Working Paper Series

#2010-055

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

Rufin Baghana

United Nations University – Maastricht Economic and social Research and training centre on Innovation and Technology
Keizer Karelplein 19, 6211 TC Maastricht, The Netherlands
Tel: (31) (43) 388 4400, Fax: (31) (43) 388 4499, email: info@merit.unu.edu, URL: http://www.merit.unu.edu

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-indifferences 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.

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 point to the underinvestment from private firms in R&D leading to a supply offer 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 latter[1]. 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 Microeconometric 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 (MDEIE[2]) is the main provincial institution for public direct funding to firms[3]. 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 level[4], 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 indictor of exposure to treatment that takes the value 1 if the firm is treated and 0 otherwise. We can write the following:

$$Y_i^0 = h^0(X_i) + \varepsilon_i, \qquad if \quad T_{gi} = 0 \qquad (i = 1,...,n)$$
$$Y_i^1 = h^1(X_i) + \tau_i + \varepsilon_i, \qquad if \quad T_{gi} = 1 \qquad (2\text{-}1)$$

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:

$$Y_i^1 - Y_i^0 = \tau_i \qquad (2\text{-}2)$$

In the second step, a productivity function Q is estimated using the following regressors: $\mathrm{Y}_\mathrm{i}^0 = \mathrm{Y}_\mathrm{i}^1 - \tau_\mathrm{i}$, 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:

$$Q_i = f(Y_i^0, \tau_i, Z_i) \qquad (2\text{-}3)$$

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, $\mathrm{Y} = \mathrm{T}_\mathrm{gi}\mathrm{Y}_\mathrm{i}^1 + (1 - \mathrm{T}_\mathrm{gi})\mathrm{Y}_\mathrm{i}^0$. 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:

$$\mathrm{ATT} = E(Y^1 | T_g = 1, X) - E(Y^0 | T_g = 1, X) \qquad (\text{ 2-4 })$$

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:

$$(Y^1, Y^0) \perp T_g V \qquad (2\text{-}5)$$

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:

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

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

$$(T_g \perp V | p(V) \qquad (2\text{-}7)$$

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

$$(Y^1, Y^0) \perp T_g V \qquad (2\text{-}8)$$

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

$$(Y^1, Y^0) \perp T_g | p(V) \qquad (2\text{-}9)$$

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:

$$\tau_{\mathrm{ATT}} = \frac{1}{\mathrm{N}_1} \sum_{\mathrm{i} \in (\mathrm{Tg}=1)} \left[\mathrm{Y}_\mathrm{i}^1 - \sum_{\mathrm{j} \in (\mathrm{Tg}=0)} \mathrm{w}_{\mathrm{ij}} \mathrm{Y}_\mathrm{j}^0 \right] \qquad (2\text{-}10)$$

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, $\sum_{j \in (Tg=0)} w_{ij=1} = 1$. 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 :

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

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:

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

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 parameter[5]. 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:

$$\tau_{i,t} = \tau_{DiD,i,t} = f\left(Y^1_{i,t}, Y^1_{i,t-1}, Y^0_{j,t}, Y^0_{j,t-1}, w_{ij}(p)\right) \qquad (2\text{-}13)$$

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:

$$Q_{it} = Ae^{\lambda_t} K_{it}^{\alpha} L_{it}^{\beta} S'^{\gamma'}_{it} S''^{\gamma''}_{it} e^{\varepsilon_{it}} \qquad (2\text{-}14)$$

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 follows[7]:

[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.

$$\Delta(q-l)_{it} = \lambda + \alpha\Delta(k-l)_{it} + \theta\Delta l_{it} + \rho'\left(\tfrac{R'}{Q}\right)_{it} + \rho''\left(\tfrac{R''}{Q}\right)_{it} + \upsilon_{it} \qquad (2\text{-}15)$$

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, $\left(\tfrac{R'}{Q}\right)$ and $\left(\tfrac{R''}{Q}\right)$ 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 firms[10].

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 probability[11]: 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 ($(L_r/L)_{t-1}$). 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 ($\text{indus}1-20$) 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.

4. EMPIRICAL ANALYSIS AND RESULTS

4.1. Assessing the input additionality of R&D grants

To assess the input additionality of R&D subsidies in Quebec, we investigated whether public subsidies induce firms in Quebec to spend more on R&D than they would without the grants. To do so, we estimated the treatment effect of public grants on firms R&D level using the CDiD matching estimator described in section 2.1. Since the choice is given to firms by the provincial and federal governments to apply each year for grants in addition to the tax credits they are entitled to claim, the treatment effect of public grants should also be evaluated each year. Hence, the CDiD matching estimator is applied at each period of the matching sample (i.e. from t=1998 to t=2004) and, for the group of treated firms, a treatment effect $\tau_{(DiD,t)}$ is obtained for each treated period. The treatment effect is not evaluated for the first year of the original sample, i.e. 1997, due to the presence of lagged variables in the estimation of the propensity score and also because, for each evaluation point, the difference-in-differences estimator requires data for at least two periods.

As another goal of this study is to see how firms of different technological level respond to public grants, we also performed matching analysis using two sub-samples of technology level breakdown (Low-medium technology and High technology). **Table 4-1** shows the description of the samples used in the matching for the full sample and for the two technologies sub-samples.

Table 4-1: Description of the samples used in the matching

Sample	All firms		Low and medium technology		High technology	
	N_1	N_0	N_1	N_0	N_1	N_0
1998	58	819	38	726	20	93
1999	56	886	41	770	15	116
2000	76	1565	60	1385	16	180
2001	56	1542	41	1352	15	190
2002	58	1925	48	1728	10	197
2003	54	2032	40	1825	14	207

2004	52	1833	39	1654	13	179

Note: N_1=Number of treated firms (firms with tax credits + R&D grants)
N_0=Number of non-treated firms (firms with tax credits only)

To obtain the propensity score, for each of the seven periods, we estimated and compared several specifications based on probit and on logit models with different set of conditioning variables. We ended-up choosing a specification based on a probit model with the following conditioning variables: the lagged log number of employees, the lagged log capital intensity, the lagged share of R&D personnel, the lagged log market share, the lagged log export share and the degree of foreign ownership. This specification was clearly superior to the logit model with the same set of covariates since it showed better statistical goodness of fit and lower value for the well-known Akaike's Information Criterion for selecting among nested econometric models.

Table 4-2 shows the results of probit models for all firms estimated for each year from 1998 to 2004 with the corresponding sub-sample. The results for the low-medium technology firms and for the high technology firms are in Table 4-3 and Table 4-4 respectively. A look at the three tables reveals that all the probit regression models are statistically significant (All firms: LR $Chi2_{1998\text{-}2004}$ = [21.35 38.81], P = [0.000 0.002]; Low-Medium technology firms: LR $Chi2_{1998\text{-}2004}$ = [13.86 30.69], P = [0.000 0.031]; High technology firms: LR $Chi2_{1998\text{-}2004}$ = [13.95 26.74], P = [0.000 0.030]). On the other hand, all the predictor variables in the models in the three samples are statistically significant at the 10% level and, across the three samples; they have the same sign at each point of time. More specifically, this indicates that the probability to get a subsidy increases with the number of employees, the capital intensity and the share of R&D personnel and decreases with the market share, the export share and the degree of foreign ownership.

Table 4-2: Probit estimates for the propensity score estimation (All firms, 1998-2004)

Parameters/statistics	Probit						
	(1998)	**(1999)**	**(2000)**	**(2001)**	**(2002)**	**(2003)**	**(2004)**
$(l)_{t-1}$	0.351***	0.196***	0.289***	0.376***	0.346***	0.430***	0.322***
	(0.115)	(0.063)	(0.086)	(0.090)	(0.093)	(0.088)	(0.097)
$(k/l)_{t-1}$	0.167**	0.0805**	0.0910**	0.145*	0.252**	0.154*	0.187*
	(0.071)	(0.040)	(0.046)	(0.077)	(0.121)	(0.092)	(0.107)
$(L_r/L)_{t-1}$	1.221***	1.406***	1.655***	0.386***	0.266*	1.028***	0.951***
	(0.470)	(0.459)	(0.533)	(0.143)	(0.144)	(0.325)	(0.348)
$(q/q_I)_{t-1}$	-0.243**	-0.110*	-0.252***	-0.195***	-0.165**	-0.244***	-0.175**
	(0.098)	(0.057)	(0.068)	(0.064)	(0.074)	(0.077)	(0.084)
$(e/q_I)_{t-1}$	-0.0464**	-0.173**	-0.0571*	-0.178**	-0.0824*	-0.0336*	-0.240*
	(0.023)	(0.085)	(0.033)	(0.074)	(0.045)	(0.018)	(0.132)
FOREIGN	-0.619**	-0.632**	-0.789***	-0.340**	-0.313*	-0.474***	-0.263*
	(0.286)	(0.276)	(0.284)	(0.150)	(0.165)	(0.150)	(0.150)
Industry dummies ($indus1 - indus20$)	included	included	included	included	included	included	included
Intercept	-5.400***	-6.047***	-6.223***	-5.977***	-6.953***	-5.668***	-6.406***
	(1.522)	(1.248)	(1.357)	(1.527)	(2.003)	(1.851)	(1.936)
Pseudo R^2	0.092	0.080	0.081	0.083	0.073	0.090	0.074
Log likelihood	-141.21	-133.89	-153.67	-185.34	-162.88	-176.74	-172.71
Number of observations	647	662	849	1140	1315	1473	1447

Notes: *** indicates significance at 1%, ** at 5%, * at 10%. Standard errors in parentheses unless otherwise indicated. Lower case symbol denote the logarithm of the corresponding variable. l :log of number of employees, k/l : capital intensity, L_r/L : share of R&D personnel, q/q_I : market share, e/q_I : export share, FOREIGN: origin country dummy (Canada=0, Foreign=1), $indus1 - indus20$: industry dummies.

