order to increase private R&D spending in targeted industries or in projects that would not be carried out without some type of subsidy. Some authors recommend rather the use of both policy instruments in a more integrated approach since there is evidence that tax incentives stimulate mainly R&D projects that involve more applied or short-term research while direct subsidies affect projects which research is more basic or long-run (Guellec and Van Pottelsberghe (2000)). In any case, the question of the goal attainment of the objectives at which direct subsidies in a context where fiscal incentives are already present may be asked. In particular, this refers mainly to the two following questions: First, do public R&D grants result in increased R&D spending (input

additionality) from firms that already receive tax credits and, second, what is the impact of these publicly funded R&D on productivity growth? This study proposes to add to the evidence on this literature by exploring empirically these questions using longitudinal data from manufacturing firms in Quebec in the range 1997-2004.

To date, these research questions have not been previously investigated with data from Quebec and, at the Canadian level, there are only a few studies that address the effectiveness of public schemes to R&D and, not surprisingly, they all focus on R&D tax credits given the emphasis of the Canadian government support programmes on the latter¹. The only Canadian studies that address the effectiveness of subsidies are the studies by Pagé (1995) and Bérubé and Mohnen (2009), but in those cases too, the focus is not on the input additionality of public R&D grants nor on their impact on firms' productivity, but rather on innovation output. To the benefit of this study, the methodological approach used by Bérubé and Mohnen (2009) has however a realistic feature in that, by evaluating the effectiveness of public R&D grants, it takes into account the effects of the other public incentives such as tax credits. The present study draws on this methodological approach for assessing the input additionality of public R&D grants. There exist a considerable body of research in other countries on the two main questions treated by this study but they traditionally have been treated separately, i.e. there are, on

¹ See McFetridge, Donald G., and Jacek P. Warda, 1983. *Canadian R & D incentives: Their adequacy and impact* (Canadian Tax Foundation, Toronto, Canada)., Mansfield, Edwin , and Lorne Switzer, 1985, How Effective Are Canada's Direct Tax Incentives for R and D?, *Canadian Public Policy* 11, 241-246.a, Mansfield, Edwin, and Lorne Switzer, 1985, The effects of R&D tax credits and allowances in Canada, *Research Policy* 14, 97-107.b, Bernstein, I. Jeffrey, 1986, The Effect of Direct and Indirect Tax Incentives on Canadian Industrial R&D Expenditures, *Canadian Public Policy* 12, 438-448., Lebeau, Daniel, 1996, Les mesures d'aide fiscale à la R-D et les entreprises québécoises, in Conseil de la science et de la technologie du Québec, ed.: *L' Efficacite des mesures d'aide fiscale à la R-D des entreprises du Canada et du Québec* (Gouvernement du Québec, Sainte-Foy, Québec)., Dagenais, Marcel, Pierre Mohnen, and Pierre Therrien, 2004, Les firmes canadiennes repondent-elles aux incitations fiscales a la recherche-developpement?, *L'Actualite Economique/Revue D'Analyse Economique* 80, 175-205., Dahlby, Bev 2005, A Framework for Evaluating Provincial R&D Tax Subsidies, *Canadian Public Policy* 31, 45-58., Hanel, Petr , Dirk Czarnitzki, and Julio Miguel Rosa, 2005, Evaluating the Impact of R&D Tax Credits on Innovation: A Microeconomic Study on Canadian Firms, in ZEW discussion paper 04-77, ed.: ZEW (Centre for European Economic Research (ZEW), Mannheim)., Baghana, Rufin, and Pierre Mohnen, 2009, Effectiveness of R&D tax incentives in small and large enterprises in Québec, *Small Business Economics* 33, 91-107. and Baghana, Rufin, 2008, Une évaluation de l'impact du programme québécois de superdéductions à la R-D, in Institut de la Statistique du Québec, ed.: *Compendium d'indicateurs de l'activité scientifique et technologique au Québec. Édition 2008* (Institut de la Statistique du Québec, Quebec).

the one hand, studies that address the question of the additionality of public subsidies and in most cases they did not take into account the effects of tax credits in country with both type of government incentives. The evidence from these studies is mixed and this casts doubt on the question whether public subsidies increase firms' private spending in R&D or to the contrary, crowds out this spending. Klette, Moen and Griliches (2000) or David, Hall and Toole (2000) provide good reviews of these studies. On the other hand, there are studies that analyze the impact of firms' total R&D spending (privately and publicly funded) on productivity growth. Most of them estimate an elasticity of the output with respect to the R&D effort or a rate of return to firms' R&D expenditures. In the latter case, on which this study focuses, empirical estimates of the rate of return vary between 20% and 50%. Examples include Griliches and Mairesse (1991), Wakelin (2001), Smith, Dilling-Hansen, Eriksson and Madsen (2004) and Maté-Garcia and Rodriguez-Fernandez (2008).

