Are R&D subsidies a substitute or a complement to privately funded R&D? Evidence from France using propensity score methods for non-experimental data[1]

Emmanuel DUGUET[2]

Revised, May 2003

Cahiers de la MSE – EUREQua 2003(75)

[1] Financial support from the Ministère de l'Industrie (DIGITIP-SESSI) is gratefully acknowledged. I thank INSEE "Marchés et Stratégies d'Entreprises" Division, Bruno Crépon, Louis de Gimel, Elisabeth Kremp and Guy Laroque for their comments as well as the participants to *the International Conference on Technological Policy and Innovation* (Paris, November 2000), to the 19th Journées de Microéconomie Appliquée (Nancy, June 2001), the GRIT Workshops as well as to the CREST, EUREQua and Fourgeaud seminars.

[2] Université de Bretagne Occidentale and EUREQua - CNRS UMR 8594 – Université de Paris I, Maison des Sciences Économiques, 106-112, bd de l'Hôpital, 75647 Paris Cedex 13.

Are R&D subsidies a substitute or a complement to privately funded R&D? An econometric analysis at the firm level

This study examines the effect of research and development subsidies on the private funding of R&D in France. We address this issue from the annual R&D survey over 1985-1997, which provides information about the R&D subsidies given by all the ministries to the firms having at least one full-time person working on R&D. In order to determine whether the supported firms would have invested the same amount of private R&D without the subsidies, we use matching methods. We show that the use of these methods is important because the global evaluations, in this paper, more often give a potential effect among the non-supported firms than a real effect among the supported firms. We first study the probability to get a subsidy. We find that this probability is increasing with size, the debt ratio and the importance of privately funded R&D. In a second step, controlling for the past public support the firms benefited from, we find that, on average, public funds add to private funds, so that there would be no significant crowding out effect.

Keywords: propensity score, non-experimental data, policy evaluation, research and development, subsidies.

JEL Classification: C14, H25, L98, O38.

Les subventions à la recherche et développement sont-elles substituables ou complémentaires à leur financement privé ? Une étude micro-économétrique.

Cette étude examine les effets des subventions à la recherche et développement (R&D) sur le financement privé de la recherche en France entre 1985 et 1997. Les données disponibles portent sur les subventions données par tous les ministères aux entreprises employant au moins un chercheur en équivalent temps-plein. Afin de déterminer si les entreprises auraient investi les mêmes montants sur fonds privés avec et sans les subventions, nous utilisons une méthode d'estimation par appariement. Nous montrons que l'emploi de cette méthode est important car les évaluations globales, dans notre application, traduisent plus souvent un effet potentiel sur les entreprises non-aidées qu'un effet réel sur les entreprises aidées. Dans un premier temps, nous étudions la probabilité d'obtenir une subvention. Cette probabilité est croissante avec la taille de l'entreprise, son ratio d'endettement et son effort de recherche privée. Dans un second temps, et en tenant compte des subventions que l'entreprise a obtenues par le passé, nous trouvons qu'en moyenne les financements publics s'ajoutent aux financements privés, de sorte qu'il n'y aurait pas d'effet d'éviction significatif.

Mots-clés : appariement sélectif, données non expérimentales, évaluation de politique microéconomique, recherche et développement, subvention.

Introduction

Economists generally agree that innovation is an important source of economic growth. A large number of empirical studies at the firm level clearly confirm this view. Originally the data available established a connection between research and development investments and the productivity differences between the firms, and, more recently the Innovation surveys conducted in several countries, have allowed for extending this result to a wider set of innovation and performance measures. Therefore, the issue of private incentives to research investment is crucial.[3]

In most situations the market will fail to provide sufficient private incentives to innovation since innovators face appropriability problems (Nelson, 1959; Arrow, 1962). The reason is that R&D has some characteristics of a public good, so that the private return on innovation will be lower than its social return. Moreover, the cost structure of innovation is specific in that it relies on a sunk cost (Cohen and Klepper, 1996).

Two categories of economic policies have been implemented in order to encourage private R&D: the first category aims to increase its private return, the second category aims to decrease its private sunk cost. In order to increase the private return on innovation, many states have introduced intellectual and industrial property rights in order to avoid imitation. However, this patent protection is not always efficient so that a combination of policies is needed. For instance, Levin et al. (1987) showed that American firms do consider patent neither as the only mean of appropriation nor as the most efficient. Lead-time or secrecy would be more efficient in securing innovation profits (Crampes, 1986).[4] It appears that, in most situations, a part of the firms patent and that, when they do, they patent a fraction of their innovations only. Globally, patenting will be efficient when knowledge can be easily codified and when innovations cannot be circumvented (Mansfield et al., 1981).

The second category of innovation policy is to reduce the private sunk cost of R&D. This policy will make more research projects accepted by private parties. But it also involves an additional advantage when the associated knowledge is difficult to protect. Since subsidies are not property rights, they do not impeach the innovations to be spread through the imitation by competitors. And the diffusion of knowledge is always good from a social perspective. This paper addresses the issue of subsidies to private R&D, that is

[3] For evidence at the firm level, see Griliches (1998) and Hall and Mairesse (1995) on the United States, Crépon, Duguet and Mairesse (1998), Crépon and Iung (1999), Duguet (2002a) on France, Lööf and Heshmati (2001) on Sweden, Jefferson et al. (2001) on China, Van Leeuwen and Klomp (2001) on the Netherlands, Lööf et al. (2001) on Finland and Norway, Arvanitis and Hollenstein (2002) on Switzerland.