Table 4-3: Probit estimates for the propensity score estimation (Low-Medium technology firms, 1998-2004)

Parameters/statistics	**Probit**						
	(1998)	**(1999)**	**(2000)**	**(2001)**	**(2002)**	**(2003)**	**(2004)**
$(l)_{t-1}$	0.307** (0.138)	0.164** (0.070)	0.203** (0.084)	0.304*** (0.108)	0.283*** (0.093)	0.388*** (0.104)	0.177** (0.089)
$(k/l)_{t-1}$	0.183* (0.097)	0.107** (0.051)	0.0969* (0.055)	0.220* (0.126)	0.143* (0.077)	0.184* (0.107)	0.195* (0.100)
$(\mathrm{L_r/L})_{t-1}$	2.193* (1.202)	1.263** (0.591)	1.042* (0.596)	0.346* (0.206)	0.238* (0.143)	0.783** (0.377)	0.999*** (0.379)
$(q/q_I)_{t-1}$	-0.274** (0.112)	-0.132** (0.066)	-0.255*** (0.067)	-0.169*** (0.065)	-0.136** (0.066)	-0.262*** (0.091)	-0.155** (0.075)
$(e/q_I)_{t-1}$	-0.131* (0.079)	-0.0745* (0.044)	-0.0623* (0.036)	-0.148* (0.080)	-0.120*** (0.038)	-0.0373** (0.018)	-0.389** (0.152)
FOREIGN	-0.625* (0.372)	-0.402* (0.242)	-0.638* (0.360)	-0.345** (0.176)	-0.325* (0.191)	-0.337* (0.187)	-0.432* (0.242)
Industry dummies (indus 1 – indus 20)	included	included	included	included	included	included	included
Intercept	-6.260*** (2.046)	-5.907*** (1.129)	-5.691*** (1.485)	-6.211*** (1.691)	-2.629** (1.300)	-5.949*** (2.144)	-6.017*** (1.370)
Pseudo R^2	0.063	0.067	0.073	0.065	0.065	0.078	0.089
Log likelihood	-92.31	-98.31	-121.94	-144.35	-140.26	-138.69	-129.76
Number of observations	548	564	734	981	1159	1308	1285

Notes: *** indicates significance at 1%, ** at 5%, * at 10%. Standard errors in parentheses unless otherwise indicated. Lower case symbol denote the logarithm of the corresponding variable. l :log of number of employees, k/l : capital intensity, $\mathrm{L_r/L}$: share of R&D personnel, q/q_I : market share, e/q_I : export share, FOREIGN: origin country dummy (Canada=0, Foreign=1), indus 1 – indus 20 : industry dummies.

Table 4-4: Probit estimates for the propensity score estimation (High technology, 1998-2004)

Parameters/statistics	Probit						
	(1998)	(1999)	(2000)	(2001)	(2002)	(2003)	(2004)
$(l)_{t-1}$	0.430*** (0.133)	0.308*** (0.116)	1.231** (0.500)	0.486** (0.210)	0.720*** (0.271)	0.801*** (0.311)	0.425** (0.173)
$(k/l)_{t-1}$	0.479** (0.208)	0.243** (0.107)	0.142* (0.084)	0.196** (0.086)	0.952*** (0.360)	0.623* (0.339)	0.342* (0.185)
$(L_r/L)_{t-1}$	2.158** (0.941)	2.512* (1.486)	3.407** (1.573)	1.483** (0.682)	1.317*** (0.411)	1.160* (0.618)	1.084* (0.630)
$(q/q_I)_{t-1}$	-0.110** (0.056)	-0.265* (0.155)	-0.697* (0.378)	-0.107* (0.065)	-0.441** (0.225)	-0.596** (0.253)	-0.286** (0.113)
$(e/q_I)_{t-1}$	-0.106** (0.048)	-0.754** (0.380)	-0.730* (0.443)	-0.249* (0.147)	-0.194* (0.118)	-0.127** (0.051)	-0.453* (0.234)
FOREIGN	-0.674* (0.357)	-0.781* (0.408)	-0.837** (0.387)	-1.382** (0.659)	-1.332** (0.658)	-1.411*** (0.361)	-0.786** (0.358)
Industry dummies (indus 1 – indus 20)	included	included	included	included	included	included	included
Intercept	-9.651*** (3.327)	-10.95*** (3.434)	-13.71*** (5.083)	-8.474*** (2.360)	-19.61*** (6.431)	-15.06** (6.251)	-9.946*** (3.046)
Pseudo R^2	0.233	0.267	0.273	0.233	0.355	0.239	0.183
Log likelihood	-36.00	-26.72	-24.69	-34.51	-16.40	-30.74	-32.84
Number of observations	99	98	115	159	156	165	162

Notes: *** indicates significance at 1%, ** at 5%, * at 10%. Standard errors in parentheses unless otherwise indicated. Lower case symbol denote the logarithm of the corresponding variable. l :log of number of employees, k/l : capital intensity, L_r/L : share of R&D personnel, q/q_I : market share, e/q_I : export share, OWN : origin country dummy (Canada=0, Foreign=1), indus 1 – indus 20 : industry dummies.

In order to check for the success of the matching in removing differences between the group of treated firms (firms with tax credits and grants) and the group of non treated firms (firms with tax credits only), one need to check for the balance in the distribution of covariates between the two groups before and after performing matching. Matching is successful if no significant differences between the means of the two groups remain after its application. We first performed the widely used before-after mean comparison t-test in the evaluation literature. It consists in carrying out a t-test comparing the means of the covariates of the two groups before and after the matching to test whether the means are different.The results of this test before the matching and after the matching for the sample of all firms are presented in Table B-2 and Table B-3 respectively. Before the matching, for all covariates but the share of R&D personnel, there is a significant difference between the two groups of means. After the matching, the differences in the means of the two groups in all covariates and in the propensity score are no longer significant. However, there are still differences in the outcome variable, R&D intensity, as the coefficients in all periods are significant at the 5% level. These differences may be attributed to the receipt of public R&D grants.

Then we computed, as suggested by Rosenbaum and Rubin (1985), the standardized difference statistic for the covariates of the two groups before and after the matching (results of test for all firms before and after matching in Table B-2 and Table B-3 respectively). This statistic is the difference of the means of the two groups as a percentage of the square root of the average of the sample variances of the two groups[12]. The results after the matching, show that standardized differences for all covariates, excluding the outcome variable, lie below the 10% threshold, that is, we therefore consider that the differences of the means of the two groups for all covariates are no longer significant. Indeed, though Rosenbaum and Rubin (1985) did not set a rule for the statistical significance of the standardized difference statistic, they considered absolute

[12] Formally, the standardized difference statistic is a percent given by $100(\bar{x}_1 - \bar{x}_{0M})/\left[(s_1^2 + s_{0R}^2)/2\right]^{1/2}$, where, for each covariate, $\bar{x}_1$ and $\bar{x}_{0M}$ are the sample means in the treated group and matched control group and s_1^2 and s_{0R}^2 are the sample variances in the treated group and control reservoir (Rosenbaum, Paul R., and Donald B. Rubin, 1985, Constructing a control group using multivariate matched sampling methods that incorporate the propensity score, *The American statistician* 39, 33-38.)

standardized differences below 10% as inconsequential. These results (for both the before-after mean comparison t-test and the standardized difference statistic) give us a great confidence in the quality of the matching. As an additional balancing test, we graphically compared the propensity scores of treated and non-treated groups before and after matching. As can be seen in Figure C-1, the results confirm the success of the LLR matching estimator in removing differences between both groups.