The novel aspect of this study is that it combines the two research questions into one analysis by taking advantage of a set of rich microdata from manufacturing firms in Quebec. More specifically, to answer the first question we use a semiparametric matching estimator, more specifically the conditional semiparametric difference-in-differences estimator and, following Bérubé and Mohnen (2009), we focus on the comparison between the intensity of R&D of firms that receive R&D grants in addition to R&D tax credits and the intensity of R&D of firms that receive only R&D tax credits. Hence, the estimated differences in the intensity of R&D of both groups permit to assess the input additionality of R&D grants in the presence of tax credits. For the second question, we subsequently use these estimated differences in a production function controlling also for other variables in order to evaluate the impact of R&D grants to firm's growth productivity.

The paper is organized as follows: Section 1 provides an overview of the R&D tax incentives in Quebec. Section 2 sketches the methodological approach which is undertaken in two steps: The first step regarding the matching estimator framework from which we shall assess the input additionality of R&D grants and the second step

regarding the production function framework from which we shall assess the impact of privately and publicly funded R&D on productivity growth. In section 3, we first describe briefly the three data sources used and the merging process of these data before explaining variables construction. Section 4 presents the empirical analysis and the results. Section 5 concludes after a discussion.

1. PUBLIC DIRECT SUBSIDIES TOWARD R&D IN QUEBEC

Over the last thirty years, most of the governmental financial assistance to R&D activities in Quebec from both the government of Quebec (the provincial government) and the Canadian government (the federal government), has been granted to firms in the form of fiscal incentives through Quebec's and federal scientific research and experimental development (SR&ED) programmes and this has resulted in a significant decrease in the amounts of grants awarded annually to firms. However, both governments maintained through years a few direct funding programmes to R&D in addition to their SR&ED programmes. At the provincial level there is no agency dedicated to the management of public funding programmes. Financial assistance is attributed to firms by the ministries and some government agencies. In most cases, the assistance from these institutions is granted for R&D expenditures that do not entitle firms to claim tax credits from SR&ED. The Ministry of Economic Development Innovation and Export Trade (MDEIE²) is the main provincial institution for public direct funding to firms³. In addition to the grants from the government of Quebec, firms can also apply for federal grants through the Industrial Research Assistance Program (NRC-IRAP) which is the most important federal programme administered by National Research Council of Canada. Firms may also benefit from grants or contracts programmes from other federal agencies acting in specific areas such as energy or military (e.g. Mines and Resources Grants, Canadian Space Agency Grants, Defence Industry Productivity Program).

Table 1-1 shows the evolution in the period 1997-2004 of R&D grants and contracts from both the Canadian federal government and the Government of Quebec to manufacturing firms performing R&D activities. The overall value of grants and contracts dropped from CAN 100 M\$ to only CAN 7 M\$ in 2004 and this fall originates essentially in a reduction of grants and contracts from the Canadian government between 1997 and 2001. Such a fall points out to policies adopted by the federal and provincial

² MDEIE: Ministère du Développement économique, de l'Innovation et de l'Exportation.

³ This role was previously played by the ministère de l'Industrie et du Commerce (MIC) and the ministère de la Recherche, de la Science et de la Technologie (MRST) which were merged in 2003 to form the ministère du Développement économique et régional et de la Recherche. The latter was replaced in turn by the MDEIE from 2006.

governments in Canada in the last years and reflects the preference over time by both levels of government for fiscal incentives policy over R&D grants incentives. This is illustrated clearly by observing the evolution of the ratio of the combined federal-provincial R&D grants and contracts to the combined federal-provincial tax credits received by firms in this period as indicated in the last line of **Table 1-1**. From more than 46 % in 1997, this ratio is no longer more than 2.2 % in 2004.

Table 1-1: Comparative evolution of public direct incentives and tax credits aimed at R&D in the manufacturing sector in Quebec, 1997-2004

	Millions of Canadian dollars							
	1997	1998	1999	2000	2001	2002	2003	2004
Provincial and federal R&D grants								
— Provincial	3	4	2	3	5	11	4	4
— Federal	97	36	30	20	5	5	3	3
Total R&D grants	100	40	32	23	10	16	7	7
R&D grants / Tax credits	46.3%	16.7%	13.7%	5.5%	3.1%	4.6%	2.1%	2.2%

Sources: Ministry of Revenue of Quebec and Statistics Canada, calculations by the author

Note: Calculations in this table have been done from the data of the survey on Research and Development in Canadian Industry (RDCI) linked to the administrative data from the Ministry of Revenue of Quebec. Firms that could not be matched in both samples were discarded. These firms represent about 9.5 % of the total number of firms.

Table 1-2 and **Table 1-3** summarize the use of the public grants by class size and by technology level. **Table 1-2** reveals an asymmetric distribution of R&D grants between large and small firms. Indeed, more than 87 % of R&D recipients are small firms (less than 250 employees) with only 12.9 % of the total value of attributed R&D grants. However, 82.8 % of the total amount of attributed R&D grants goes to the largest firms (more than 500 employees) present in the sample in the proportion of only 8.6 %. The pattern of this asymmetric distribution is also comparable to that of the use of tax credits by firms in the manufacturing sector as stated by Baghana and Mohnen (2009) and refers to questions concerning firms accessibility to public funding. When questioned about this issue, most companies' administrators of the small enterprises invoke reasons such as the complexity of the application procedure of the government programmes, the lack of time

and human resources necessary to follow up technological activities, and the high consultancy that exceed in some cases the benefits of the programme, etc.

Table 1-2: Use of public grants by class size in the manufacturing sector, 1997-2004

	% Grants	Value of grants			Ratio grants /tax
	n	% Québec	% federal	Total	%
≥ 500 employees	8.6	43.2	89.9	82.8	17.2
250 to 499 employees	4.2	11.5	3.0	4.3	4.9
50 to 249 employees	28.5	29.6	3.8	7.7	3.2
< 50 employees	58.7	15.7	3.3	5.2	2.3
TOTAL	100.0	100.0	100.0	100.0	9.6

Sources: Ministry of Revenue of Quebec and Statistics Canada, compiled by the Institut de la Statistique du Québec and by the author

Turning now to the distribution of R&D grants by technology group in the manufacturing sector, it can be stated in **Table 1-3** that, more than 74.4 % of the firms in the sample are in groups other than the high technology group making up only 15.1 % of the total value of attributed R&D grants. However, the high technology group which represents 25.6 % of the total number of R&D grants recipients gets almost 85.0 % of the total amount of attributed R&D grants. Clearly, these differences suggest that most federal and provincial direct support programmes are used by the high technology group which is R&D-intensive.

Table 1-3: Use of public grants by technology level⁴, 1997-2004

	% Grants	Value of grants			Ratio grants /tax
	n	% Québec	% federal	Total	%
High-technology	25.6	41.6	92.8	84.9	16.5
Medium-high-technology	30.7	21.0	4.9	7.4	4.6
Medium-low-technology	34.5	35.3	1.9	7.0	2.3
Low-technology	9.2	2.1	0.4	0.7	1.2
TOTAL	100.0	100.0	100.0	100.0	9.6

Sources: Ministry of Revenue of Quebec and Statistics Canada, compiled by the Institut de la Statistique du Québec and by the author

⁴ Based on OECD classification of manufacturing industries using International Standard Industrial Classification (ISIC) rev.3 activity breakdown (see classification in **Table D-4**)

2. METHODOLOGICAL APPROACH

The methodology comprises two steps, each investigating one of the two main research questions of this study, i.e. whether public R&D grants result in increased R&D spending from firms (input additionality), and what is the impact of privately and publicly funded R&D on productivity growth. In the first step, input additionality is evaluated by computing an overall average effect on knowledge capital using the individual effects for each recipient firm estimated by the semiparametric matching estimator. The estimated individual effects are then subtracted from firm's knowledge capital in order to isolate, on the one hand, the knowledge capital induced by R&D expenditures funded out of firms' own pockets and by R&D tax credits and, on the other hand, the knowledge capital induced by R&D expenditures funded by public grants. In the second step, a productivity growth function is estimated and the two knowledge capital components from the first step are included along with other variables as regressors in the estimation. This two-step approach has been previously used by Czarnitzki and Licht (2006). However, our study is departing from theirs by the matching scheme used. Furthermore, their study is based on a different specification for the production function as it focuses on innovation output. Indeed, they introduce the two knowledge capital components resulting from the matching process as innovation input in a Griliches invention production function linking innovation output to innovation input. The focus of this study is rather on the growth productivity of firms.