We performed matching using PSMATCH2 Stata module by Leuven and Sianesi (2003). For each of the seven periods, we allowed matching only among treatment and control observations pertaining to the same industrial sector using NAICS 3 digit code. Given that the seven subsamples for each period are constructed from an unbalanced panel, certain firms are observed in more than one cross-section and, in several cases these firms switch from the status of treated (Tax credits + R&D grants) to the non-treated status (Tax credits only) from a given year to another and vice-versa. If we assume that the effect of treatment on a firm last more than one year, then a non-treated firm in a given year may not be a reliable match if it had the status of a treated firm in the previous year. For this reason and in order to ensure group homogeneity, we imposed an additional restriction on each comparison group constructed in the matching for the seven subsamples by limiting it only to firms that maintained a status of a non-treated firm in any year it is observed along the study period. Additionally, we imposed the common support region which consists in dropping treatment observations whose propensity score is higher than the maximum or less than the minimum propensity score of the observations in the control group. As a consequence, after the matching, the remaining observations in the three samples used are as follows: All firms, 7117 observations, 401 treated and 6716 non-treated; Low-medium technology sample, 6147 observations, 299 treated and 5848 non-treated; High technology, 876 observations, 99 treated and 876 non-treated. The weighting function used for the local linear regression matching is the triweight or tricube kernel[13] and the bandwidth is 0.08[14].

[13] We also experimented with biweight (quartic) kernel, gaussian kernel and uniform kernel and it did not change significantly the results. A good review of weightings functions can be found in Li, Qi, and Jeffrey S. Racine, 2007. *Nonparametric Econometrics: Theory and Practice* (Princeton University Press).

Before turning to the findings from the application of the CDiD estimator, we should note that we also performed the CDiD using the Kernel matching estimator, but there were no much difference in the estimates reported. This must be caused by the fact that in our data there is substantial overlap in the distribution of the propensity score between the comparison and treatment groups. Dehejia and Wahba (2002) explain that in such case, most of the matching algorithms will yield similar results[15].

The findings from the CDiD estimator for all firms and by technology level are reported in Table 4-5. Since weighted average outcomes for the non-treated firms enter the LLR matching estimator equation (see section 2.1.2), the standard errors reported in the matching process for the CDiD coefficients are invalid. To deal with this issue, we bootstrapped these standard errors for all the CDiD estimates with 1000 replications. In sum, the results show an additionality effect of R&D grants on firms' private investment on R&D. More precisely, the average effect of treatment on R&D intensity for the full sample varies between 2% and 11% from 1998 to 2004. For the sub-sample of low-medium technology firms, which contains more than 86% of the observations of the full sample, the results are also quite similar. However, for the sub-sample of high technology firms, the CDiD estimates are a bit higher than the ones in the two previous cases, i.e. these estimates vary between 4% and 13%. These results suggest the exclusion of possible crowding-out effects in a context where firms are given the choice of adding R&D grants to R&D tax credits. Furthermore, the results in the case of high technology firms suggest that in a fiscal environment such as that of Canada, high technology firms may be more responsive than other firms to direct public subsidies. Finally, as can be seen in these results, the magnitude of the response decreases between 1998 and 2004 in the three cases. Since the differences in the number of subsidized firms are not very

[14] We also tested sensitivity of findings to various specifications on bandwidth but did not find significant differences.

[15] We find that the results from the LLR matching estimator exhibit significantly less bias error than the kernel-based matching estimator. Even though, differences in the estimates of the both matching estimators are not very important, the overall comparison shows that LLR matching is the best estimator for our data since it is more successful in removing differences between the group of firms with tax credits and grants and the group of firms with tax credits only.

important in the three samples, we attribute such a decline to the diminution in the value of grants observed in Quebec between 1998 and 2004 (see **Table 1-1**).

Table 4-5: Estimated effect of treatment on R&D intensity based on conditional difference-in-differences and local regression matching (bandwidth: 0.08)

	All firms	Low and medium technology	High technology
1998	N_1=55, N_0=410	N_1=37 N_0=375	N_1=20, N_0=90
	0.1122***	0.1022***	0.1269***
	(0.0002)	(0.0008)	(0.0021)
1999	N_1=56, N_0=477	N_1=40 N_0=413	N_1=14, N_0=98
	0.1108***	0.1001***	0.1248***
	(0.0002)	(0.0007)	(0.0021)
2000	N_1=73, N_0=1063	N_1=58 N_0=933	N_1=15, N_0=112
	0.0996***	0.0979***	0.1067***
	(0.0005)	(0.0008)	(0.0037)
2001	N_1=55, N_0=1014	N_1=40, N_0=899	N_1=15, N_0=120
	0.0503***	0.0407***	0.0528***
	(0.0001)	(0.0003)	(0.0031)
2002	N_1=58, N_0=1270	N_1=46, N_0=1033	N_1=10, N_0=118
	0.0500***	0.0399***	0.0511***
	(0.0002)	(0.0005)	(0.0036)
2003	N_1=53, N_0=1302	N_1=40,N_0=1116	N_1=12, N_0=122
	0.0243	0.0201	0.0389***
	(0.0013)	(0.0020)	(0.0026)
2004	N_1=51, N_0=1180	N_1=38 N_0=1079	N_1=13, N_0=117
	0.0245	0.0233	0.0392***
	(0.0010)	(0.0017)	(0.0026)

Note: *** indicates significance at 1%, ** at 5%, * at 10%. Bootstrap standard errors in parentheses with 1000 replications.

4.2. Assessing the impact of privately and publicly funded R&D on productivity growth

We estimated equation (2-15) described in section 2.2. For the two components of R&D intensity that enter that equation i.e., the R&D intensity induced by R&D expenditures funded by R&D grants (R''/Q) and the R&D intensity induced by private R&D

expenditures and by tax credits (R'/Q), we take the two variables estimated in the previous section on matching, respectively $R''/Q = \tau_{DiD}$ and $R'/Q = Y^0 = Y^1 - \tau_{DiD}$. However, these variables should not enter the equation as contemporary variables since the effect of a shock on R&D intensity is likely to take a certain time before it is fully reflected in firm's productivity growth. It is often argued that a lag structure of several periods should be included in order to account for the adjustment of productivity growth. According to the Bureau of Labor Statistics (1989) survey of the literature relating R&D to productivity growth, the productivity gains should lag the R&D outlays by at least one year. Hence, we therefore included one-period lagged values of R''/Q and R'/Q in the regression due to relatively small number of observations in the subsample of treated firms.

We used value added as proxy for firm's output Q and we corrected our data to prevent bias in the estimates that can arise from double counting as warned by Hall and Mairesse (1995). To do so, we respectively subtracted R&D capital[16] from the assets and R&D personnel from the number of employees in the production function and, added the materials component of R&D expenditure into value added. Value added, the capital stock and R&D expenditures have been deflated respectively by the industry output price deflator at the 3-digit NAICS code, the machinery and equipment price indexes and by the Jaffe-Griliches R&D deflator (see Bureau of Labor Statistics (1989) for the latter).

For the estimation procedure, we first performed OLS regression of equation (2-15) with and without time and industry dummies. We also performed the Newey–West two-step efficient GMM estimation of this equation to allow for the fact that R&D intensity components (R''/Q and R'/Q) may be endogenous, i.e. the fact that R&D intensity and productivity may be mutually dependent. Indeed, when this hypothesis is verified, OLS estimates are biased because the assumption requiring that explanatory variables be independent is no longer valid. To correct for this problem, several statistical techniques

[16] For details about the construction of this variable, see 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.

such as indirect least squares, instrumental variable and two- and three-stage least squares may be applied. Hence, to verify whether R&D intensity components are endogenous, we ran the C or “difference–in–Sargan” test and for the two components, the test rejected the null that they are not endogenous regressors in the equations for the two-step GMM estimator. This confirmed the need to instrument both R''/Q and R'/Q. To this end, we experimented with several sets of variables and we ended-up using the following set: the second and third lagged values of R&D intensity, the lagged value of the share of R&D personnel and lagged value of the sum of provincial and federal grants. Table 4-6 shows the estimation results of equation (2-15) estimated by OLS and GMM respectively. It should be noted that, for comparison purposes, since the estimation by GMM is Heteroscedasticity-consistent (HC), we estimated the OLS using the Huber–White sandwich estimator of variance, which is Heteroscedasticity-consistent, in place of the traditional calculation. For each case we ran the estimation with and without the dummies for time and for industry. For this last part of the analysis, only the larger sample of all firms is considered. We excluded the sub-samples of low-medium and high technology firms due to the important reduction in the number of observation in the group of treated firms after differencing.