Before describing in detail the matching estimator and the production function frameworks used respectively in the first and second step of the methodological approach, let's first present formally this two-step approach. As starting point, consider a prototypical model of economic choice. Let Y_i^0 and Y_i^1 denote two potential outcomes of the firm i respectively with treatment and without treatment, that is, in our case, the outcome of firm i if it claims only tax credits and its outcome when it receives both tax credits and R&D grants. $T_{gi} = \{0,1\}$, is the indicator of exposure to treatment that takes the value 1 if the firm is treated and 0 otherwise. We can write the following:

$$\begin{aligned} Y_i^0 &= h^0(X_i) + \varepsilon_i, & \text{if } T_{gi} &= 0 \\ Y_i^1 &= h^1(X_i) + \tau_i + \varepsilon_i, & \text{if } T_{gi} &= 1 \end{aligned} \quad (i = 1, \dots, n) \quad (2-1)$$

where τ_i is the impact or the effect of the treatment to be estimated. The aggregation of these individual effects yields the widely known ATT (average treatment effect on the treated). For the estimation of this effect, we adopt the approach of the matching estimator. The advantage of using the matching estimator is that it allows $h(X)$ to have unknown functional form rather than being, for instance, a linear combination of X such as in the case of parametric sample selection models.

For a firm i , the gain from treatment is given by:

$$Y_i^1 - Y_i^0 = \tau_i \quad (2-2)$$

In the second step, a productivity function Q is estimated using the following regressors:

$Y_i^0 = Y_i^1 - \tau_i$, i.e. the knowledge capital, taken as R&D intensity, induced by private R&D expenditures and by R&D tax credits, τ_i , the knowledge capital induced by R&D expenditures funded by public grants and others variables Z_i affecting productivity. This relationship is given by:

$$Q_i = f(Y_i^0, \tau_i, Z_i) \quad (2-3)$$

2.1. Step 1: The matching estimator framework

2.1.1. The fundamental evaluation problem and the selection bias problem

The estimation of equation (2-2) is not straightforward because only one of the two possible outcomes is observed at the same time i.e. Y_i^0 or Y_i^1 . In other words, $Y = T_{gi}Y_i^1 + (1 - T_{gi})Y_i^0$. This is known in the literature as the fundamental evaluation problem.

In order to solve this issue, treatment impacts measures are approximated by means impacts. Hence, equation (2-2) is rewritten as the average treatment effect on the treated firms (ATT) and is given by:

$$ATT = E(Y^1|T_g = 1, X) - E(Y^0|T_g = 1, X) \quad (2-4)$$

where X is a set of comparable characteristics. If the mean $E(Y^1|T_g = 1, X)$ is known, it is not the case for $E(Y^0|T_g = 1, X)$ the no-treatment outcome of programme participants. This unknown outcome may be approximated by the outcome of the nonparticipants $E(Y^0|T_g = 0, X)$ if such information is available. However, another important issue arises when using this approximation. Indeed, unless the case of randomly selected firms, the outcome of treated firms in case of non-treatment is expected to differ from the outcome of nonparticipants i.e. $E(Y^0|T_g = 1, X) \neq E(Y^0|T_g = 0, X)$. As a consequence, equation (2-4) using such an approximation can possibly be biased. This problem is known as the selection bias. Several models in the statistics and econometrics literature propose identifying assumptions to estimate $E(Y^0|T_g = 1, X)$ i.e. what programme participants would experience had they not participated.

2.1.2. The matching methods

To address the fundamental evaluation problem and the selection bias problem, different models referred to as models of economic choice have been proposed, among them are the matching estimators. Matching is based, as stated by Heckman, Ichimura and Todd (1997), on the intuitively attractive idea of contrasting the outcomes of programme participants with the outcomes of “comparable” nonparticipants. There are two approaches to matching, first, the parametric matching approach and, second, the non parametric approach to matching. While in the former approach one needs to estimate consistently the two counterfactuals’ situations by assuming specific functional forms, the latter approach, which is used in this study, has the advantage of not requiring doing so.