Table 4-6: Estimation results of OLS-HC and GMM models for the productivity growth (All firms)

Parameters/statistics	OLS-HC		GMM	
	(1)	(2)	(3)	(4)
$\Delta(q/l)_t$				
$(R''/Q)_{t-1}$	-0.154*** (0.038)	-0.156*** (0.038)	0.367*** (0.097)	0.134* (0.075)
$(R'/Q)_{t-1}$	-0.0387* (0.023)	-0.0406* (0.023)	0.426*** (0.108)	0.322*** (0.048)
$\Delta(k/l)_t$	0.112*** (0.014)	0.110*** (0.014)	0.0385** (0.018)	0.0971*** (0.013)
Δl_t	-0.198*** (0.019)	-0.200*** (0.019)	-0.0616*** (0.022)	-0.190*** (0.014)
Year dummies $(t_1 - t_6)$	Not included	Included	Included	Not included
Industry dummies (indus 1 – indus 20)	Not included	Included	Included	Not included

Intercept	0.181 (0.145)	0.263*** (0.031)	0.0667 (0.068)	0.221*** (0.032)
F test (p-value)	6.61 0.000	34.49 0.000	2.16 0.001	124.32 0.000
Log likelihood	-8315.73	-8337.55	-8432.84	-7251.93
AIC	16687.45	16685.10	16911.68	14513.86
Test of underidentification: Kleibergen-Paap rank LM statistic (p-value)			23.92 0.000	24.46 0.000
Test of weak identification: Cragg-Donald F statistic (at 5% level of confidence)			5.98	6.33
— IV relative bias to OLS			Rejected >30%	Rejected >20%
— Size bias			Not rejected >25%	Rejected >25%
Test of overidentification: Hansen J-statistic (p-value)			3.99 0.136	4.10 0.129
Number of observations	7406	7406	6071	6316

Notes: Estimation period is 1998-2004. *** indicates significance at 1%, ** at 5%, * at 10%. Standard errors in parentheses unless otherwise indicated. Lower case symbol denote the logarithm of the corresponding variable. $\Delta(q / l)$: labor productivity growth rate; R' / Q : R&D intensity (induced by private R&D expenditures and by tax credits); R'' / Q : R&D intensity induced by R&D expenditures funded by public grants; $\Delta(k / l)$: capital-labor growth rate; Δl : employment growth rate.

Comparing the estimated models, GMM estimated model 4 is clearly superior and its relative lower AIC value (14513.86) among the estimated models adds to this evidence. Furthermore, model 4 is the one having the best fit since all the parameters of interest (i.e. the coefficient in the two R&D intensity components, in capital labor growth rate and in employment growth rate) are significant at the 10% level of confidence. We therefore selected model 4 as our preferred model.

We ran several additional tests in order to check for the robustness of the specification of the selected model. First, as time and industry dummy variables were individually insignificant based on t-tests with very high p values, we proceeded with testing for the significance of a subset of coefficients using the Wald test. We could not reject the null that industry dummies, in the one hand, and time dummies, in the other hand, are jointly insignificant. In consequence, these dummies were excluded from model 4

To check the relevance of our excluded instruments, we performed the Kleibergen and Paap (2006) rank LM statistic test of underidentification which is robust to heteroskedasticity, autocorrelation or clustering. The reported statistic of 24.46 falls within the region of rejection as the p-value (0.000) strongly reject the null that the equation is underidentified at the 1% significance level. We further proceeded with testing for the presence of weak instruments since it is well known that even if the equation is identified, there may still be weak instrument problems often responsible of poor performance of estimators. We further proceeded with testing for the presence of weak instruments using the Cragg-Donald F statistic and critical values tabulated by Stock and Yogo (2005). At the 5% significance level, the test rejected the null hypothesis that the instruments are weak in both alternative definitions offered by Stock and Yogo (2005) i.e. first, the bias of the IV estimator, relative to the bias of the OLS, could exceed the threshold of 30% and, second, the bias in the size of the Wald test based in the IV statistics could exceed 25%. Finally, with a statistic value of 6.33, we fail to reject the Hansen J overidentification test, giving us greater confidence that our instrument set is relevant.

Regarding the significance of the coefficients, as can be seen, all the coefficients in our preferred econometric model (model 4) are significant at the 1% level with an exception for the coefficient related to the variable of R&D expenditures funded by public grants which is significant at the 10% level. Furthermore, except for the coefficient of employment growth that have a negative sign, all the coefficients have a positive sign, suggesting a positive effect on labor productivity growth rate. The result for the estimated coefficient for the capital-labor growth rate shows a statistically significant role and amounts to 0.097, which is quite close to the results reported in the literature. The estimated coefficient for the employment growth rate exhibits a negative sign and the t-test strongly rejects the null hypothesis that it is equal to zero. As pointed in section 2.2, this means that constant return to scale is rejected and that firms are facing decreasing return to scale in Quebec. Similar results have been reported by Griliches and Mairesse (1991) for US and Japan, Wakelin (2001) for UK, Smith, Dilling-Hansen, Eriksson and Madsen (2004) for Denmark and Maté-Garcia and Rodriguez-Fernandez (2008) for Spain.

The rates of return to R&D expenditures funded by public grants and to R&D expenditures privately funded and by tax credits are respectively 0.134 and 0.322. This means that for each additional dollar of public R&D grant, output increases by 0.134 dollars in the former case and that, for each additional dollar of the remaining R&D expenditures (funded privately and by tax credits), output increases by 0.322 dollars. It is noteworthy that, although these positive results support direct subsidies programmes, they also show 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. From a comparative point of view with the empirical literature, the estimated rates of return of 0.134 and 0.322 are consistent with most studies since they fall within the 0.2-0.5 interval of the values estimated by these studies. We should note however that this comparison must be taken with caution since these studies don't break R&D spending in various components as in the case of this study. The ratio of both rates of return is of 0.42 which indicates that public R&D grants represent 42% of the productivity induced by both private funds and R&D tax credits.

5. CONCLUSION

Investigating the effectiveness of public grants as a significant driver in firms' private spending for R&D and the impact of the latter on productivity, have become important policy issues in a context of tax credits. Using longitudinal data through the period 1997 to 2004 from manufacturing firms in Quebec, we analyse these issues in two steps: In the first step, we assess the input additionality of public R&D grants in terms of increased R&D spending from firms that already receive tax credits and, in the second step, in terms of productivity growth. As far we know, this is the first study that integrates these issues in one analysis framework by using micro data from Quebec.

Our results 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 and in terms of growth productivity than firms that use only tax credits for R&D. In particular, we find in the first step of the study that the R&D intensity of firms in the former group is higher than that of firms in the latter group. Our results also support the hypothesis that high-tech firms benefit the most from public grants financing. Indeed, we find that the magnitude of the impact of R&D grants in the intensity of R&D of high-tech firms is greater than that of low and medium technology firms.

In the second step of the study, the first main result is that the rate of return to R&D expenditures funded by public R&D grants is of 0.134 i.e. for each additional dollar of public R&D grant, output increase by 0.134 dollars. This finding permits to 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. The second main result is that public R&D grants represent 42% of the productivity induced by both private funds and R&D tax credits. This ratio provides support to the use of public R&D grants as additional policy instrument to tax credits. In other words, since R&D grants and R&D tax credits work well together, the question for policy makers is less whether to choose between R&D grants and R&D tax credits. The question is rather the capacity of the public agency responsible of the attribution of R&D grants to identify the projects for which R&D tax credits fail to offer an incentive for the private projects and to find a suitable level of additional funding in a context where firms already benefit from fiscal incentives.

REFERENCES

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.

Arrow, Kenneth, 1962, Economic Welfare and the Allocation of Resources for Invention, in *The Rate and Direction of Inventive Activity* (Princeton University Press).

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).

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.

Bernstein, I. Jeffrey, 1986, The Effect of Direct and Indirect Tax Incentives on Canadian Industrial R&D Expenditures, *Canadian Public Policy* 12, 438-448.

Bérubé, Charles, and Pierre Mohnen, 2009, Are firms that receive RD subsidies more innovative, *Canadian Journal of Economics* 42, 206-225.

Bureau of Labor Statistics, 1989, The impact of research and development on productivity growth, BLS Bulletin 2331 (Bureau of Labor Statistics, Washington, DC).

Cleveland, William S., 1979, Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association* 74, 859-836.

Cochran, William G., and Donald B. Rubin, 1973, Controlling Bias in Observational Studies: A Review *Sankhyā: The Indian Journal of Statistics* 35, 417-446.

Czarnitzki, Dirk, and Georg Licht, 2006, Additionality of Public R&D Grants in a Transition Economy: The Case of Eastern Germany, *Economics of Transition* 14, 101-31.

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.

David, Paul A., Bronwyn H. Hall, and Andrew A. Toole, 2000, Is public R&D a complement or substitute for private R&D? A review of the econometric evidence, *Research Policy* 29, 497-529.

Dehejia, Rajeev H. , and Sadek Wahba, 2002, Propensity Score-Matching Methods For Nonexperimental Causal Studies, *The Review of Economics and Statistics* 84, 151-161.

Fan, Jianqing, 1992, Design Adaptive Nonparametric Regression, *Journal of the American Statistical Association* 87, 998-1004.

Griliches, Zvi, 1973, Research expenditures and growth accounting, in B.R. Williams, ed.: *Science and technology in economic growth* (Macmillan, London).

Griliches, Zvi , and Jacques Mairesse, 1991, R&D and Productivity Growth: Comparing Japanese and U.S. Manufacturing Firms, (National Bureau of Economic Research, Inc).

Guellec, Dominique, and Bruno Van Pottelsberghe, 2000, The Impact of Public R&D Expenditure on Business R&D, in OECD, ed.: OECD Science, Technology and Industry Working Papers (OECD, Directorate for Science, Technology and Industry, Paris).

Hall, Bronwyn H., and Jacques Mairesse, 1995, Exploring the relationship between R&D and productivity in French manufacturing firms, *Journal of Econometrics* 65, 263-293.