From the introduction of the hypothesis of conditional independence assumption (CIA) by Rubin (1977), non parametric matching methods have been improved and became one of the most important models used in the evaluation literature especially in the labor market area. The CIA states that for individuals with the same set of exogenous characteristics V , the participation and the potential outcome are independent. When this condition is satisfied, the selection bias is eliminated and the outcome of treated firms in case of non-treatment equals the outcome of nonparticipants that is, $E(Y^0|T_g = 1, X) = E(Y^0|T_g = 0, X)$. The CIA is expressed as follows:

$$(Y^1, Y^0) \perp T_g | V \quad (2-5)$$

where V is the multidimensional vector of pre-treatment characteristics. This condition states the independence of outcomes (Y^1, Y^0) from the treatment variables given a set of conditioning variables V . We recall that the vector V is formed of observable variables X that affect the outcomes of other observable variables Z determining choices. However, the non-parametric matching methods do not make any distinction between the X and Z variables.

In order to control for the multidimensional problem, Rosenbaum and Rubin (1983) introduced in their seminal paper the method of propensity score matching. Propensity score is defined as the probability that a treatment was assigned to a unit given information in the control variables. In other words, it is the probability that an individual takes treatment. Matching on the propensity is a useful alternative because it reduces the multidimensional problem to one dimension. Formally, the propensity score is given by:

$$p(V) = \Pr\{T_g = 1|V\} = E\{T_g|V\} \quad (2-6)$$

To derive this equation, Rosenbaum and Rubin (1983) assume the following:

Lemma 1: Balancing of pre-treatment variables given the propensity score. If $p(V)$ is the propensity score, then

$$(T_g \perp V | p(V)) \quad (2-7)$$

Lemma 2: Unconfoundedness given the propensity score. Suppose that assignment to treatment is unconfounded, i.e.

$$(Y^1, Y^0) \perp T_g | V \quad (2-8)$$

Then under (2-7) and (2-8) it can be showed that assignment to treatment is unconfounded given the propensity score, i.e.

$$(Y^1, Y^0) \perp T_g | p(V) \quad (2-9)$$

The propensity score $p(V)$ can be estimated using any standard probability model. Many matching estimators based on propensity score have been developed for the estimation of the average effect of treatment (ATT), which is the most used evaluation parameter. A generalized form for the ATT may be written as follows:

$$\tau_{\text{ATT}} = \frac{1}{N_1} \sum_{i \in (Tg=1)} \left[Y_i^1 - \sum_{j \in (Tg=0)} w_{ij} Y_j^0 \right] \quad (2-10)$$

where w_{ij} is a weight that is function of the propensity score $p(V)$. This weight function gives higher weights to nonparticipants j with propensity score closer to that of the participant i and a lower weight to nonparticipants j with distant propensity score. Hence, for all i , $\sum_{j \in (Tg=0)} w_{ij} = 1$. In this setting, the treated are matched with a weighted average of

all controls with weights that are inversely proportional to the distance between the propensity scores of treated and controls.

The most known matching estimators are Nearest neighbor matching, Radius matching estimator (Cochran and Rubin (1973)), Kernel matching and Local linear regression matching estimator. These estimators differ from each other by the weights they attach to individuals of the comparison group. A good review of matching estimators can be found in Smith and Todd (2003).

These traditional cross-section matching estimators however have been criticized since they do not allow controlling for the potential bias that may arise from unobservable time trends which are common across treatment and comparison groups. Indeed, if the outcome variable change over time due to reasons unrelated to the participation decision, these estimators will be biased. Various alternate estimators have been proposed to eliminate this potential bias and one of the most compelling estimators is the difference-in-differences estimator (DiD) (see for example Heckman and Robb (1985)). The DiD is based in the simple idea of subtracting the before-after change in nonparticipant outcomes from the before-after change in the participant outcome. Hence, relying on the assumption of time-invariant linear selection, this double-difference permits to eliminate the bias from common time trends. Reformulating equation (2-10) accordingly, we get the DiD estimator :

$$\tau_{DiD} = \frac{1}{N} \sum_{i \in (Tg=1)} \left[Y_{i,t1}^1 - Y_{i,t0}^1 - \sum_{j \in (Tg=0)} w_{ij} (Y_{j,t1}^0 - Y_{j,t0}^0) \right] \quad (2-11)$$