Hanel, Petr , Dirk Czarnitzki, and Julio Miguel Rosa, 2005, Evaluating the Impact of R&D Tax Credits on Innovation: A Microeconometric Study on Canadian Firms, in ZEW discussion paper 04-77, ed.: ZEW (Centre for European Economic Research (ZEW), Mannheim).

Heckman, James, Hidehiko Ichimura, Jeffrey Smith, and Petra Todd, 1998, Characterizing Selection Bias Using Experimental Data.

Heckman, James J, Hidehiko Ichimura, and Petra E Todd, 1997, Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme, *Review of Economic Studies* 64, 605-54.

Heckman, James J., and Richard Robb, Jr., 1985, Alternative methods for evaluating the impact of interventions : An overview, *Journal of Econometrics* 30, 239-267.

Kleibergen, Frank, and Richard Paap, 2006, Generalized reduced rank tests using the singular value decomposition, *Journal of Econometrics* 133, 97-126.

Klette, Tor Jakob, Jarle Moen, and Zvi Griliches, 2000, Do Subsidies to Commercial R&D Reduce Market Failures? Microeconometric Evaluation Studies, *Research Policy* 29, 471-95.

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).

Leuven, Edwin , and Barbara Sianesi, 2003, PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing, (Boston College Department of Economics).

Li, Qi, and Jeffrey S. Racine, 2007. *Nonparametric Econometrics: Theory and Practice* (Princeton University Press).

Mansfield, Edwin , and Lorne Switzer, 1985, How Effective Are Canada's Direct Tax Incentives for R and D?, *Canadian Public Policy* 11, 241-246.

Mansfield, Edwin, and Lorne Switzer, 1985, The effects of R&D tax credits and allowances in Canada, *Research Policy* 14, 97-107.

Maté-Garcia, Jorge J., and José M. Rodriguez-Fernandez, 2008, Productivity and R&D: an econometric evidence from Spanish firm-level data, *Applied Economics* 40, 1827 - 1837.

McFetridge, Donald G., and Jacek P. Warda, 1983. *Canadian R & D incentives: Their adequacy and impact* (Canadian Tax Foundation, Toronto, Canada).

Nelson, Richard, R. , 1959, The Simple Economics of Basic Scientific Research, *Journal of Political Economy* 67, 297-306.

Pagé, L., 1995, L'incidence des subventions sur les dépenses privées en recherche et développement:une approche microéconomique, Université de Sherbrooke (Université de Sherbrooke, Sherbrooke).

Rosenbaum, Paul R., and Donald B. Rubin, 1983, The central role of the propensity score in observational studies for causal effects, *Biometrika* 70, 41-55.

Rosenbaum, Paul R., and Donald B. Rubin, 1985, Constructing a control group using multivariate matched sampling methods that incorporate the propensity score, *The American statistician* 39, 33-38.

Rubin, Donald B., 1977, Assignment to Treatment Group on the Basis of a Covariate, *Journal of Educational and Behavioral Statistics* 2, 1-26.

Smith, Jeffrey, and Petra Todd, 2003, Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?, (University of Western Ontario, CIBC Human Capital and Productivity Project).

Smith, Jeffrey, and Petra Todd, 2005, Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?, *Journal of Econometrics* 125, 305-353.

Smith, Valdemar , Mogens Dilling-Hansen, Tor Eriksson, and Erik Strøjer Madsen, 2004, R&D and productivity in Danish firms: some empirical evidence, *Applied Economics* 36, 1797-1806.

Stock, James H., and Motohiro Yogo, 2005, Testing for weak instruments in linear IV regression, in Donald W. K. Andrews, and James H. Stock, eds.: *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg* (Cambridge University Press, Cambridge).

Stone, Charles J., 1977, Consistent Nonparametric Regression, *Annals of Statistics* 5, 595-620.

Wakelin, Katharine, 2001, Productivity growth and R&D expenditure in UK manufacturing firms, *Research Policy* 30, 1079-1090.

APPENDICES

A. DESCRIPTION OF VARIABLES

Table A-1: Variable constructions and descriptive statistics

Variable	Construction	Mean	S.D.
L	Firm's number of employees	111.09	747.65
l	Firm's number of employees, in logarithm	3.05	1.61
K	Tangible assets in real terms, in million 2002 CAN$	40.00	540.00
k/l	Capital intensity, constructed as firm's real assets divided by the number of employees, in logarithm. Firm's real assets is deflated by the machinery and equipment price indexes	11.74	1.17
l_r/l	Share of R&D personnel, constructed as firm's R&D personnel divided by the number of employees	0.19	0.27
q	Firm's value added in real terms, in logarithm, deflated by the industry output price deflator at the 3-digit NAICS code	15.09	1.78
q_I	Industry sales in real terms, (at 3-digit NAICS classification), in logarithm	21.74	0.60
q/q_I	Market share, constructed as firm's real sales divided by industry's real sales, in logarithm	-6.65	1.85
e	Firm exportations in real terms, in logarithm	14.94	1.72
e/q_I	Export share, constructed as firm's real exportations divided by industry's real sales, in logarithm	-0.13	0.57
R	R&D expenditures, in million 2002 CAN$	0.59	6.8
R / Q	R&D intensity, constructed as firm's real R&D expenditures divided by firm's real value added. Firm's real R&D is deflated by the R&D deflator	0.06	0.11
Year dummies $(t_1 - t_6)$	6 year dummies: t1-t7 for years 1998, 1999, 2000, 2001, 2002, 2003 and 2004		
Industry dummies (indus1 – indus 20)	20 industry dummies: indus1-Indus20 for industries at NAICS 3 digit code level: 311, 312, 313, 314, 315, 316, 321, 322, 323, 324, 325, 326, 327, 331, 332, 333, 334, 335, 336, 337		
q/l	Labor productivity, constructed as firm's real value added divided by the number of employees, in logarithm.	12.03	1.17
$\Delta(q/l)$	Labor productivity growth rate, constructed as the logarithm of the change of the ratio of firm's real value added to the number of employees	0.02	0.80
$\Delta(k/l)$	Capital-labor growth rate, constructed as the logarithm of the change of the ratio of firm's real capital to the number of employees	0.05	0.78
Δl	Employment growth rate, constructed as the logarithm of the change of the number of employees	0.06	0.69

Note: Lower case symbol denote the logarithm of the corresponding variable. The descriptive statistics are sample means for the years 1997-2004. The base year is 2002. Number of observations: 11842 (3815 firms).

B. BALANCING TESTS RESULTS

Table B-2: Balancing tests before propensity score matching for the Average effect of treatment (ATT) on R&D intensity (**All firms**)

Variables / Statistics	1998		1999		2000		2001		2002		2003		2004	
	N_1=58, N_0=819		N_1=56, N_0=886		N_1=76 N_0=1565		N_1=56 N_0=1542		N_1=58, N_0=1925		N_1=54, N_0=2032		N_1=52, N_0=1833	
	t-stat	**Std diff (%)**	**t-stat**	**Std diff (%)**	**t-stat**	**Std diff (%)**	**t-stat**	**Std diff (%)**	**t-stat**	**Std diff (%)**	**t-stat**	**Std diff (%)**	**t-stat**	**Std diff (%)**
L	3.768***	26	4.418***	29	1.811*	16	1.959*	22	3.016***	22	5.243***	25	3.245***	24
	0.000		0.000		0.070		0.050		0.003		0.000		0.001	
K / L	6.407***	34	-1.754*	-22	-2.278**	-25	1.791*	16	-1.955*	-23	7.332***	39	9.820***	45
	0.000		0.080		0.023		0.074		0.051		0.000		0.000	
L_r / L	1.421	21	-0.640	-11	0.309	4	0.707	10	0.117	2	2.442**	33	1.567	22
	0.156		0.522		0.758		0.479		0.907		0.015		0.117	
Q / Q_I	1.715*	19	2.170**	21	3.002***	17	2.275**	18	8.695***	32	9.309***	33	5.678***	28
	0.087		0.030		0.003		0.023		0.000		0.000		0.000	
E/Q_I	-2.288**	-26	2.503**	19	3.387***	15	4.267***	23	1.833*	11	-3.038***	-36	3.930***	19
	0.022		0.012		0.001		0.000		0.067		0.002		0.000	
FOREIGN	-3.034***	-32	-2.481**	-27	-1.972**	-19	-1.670*	-18	-2.878***	-26	-1.786*	-18	-1.989**	-19
	0.002		0.013		0.049		0.095		0.004		0.074		0.047	
PSCORE	6.768***	71	4.385***	50	3.526***	54	5.856***	45	6.577***	54	7.359***	59	7.496***	66
	0.000		0.000		0.000		0.000		0.000		0.000		0.000	
R / Q	3.059***	35	2.829***	32	8.143***	60	4.889***	45	2.674***	31	7.025***	57	2.161**	29
	0.002		0.005		0.000		0.000		0.008		0.000		0.031	

Note: *** indicates significance at 1%, ** at 5%, * at 10%. Standard errors in parentheses unless otherwise indicated. The descriptive statistics are sample means for the years 1998-2004. The base year is 2002. L : number of employees, K/L : capital intensity, L_r/L : share of R&D personnel, Q/Q_I : market share, E/Q_I : export share, OWN : origin country dummy (Canada=0, Foreign=1), PSCORE : propensity score, R / Q : R&D intensity

Table B-3: Balancing tests after propensity score matching for the Average effect of treatment (ATT) on R&D intensity (**All firms**)

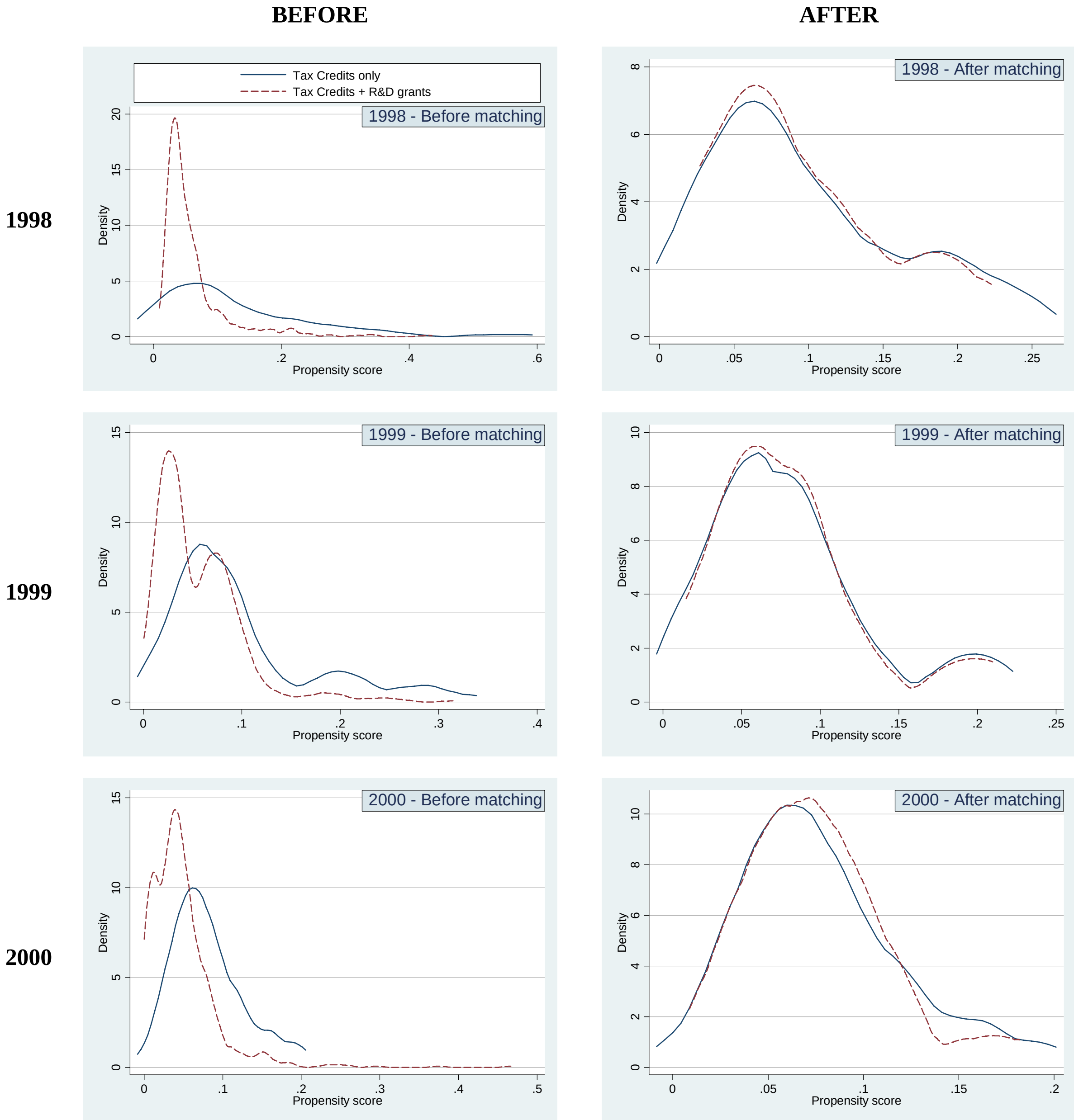
Variables / Statistics	1998		1999		2000		2001		2002		2003		2004	
	N_1=55, N_0=410		N_1=56, N_0=477		N_1=73 N_0=1063		N_1=55 N_0=1014		N_1=58, N_0=1270		N_1=53, N_0=1302		N_1=51, N_0=1180	
	t-stat	**Std diff (%)**	**t-stat**	**Std diff (%)**	**t-stat**	**Std diff (%)**	**t-stat**	**Std diff (%)**	**t-stat**	**Std diff (%)**	**t-stat**	**Std diff (%)**	**t-stat**	**Std diff (%)**
L	0.350	3	-0.275	-3	1.551	3	-0.516	-25	0.540	5	0.885	5	0.037	1
	0.727		0.784		0.125		0.607		0.591		0.380		0.970	
K / L	-0.945	-116	-0.536	-25	-0.965	-28	0.806	9	-0.277	-7	0.870	9	-0.134	-7
	0.348		0.594		0.338		0.423		0.783		0.388		0.894	
L_r / L	-0.933	-27	1.212	7	-0.614	-5	0.358	8	-0.938	-20	0.305	7	1.518	9
	0.354		0.230		0.541		0.721		0.352		0.761		0.134	
Q / Q_I	0.419	8	-0.104	-3	1.515	7	0.307	8	1.221	6	1.006	5	0.955	8
	0.677		0.918		0.134		0.760		0.227		0.319		0.343	
E/Q_I	-0.924	-21	1.472	4	1.440	0	1.478	4	-0.067	-1	-1.144	-19	0.505	8
	0.359		0.146		0.154		0.143		0.947		0.257		0.616	
FOREIGN	-0.745	-20	0.000	0	0.000	0	0.225	6	0.000	0	-0.501	-34	0.736	9
	0.459		1.000		1.000		0.823		1.000		0.618		0.465	
PSCORE	0.225	4	0.051	1	0.216	6	0.053	1	0.161	4	0.165	3	0.512	8
	0.822		0.959		0.830		0.958		0.872		0.870		0.610	
R / Q	1.835*	45	1.875*	46	1.728*	29	1.913*	45	1.751*	25	1.961*	44	1.739*	41
	0.071		0.066		0.088		0.059		0.085		0.055		0.087	

Note: *** indicates significance at 1%, ** at 5%, * at 10%. Standard errors in parentheses unless otherwise indicated. The descriptive statistics are sample means for the years 1998-2004. The base year is 2002. L : number of employees, K/L : capital intensity, L_r/L : share of R&D personnel, Q/Q_I : market share, E/Q_I : export share, OWN : origin country dummy (Canada=0, Foreign=1), PSCORE : propensity score, R / Q : R&D intensity.

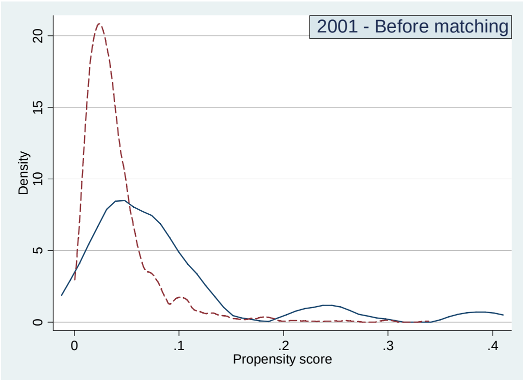
C. DISTRIBUTION OF PROPENSITY SCORE

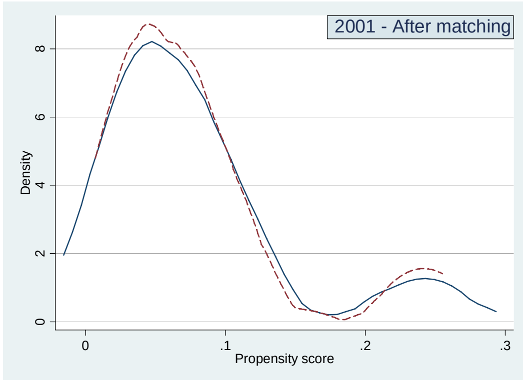
Figure C-1: Distribution of propensity score before and after matching – 1998-2004, **All firms**

Distribution of propensity score before and after matching – 1998-2004, All firms