The application of the simplest DiD estimator is based on the simple comparison of the outcomes of the treatment group and of the control group between pre-treatment period t0 and post-treatment period t1. Heckman, Ichimura, Smith and Todd (1998) introduced an estimator that combines this standard DiD estimator with any of the matching estimators based on propensity score mentioned above, namely the conditional semiparametric difference-in-differences estimator (CDiD). More specifically, the CDiD estimator is implemented using equation (2-11), but in this case the weights w_{ij} are based on the propensity score $p(V)$. This is the key difference and the advantage of the CDiD over the standard DID since, by using $p(V)$ in equation (2-11), the CDiD allows not only selection on both observables and on unobservables, but it also resolves the multidimensional problem mentioned above. Another advantage of the CDiD matching estimator is that, unlike traditional cross-section matching estimators, it combines in one step the matching process with the estimation of the effect of the treatment. Smith and Todd (2005) show that the CDiD estimator is more robust than standards DiD estimators. The CDiD may be based on any of the various propensity matching estimators previously cited.

Following Heckman, Ichimura, Smith and Todd (1998), we apply in this study the CDiD with weights based on the local linear regression matching estimator (LLR henceforth). These weights are defined by Heckman, Ichimura, Smith and Todd (1998) as:

$$w_{ij} = \frac{G_{ij} \sum_{k \in (Tg=0)} G_{ik} (p_k - p_i)^2 - [G_{ij} (p_j - p_i)] \left[\sum_{k \in (Tg=0)} G_{ik} (p_k - p_i) \right]}{\sum_{j \in (Tg=0)} G_{ij} \sum_{k \in (Tg=0)} G_{ij} (p_k - p_i)^2 - \left(\sum_{k \in (Tg=0)} G_{ik} (p_k - p_i) \right)^2} \quad (2-12)$$

where p is the propensity score, $G_{ij} = G((p_i - p_j)/a_N)$ is a kernel function and a_N is a bandwidth parameter⁵. Intuitively, w_{ij} uses all the observations in the control group and place higher weight on units close in terms of propensity score p and lower weights on more distant observations. Local linear weights are superior to any conventional kernel weights since local linear estimators converge at a faster rate at boundary points and adapt better to different data densities (Heckman, Ichimura and Todd (1997), Fan (1992))⁶.

To summarize, for the application of the CDiD estimator, we apply equation (2-11) using the weights defined by (2-12). For each treated firm we obtain at each period an estimate of the effect of the treatment $\tau_{i,t}$, now defined by $\tau_{DiD,i,t}$ and given by:

$$\tau_{i,t} = \tau_{DiD,i,t} = f(Y_{i,t}^1, Y_{i,t-1}^1, Y_{j,t}^0, Y_{j,t-1}^0, w_{ij}(p)) \quad (2-13)$$

2.2. Step2: The production function framework

To assess the effectiveness of the public direct support in the growth productivity of firms, we use a production function linking output to physical capital, to labour and to

⁵ The bandwidth parameter a_N is assumed to converge toward zero as n and na_N converge toward infinity.

⁶ It is worth noting that LLR matching estimator is a generalized version of the kernel matching estimator and both are special cases of local polynomial estimators (Cleveland, William S., 1979, Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association* 74, 859-836, Stone, Charles J., 1977, Consistent Nonparametric Regression, *Annals of Statistics* 5, 595-620.)

knowledge capital. The latter, which is our variable of interest, enters the production function in two separate components, that is on the one hand, the share of knowledge capital induced by R&D expenditures funded privately and by tax credits, and on the other hand, the share of knowledge capital induced by R&D expenditures funded by public grants. These two components are taken from the estimation of the average effect of treatment described in the previous section.

For the production function, let's consider a typical a Cobb-Douglas production function given by:

$$Q_{it} = Ae^{\lambda t} K_{it}^{\alpha} L_{it}^{\beta} S_{it}^{\gamma'} S_{it}^{\gamma''} e^{\varepsilon_{it}} \quad (2-14)$$

where i denotes firms and t years, Q is the output, K is physical capital, L is labour, S' is the composite share of knowledge capital induced by R&D expenditures privately funded and funded by tax credits, S'' is the share of knowledge capital induced by R&D expenditures funded by public grants, A is a constant, λ is a scale factor measuring the rate of disembodied technical change, α is the elasticity of output with respect to physical capital, β the elasticity of output with respect to labour, γ' the elasticity of output with respect to knowledge capital induced by private R&D and by tax credits, γ'' the elasticity of output with respect to knowledge capital induced by public grants, and finally ε is the error term that takes account of all the other factors such as measurement errors, firms technological heterogeneity and all other factors not accounted for explicitly in the inputs.

For the estimation of equation (2-14), since in the first step of the study we used the R&D intensity as a proxy for the knowledge capital, equation (2-14) must be transformed in such a way that it includes R&D intensity in its functional form instead of knowledge capital. This is fulfilled by adapting to our needs the derivation proposed by Griliches (1973) into a labour productivity equation as follows⁷:

⁷ The labour productivity equation is an extended version of the Cobb-Douglas production function based on the proposition of Griliches, Zvi, 1973, Research expenditures and growth accounting, in B.R. Williams, ed.: *Science and technology in economic growth* (Macmillan, London). See appendix E for the full derivation of this equation.

$$\Delta(q-l)_{it} = \lambda + \alpha\Delta(k-l)_{it} + \theta\Delta l_{it} + \rho' \left(\frac{R'}{Q}\right)_{it} + \rho'' \left(\frac{R''}{Q}\right)_{it} + v_{it} \quad (2-15)$$

where lower case letters denote the logarithms of variables. ρ' is the rate of return on R&D induced by private R&D and tax credits, and ρ'' is the rate of return induced by public grants. In the same manner, $\left(\frac{R'}{Q}\right)$ and $\left(\frac{R''}{Q}\right)$ are respectively the R&D intensity induced by private R&D and tax credits, and by public grants. $\theta = \alpha + \beta - 1$ is the measure of returns to scale and, if we assume constant returns to scale, this measure should equal zero. However, if this measure is less than or greater than zero, then decreasing, respectively increasing, returns to scale prevail.

3. DATA AND VARIABLES

3.1. DATA

The data consist of a longitudinal sample of 3821 manufacturing firms over the period 1997-2004 (11 884 observations) obtained by linking three micro data files. Such linking was necessary since only a part of the required information is available in each file. The first file is sourced from the Statistics Canada annual survey on Research and Development in Canadian Industry (RDCI)⁸ which is a census with cross sectional design that collects data on all Canadian firms known or believed to perform or fund R&D. In order to reduce the reporting burden on firms, only firms performing or funding more than \$1 million in R&D are surveyed. For all the other firms, the data are extracted from administrative data from Canada Revenue Agency (CRA). For the purpose of this study, the main information collected from this file is the information on public subsidies. Indeed, the RDCI survey is the only source of information that collects information on the various types of government direct incentives (grants, loans, procurements, venture capital investments or other instruments) from both provincial governments and the federal government including its agencies. The second file comes from the Survey of Manufactures (ASM)⁹, conducted annually by Statistics Canada since several decades and its sampling frame comprises all establishments primarily engaged in manufacturing activities, that is, all establishments classified to sectors 31, 32 and 33 under the North American Industry Classification System (NAICS). Data collected directly from firms' respondents by questionnaires and from administrative files from CRA permit to compile financial information such as employment, wages, total cost of materials, total sales of manufactured products, inventories, value added by manufacturing and capital expenditures. The third linked file comes from administrative data from the Ministry of Revenue of Quebec (Revenu Québec) that collects fiscal data from R&D performers claiming tax credits located in Quebec or have R&D conducted on their behalf in Quebec. Hence, the addition of this file to our dataset allowed us to gather records of the

⁸ Detailed description of the RDCI survey can be found in the Industrial Research and Development (2002 Intentions)

⁹ Starting from 2004, this survey has been amalgamated to the Annual Survey of Forestry to form the Annual Survey of Manufactures and Logging (ASML).

amounts of SR&ED or other fiscal incentives to R&D from the government of Québec effectively received by firms¹⁰.

For both, the RCDI data survey and the Revenu Québec's administrative data, the reporting unit is the company or enterprise. However, the reporting unit in the case of the ASM survey is the establishment. We grouped the establishments in the ASM data file by enterprise in the cases of multi-establishments prior to performing the linking of the three data file. After matching the three files, discarding cases of firms' merger and acquisition and cleaning outliers in the data, we obtained a sample of 3815 firms with 11 842 observations of which, 485 cases received both tax credits and R&D grants and, the remaining 11 357 cases received tax credits only. However, it should be noted that because of the methodological approach adopted, the matching sample used corresponds to the years 1998 to 2004 with 3749 firms (11012 observations, of which 410 are treated and 10602 are non-treated). Detailed description of this sample is presented in section 4.1.