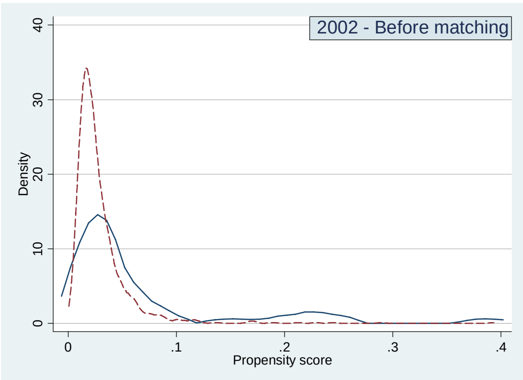


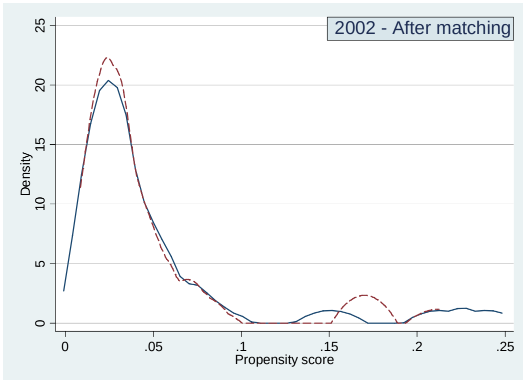
2001



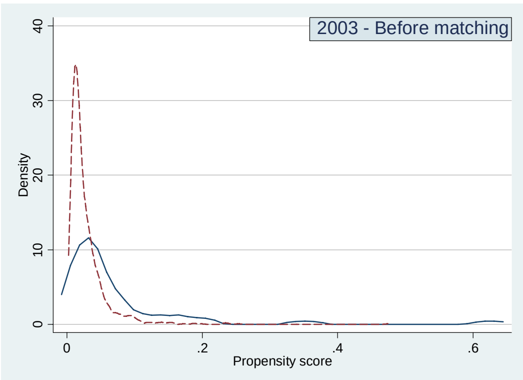


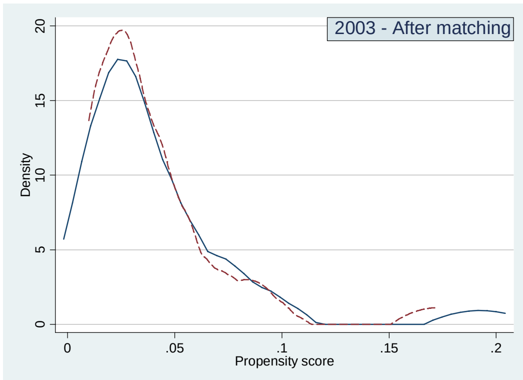
2002



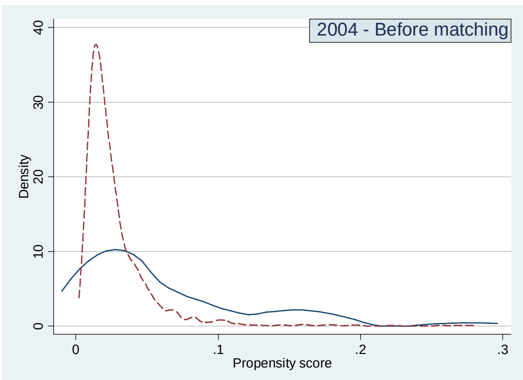


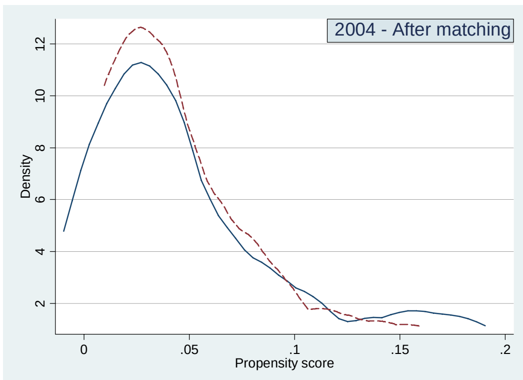
2003





2004





D. CLASSIFICATION OF MANUFACTURING INDUSTRIES

Table D-4: OECD Classification of Manufacturing Industries Based on Technology using ISIC* rev.3 activity breakdown

		ISIC rev.3 code
1	**High-technology industries**	
	Aircraft and spacecraft	353
	Pharmaceuticals	2423
	Office, accounting and computing machinery	30
	Radio, TV and communciations equipment	32
	Medical, precision and optical instruments	33
2	**Medium-high-technology industries**	
	Electrical machinery and apparatus, n.e.c.	31
	Motor vehicles, trailers and semi-trailers	34
	Chemicals excluding pharmaceuticals	24 excl. 2423
	Railroad equipment and transport equipment, n.e.c.	352 + 359
	Machinery and equipment, n.e.c.	29
3	**Medium-low-technology industries**	
	Building and repairing of ships and boats	351
	Rubber and plastics products	25
	Coke, refined petroleum products and nuclear fuel	23
	Other non-metallic mineral products	26
	Basic metals and fabricated metal products	27-28
4	**Low-technology industries**	
	Manufacturing, n.e.c.; Recycling	36-37
	Wood, pulp, paper, paper products, printing and publishing	20-22
	Food products, beverages and tobacco	15-16
	Textiles, textile products, leather and footwear	17-19

*ISIC: International Standard Industrial Classification

Source: OECD, 2005

Table D-5: Classification of Manufacturing Industries Based on the type of activity using NAICS* 3 digit code breakdown

Dummy		NAICS 3 digit code
Indus 01	Food Manufacturing	311
Indus 02	Beverage and Tobacco Product Manufacturing	312
Indus 03	Textile Mills	313
Indus 04	Textile Product Mills	314
Indus 05	Clothing Manufacturing	315
Indus 06	Leather and Allied Product Manufacturing	316
Indus 07	Wood Product Manufacturing	321
Indus 08	Paper Manufacturing	322
Indus 09	Printing and Related Support Activities	323
Indus 10	Petroleum and Coal Products Manufacturing	324
Indus 11	Chemical Manufacturing	325
Indus 12	Plastics and Rubber Products Manufacturing	326
Indus 13	Non-Metallic Mineral Product Manufacturing	327
Indus 14	Primary Metal Manufacturing	331
Indus 15	Fabricated Metal Product Manufacturing	332
Indus 16	Machinery Manufacturing	333
Indus 17	Computer and Electronic Product Manufacturing	334
Indus 18	Electrical Equipment, Appliance and Component Manufacturing	335
Indus 19	Transportation Equipment Manufacturing	336
Indus 20	Furniture and Related Product Manufacturing	337
Indus 21	Miscellaneous Manufacturing	339

*NAICS: North American Industry Classification System

Source: Statistics Canada

E. DERIVATION OF THE LABOUR PRODUCTIVITY EQUATION

Consider a typical a Cobb-Douglas production function given by:

$$Q_{it} = Ae^{\lambda_t} K_{it}^{\alpha} L_{it}^{\beta} S'^{\gamma'}_{it} S''^{\gamma''}_{it} e^{\varepsilon_{it}} \qquad (5\text{-}1)$$

where:

i = firms;
t = years;
Q = output;
A = constant;
λ = scale factor measuring the rate of disembodied technical change;
α = elasticity of output with respect to physical capital;
β = elasticity of output with respect to labour;
γ' = elasticity of output with respect to knowledge capital induced by private R&D and by tax credits;
γ' = he elasticity of output with respect to knowledge capital induced by public grants;
K = physical capital;
L = labour;
S' = share of knowledge capital induced by R&D expenditures funded privately and by tax credits;
S'' = share of knowledge capital induced by R&D expenditures funded by public grants;
ε = error term.

If we take the logarithm and subtract both sides of equation (5-1) by the logarithm of labour, we get:

$$q_{it} - l_{it} = a + \lambda t + \alpha k_{it} + \beta l_{it} + \gamma' s'_{it} + \gamma'' s''_{it} + l_{it} + \varepsilon_{it} \qquad (5\text{-}2)$$

By assuming that the constant returns to scale coefficient is given by $\theta = \alpha + \beta - 1$, this measure should equal zero in the case of constant return to scale. However, if this measure is less than zero or greater than zero, then increasing return to scale and decreasing return to scale will be assumed respectively. If we substitute $\beta = \theta + 1 - \alpha$ into equation (5-2) and re-order terms, we get the following expression:

$$q_{it} - l_{it} = a + \lambda t + \alpha (k - l)_{it} + \theta l_{it} + \gamma' s'_{it} + \gamma'' s''_{it} + \varepsilon_{it} \qquad (5\text{-}3)$$

First-differencing (5-3) yields the following:

$$\Delta(q-l)_{it} = \lambda + \alpha\Delta(k-l)_{it} + \theta\Delta l_{it} + \gamma'\Delta s'_{it} + \gamma''\Delta s''_{it} + \varepsilon_{it} \qquad (5\text{-}4)$$

Since, in the first step of the study, we used the capital intensity as a proxy for the knowledge capital, we need to transform equation (5-4) in a way such that it includes capital intensity in its functional form instead of knowledge capital. This is done by using the usual relation linking the output elasticity with regard to knowledge capital to the rate of return, which is:

$$\gamma = \rho\left(\tfrac{S}{Q}\right)_{it} \qquad (5\text{-}5)$$

where ρ is the rate of return, and by assuming that the growth rate of the knowledge capital is given by:

$$\Delta s_{it} = \left(\tfrac{\partial S}{S}\right)_{it} \qquad (5\text{-}6)$$

Equations (5-5) and (5-6) give:

$$\gamma\Delta s_{it} = \rho\left(\tfrac{S}{Q}\right)_{it}\left(\tfrac{\partial S}{S}\right)_{it} \qquad (5\text{-}7)$$

The change in the knowledge capital ∂S may be approximated by current R&D expenditures R if we assume that there is no depreciation in the knowledge capital. Then we can write that $\gamma'\Delta s'_{it} = \rho'\left(\tfrac{R'}{Q}\right)_{it}$ and $\gamma''\Delta s''_{it} = \rho'\left(\tfrac{R''}{Q}\right)_{it}$, and by substituting these expressions into equation (5-4) we get the labour productivity growth as follows:

$$\Delta(q_{it} - l_{it}) = \lambda + \alpha\Delta(k-l)_{it} + \theta\Delta l_{it} + \rho'\left(\tfrac{R'}{Q}\right)_{it} + \rho''\left(\tfrac{R''}{Q}\right)_{it} + \upsilon_{it} \qquad (5\text{-}8)$$

It is worth mentioning that equation (5-8) in which the main parameter of interest is the rate of return ρ has been widely used in the literature as an alternative to equation (5-4) in which the parameter of interest is instead the elasticity γ. This has been particularly the case because of the lack of data to construct the knowledge capital variable.