3.2. Variables for the estimation of ATT (matching framework)

3.2.1. Treatment variable

The treatment variable, denoted T_g , is a dummy variable that takes the value of 1 if a firm received both R&D tax credits and R&D grants from either the Canadian government or from the government of Quebec. In the case where the firm received only R&D tax credits from either the Canadian government or from the government of Quebec, the treatment variable takes the value of 0. Thus, T_g is a binomial variable that captures the program participation status of a firm and is, for that, the dependant variable in the regression for the estimation of the propensity score.

¹⁰ Only the information related to SR&ED and other fiscal incentives to R&D from the government of Québec was available to us. However, by using official formulas and both, the RCDI data and the Revenu Québec's administrative data, we estimated the amounts of fiscal incentives to R&D received by firms from the federal government. This procedure is explained in more details in Baghana, Rufin, and Pierre Mohnen, 2009, Effectiveness of R&D tax incentives in small and large enterprises in Québec, *Small Business Economics* 33, 91-107.

3.2.2. Selection variables

We take the variables that influence both the probability, for a firm that receives tax credits, of getting an R&D grant (or R&D contract) and the investment in private R&D. These selection variables are taken at the beginning of the period prior to the decision to participate (or not), that is in our case, the first lag of these variables. As our dataset is rich enough, we included most variables often used to estimate this probability¹¹: Firm size is widely considered as one of the variables that affect positively this probability and the latter is likely higher for larger firms. Indeed, larger firms are less constrained by factors such as fixed cost barrier to R&D, the lack of time and human resources necessary to follow up technological activities. We use the lagged log of the number of employees ($(l)_{t-1}$) as proxy for the size of the firm. The intensity of capital ($(k/l)_{t-1}$), constructed as the lag of the logarithm of firm's real assets divided by the number of employees, is added to capture the influence of the capital structure in the willingness of firms to engage in R&D. The share of technical personnel working in R&D activities (engineers, scientists) is a good indicator of firms' involvement in innovation activities. We therefore added a variable constructed as the lagged value of the ratio of firm's R&D personnel to number of employees ($(L_T/L)_{t-1}$). Almus and Czarnitzki (2003) point out the importance of also including variables that control for market competition. For that, we use respectively, the market share ($(q/q_I)_{t-1}$), constructed as the lagged value of the ratio of firm's real sales to industry's real sales (at 3-digit NAICS classification) in logarithm and the export share ($(e/q_I)_{t-1}$) constructed as the lagged value of the ratio of firm's real exportations to industry's real sales in logarithm. We include in the regression a country of control dummy (FOREIGN) that equals 1 if the country of control of the firm is not Canada and 0 otherwise. It is often advanced in the literature that a firm with a foreign owned parent company will usually not apply for grants when its demand in R&D is adequately supplied by the latter. If this is true for most of the firms owned by a foreign

¹¹ We present only variables retained in the final estimation. Other variables that were not statistically significant were discarded in the final estimations. Among these variables are variables often used in the literature: age of firm, control variables for market competition such as the concentration ratio or the import ratio that captures the competitive pressure of foreign firms on the market (see example Almus, Matthias, and Dirk Czarnitzki, 2003, The Effects of Public R&D Subsidies on Firms' Innovation Activities: The Case of Eastern Germany, *Journal of Business and Economic Statistics* 21, 226-36.)

company, we should expect a negative sign in the coefficient of the dummy FOREIGN. We control for technological differences by including twenty dummies (indus1–20) to take account of the firm specific effects in the twenty one sectors of the NAICS at the 3 digit code level (see NAICS classification in Table D-5).

3.2.3. Outcome variable

Sometimes referred to as response variables, outcome variables are the outputs of models of economic choice such as the model described in section 2.1. The response variable of interest in this study is the intensity of R&D constructed as the ratio of firm's real R&D expenditures R to firm's output Q . In what follows, we denote it R/Q .

3.3. Variables entering the production function

The main variables entering the production function are those appearing explicitly in equation (2-15) described in section 2.2, i.e. the labor productivity growth rate ($\Delta(q/l)_t$), the capital-labor growth rate ($\Delta(k/l)_t$), the R&D intensity induced by private R&D and tax credits $(R'/Q)_t$, the R&D intensity induced by public grants $(R''/Q)_t$ and the employment growth rate (Δl_t).

In addition to these variables, we add to the regression other variables in order to control the effects of factors not explicitly accounted for. Six indicator variables denoted t1 to t6 are added to take account of time effects in the study period (1998-2004) and 20 group effect dummies as described in the previous section.