The UNU-MERIT WORKING Paper Series

2010-01 *Endogenous Economic Growth through Connectivity* by Adriaan van Zon and Evans Mupela
2010-02 *Human resource management and learning for innovation: pharmaceuticals in Mexico* by Fernando Santiago
2010-03 *Understanding multilevel interactions in economic development* by Micheline Goedhuys and Martin Srholec
2010-04 *The Impact of the Credit Crisis on Poor Developing Countries and the Role of China in Pulling and Crowding Us Out* by Thomas H.W. Ziesemer
2010-05 *Is there complementarity or substitutability between internal and external R&D strategies?* by John Hagedoorn and Ning Wang
2010-06 *Measuring the Returns to R&D* by Bronwyn H. Hall, Jacques Mairesse and Pierre Mohnen
2010-07 *Importance of Technological Innovation for SME Growth: Evidence from India* by M. H. Bala Subrahmanya, M. Mathirajan and K. N. Krishnaswamy
2010-08 *Economic Adversity and Entrepreneurship-led Growth: Lessons from the Indian Software Sector* by Suma Athreye
2010-09 *Net-immigration of developing countries: The role of economic determinants, disasters, conflicts, and political instability* by Thomas H.W. Ziesemer
2010-10 *Business and financial method patents, innovation, and policy* by Bronwyn H. Hall
2010-11 *Financial patenting in Europe* by Bronwyn H. Hall, Grid Thoma and Salvatore Torrisi
2010-12 *The financing of R&D and innovation* by Bronwyn H. Hall and Josh Lerner
2010-13 *Occupation choice: Family, Social and Market influences* by Ezequiel Tacsir
2010-14 *Choosing a career in Science and Technology* by Ezequiel Tacsir
2010-15 *How novel is social capital: Three cases from the British history that reflect social capital* by Semih Akcomak and Paul Stoneman
2010-16 *Global Players from Brazil: drivers and challenges in the internationalization process of Brazilian firms* by Flavia Carvalho, Ionara Costa and Geert Duysters
2010-17 *Drivers of Brazilian foreign investments – technology seeking and technology exploiting as determinants of emerging FDI* by Flavia Carvalho, Geert Duysters and Ionara Costa
2010-18 *On the Delivery of Pro-Poor Innovations: Managerial Lessons from Sanitation Activists in India* by Shyama V. Ramani, Shuan SadreGhazi and Geert Duysters
2010-19 *Catching up in pharmaceuticals: a comparative study of India and Brazil* by Samira Guennif and Shyama V. Ramani
2010-20 *Below the Radar: What does Innovation in Emerging Economies have to offer other Low Income Economies?* by Raphael Kaplinsky, Joanna Chataway, Norman Clark, Rebecca Hanlin, Dinar Kale, Lois Muraguri, Theo Papaioannou, Peter Robbins and Watu Wamae
2010-21 *Much ado about nothing, or sirens of a brave new world? MNE activity from developing countries and its significance for development* by Rajneesh Narula
2010-22 *From trends in commodities and manufactures to country terms of trade* by Thomas H.W. Ziesemer
2010-23 *Using innovation surveys for econometric analysis* by Jacques Mairesse and Pierre Mohnen

2010-24 *Towards a New Measurement of Energy Poverty: A Cross-Community Analysis of Rural Pakistan* by Bilal Mirza and Adam Szirmai

2010-25 *Discovery of the flower industry in Ethiopia: experimentation and coordination* by Mulu Gebreeyesus and Michiko Iizuka

2010-26 *CSR and market changing product innovations: Indian case studies* by Shyama V. Ramani and Vivekananda Mukherjee

2010-27 *How firms innovate: R&D, non-R&D, and technology adoption* by Can Huang, Anthony Arundel and Hugo Hollanders

2010-28 *Sure Bet or Scientometric Mirage? An Assessment of Chinese Progress in Nanotechnology* by Can Huang and Yilin Wu

2010-29 *Convergence of European regions: a reappraisal* by Théophile T. Azomahou, Jalal El ouardighi, Phu Nguyen-Van and Thi Kim Cuong Pham

2010-30 *Entrepreneurship and the National System of Innovation: What is Missing in Turkey?* by Elif Bascavusoglu-Moreau

2010-31 *Keeping the eclectic paradigm simple: a brief commentary and implications for ownership advantages* by Rajneesh Narula

2010-32 *Brazilian Aerospace Manufacturing in Comparative Perspective: A Brazil/USA Comparison of Output and Productivity* by Daniel Vertesy and Adam Szirmai

2010-33 *Economic restructuring and total factor productivity growth: Tunisia over the period 1983-2001* by Sofiane Ghali and Pierre Mohnen

2010-34 *Impact of government support on R&D and innovation* by Abraham Garcia and Pierre Mohnen

2010-35 *Product, process and organizational innovation: drivers, complementarity and productivity effects* by Michael Polder, George van Leeuwen, Pierre Mohnen and Wladimir Raymond

2010-36 *Entrepreneurship Development and the Role of Economic Transition in Entrepreneurial Activities in China* by Ying Zhang and Geert Duysters

2010-37 *Pro-Poor, Entrepreneur-Based Innovation and it's Role in Rural Development* by Lina Sonne

2010-38 *Financing pro-poor entrepreneur-based innovation: A review of existing literature* by Lina Sonne

2010-39 *India's Rural Financial System: Does it Support Pro-Poor Innovation?* by Lina Sonne

2010-40 *How effective are level-based R&D tax credits? Evidence from the Netherlands* by Boris Lokshin and Pierre Mohnen

2010-41 *Analysing Multidimensional Poverty in Guinea: A Fuzzy Set Approach* by Fatoumata Lamarana Diallo

2010-42 *Bottom-up, Bottom-line: Development-Relevant Enterprises in East Africa and their Significance for Agricultural Innovation* by Andy Hall, Norman Clark and Andy Frost

2010-43 *Assessing Innovations in International Research and Development Practice* by Laxmi Prasad Pant

2010-44 *Research Into Use: Investigating the Relationship between Agricultural Research and Innovation* by Andy Hall, Jeroen Dijkman and Rasheed Sulaiman V.

2010-45 *The global institutional frameworks and the diffusion of renewable energy technologies in the BRICS countries* by Isabel Maria Freitas, Eva Dantas & Michiko Iizuka

2010-46 *The role of patent protection in (clean/green) technology transfer* by Bronwyn H. Hall and Christian Helmers

2010-47 *Localisation Strategies of Firms in Wind Energy Technology Development* by Radhika Perrot and Sergey Filippov

2010-48 *The R&D activity of multinational enterprises in peripheral economies: evidence from the EU new member states* by Rajneesh Narula and José Guimón

2010-49 *Bridging the Financing Gap for Pro-Poor Innovation: Towards a Framework* by Lina Sonne

2010-50 *Efficient Development Portfolio Design for Sub Saharan Africa* by Adriaan van Zon and Kirsten Wiebe

2010-51 *Global excellence at the expense of local relevance, or a bridge between two worlds? Research in science and technology in the developing world* by Helena Barnard, Robin Cowan, Moritz Müller

2010-52 *Innovation strategy, firm survival and relocation: The case of Hong Kong-owned manufacturing in Guangdon province, China* by Naubahar Sharif and Can Huang

2010-53 *Determinants of PRO-industry interactions in pharmaceutical R&D: the case of Mexico* by Fernando Santiago Rodriguez and Gabriela Dutrenit

2010-54 *New Organisational and institutional vehicles for managing innovation in South Asia: Opportunities for using research for technical change and social gain* by Vamsidhar Reddy, T.S., Andy Hall and Rasheed Sulaiman V.

2010-55 *Public R&D subsidies and productivity: Evidence from firm-level data in Quebec* by Rufin Baghana