[4] These results have been found in other countries, see Duguet and Kabla (1998) for France and Arundel and Kabla (1997) for other European countries.

of research that is managed by the firms themselves.[5] The main difference with the research conducted in public institutions is that while public laboratories mostly perform fundamental and applied research, private firms mostly perform applied research and development.[6]

The goal of the subsidy policy is clear: by granting a subsidy, one hopes that additional research projects will take place compared to the ones that would have been done without public support. But there are two limits to this policy: On the one hand, the *private* return on development is *a priori* higher than the private return on fundamental research, so that the justification of private R&D subsidies is lower than the justification of public research. On the other hand, private firms have access to other funding opportunities, so that a substitution between R&D subsidies and private R&D cannot be excluded (David, Hall and Toole, 2000; Wallsten, 2000).

The latter argument illustrates the very important role that the private return on innovation has on the efficiency of a subsidy policy. It also implies that one should be very careful about the nature of the performance variables used in evaluation studies. In general, it is clear that the less the private return on R&D the more a subsidy will be useful. It is the reason why standard performance measures will not be convenient for evaluation studies. For instance, productivity may be a good measure when R&D is privately funded, but it may not when it is publicly funded because the justification of the subsidy is precisely that the private return is to low to justify the R&D project from a private viewpoint. It would not be surprising that, in the latter case, we find that the private performance of subsidized firms is smaller than the performance of the fully privately funded firms. In fact, such a finding could merely mean that the subsidies were given to the right projects. This argument suggests that the evaluations that rely on private returns only could underestimate the effect of research policies.

The simplest variable to use for the evaluation is necessarily privately funded R&D itself. This must be so because there can be no effect of a policy if there is a full substitution between public and private funds.[7]

The evaluation methodology is equally important. What we need to evaluate is the difference between the private funding that firms have made with the subsidy and the private funding they would have made without it. In order to evaluate this difference, we use the propensity score method introduced by Rubin (1974), where subsidized firms are matched with the most similar unsubsidized firms.

[5] Notice that this research can be made by the firm itself or in cooperation with other private or public institutions.

[6] For evidence in France, see Crépon and Duguet (1996).

[7] This explains why such a large number of studies focus on this variable. See David, Hall and Toole (2000) for a survey.

The second section presents a theoretical analysis of the effect of subsidies and links it to recent previous studies in the field. The data are presented in the third section and the methodology in section 4. The last section presents our results.

1 Theoretical background and previous empirical studies

This section presents a simple model of R&D subsidies and a discussion of the results of previous studies. The estimates that we provide in this paper can be interpreted as a non-parametric estimation of this model. However, it is important to notice that none of the assumptions of this model is needed to perform the estimation. In a first step, we establish the conditions under which substitution happens between private R&D and subsidies. In a second step, we derive the empirical counterpart implied by the theoretical model in order to easier the interpretation of the empirical analysis that follows.

1.1 The model

We consider a situation in which a ministry (or an agency) offers subsidies to innovative firms. Any firm can apply for a subsidy and must provide information on its research project, including the amount of privately funded R&D. After considering this information, the ministry decides the amount of public support it grants to the firm on an individual basis.[8] The projects that do no include any privately funded R&D are excluded from the subsidy system.[9] We represent the functioning of this public support system by a Nash (1953) bargaining process on both the amount of privately funded R&D and the amount of the subsidy.

On the one hand, firms maximize profits and can apply for a subsidy even if they do not need it. It will be the case when the private return on R&D is high. On the other hand, we assume that the ministry seeks to maximize the total amount of R&D that the firms perform. This assumption can be justified by the goal of maximizing the number of innovations, which depend on the total amount invested. Moreover, it accounts for possible substitution effects, since a decrease of private R&D lowers total R&D.

We distinguish two cases, which correspond to two different threat points of the bargaining process. In the first case, the private return on R&D is so low that no innovation project would be undertaken with private funds. The subsidies should be efficient. In the second case, the private return on R&D is high enough for the research to be privately funded and we expect substitution to take place between public and private funds.

[8] Decisions taken on an individual basis creates an important difference with the R&D tax credit that is attributed on an automatic basis.

[9] This assumption was motivated by our data.

The goal of our model is to determine the difference between the amount of privately funded R&D when the firm has a subsidy (denoted R_1) and the amount of privately funded R&D the firm would have invested without a subsidy (denoted R_0). The effect of the subsidy is denoted:

$$c = R_1 - R_0$$

The profit function of the firm is given by:

$$\Pi(R, A) = V(R + A) - R$$

Where R is privately funded R&D, A the subsidy and V the revenue of total research that can be obtained either directly by a production activity or indirectly through the trade of property rights. We assume that the revenue function is given by:

$$V(x) = \begin{cases} \phi x & \text{f } x < \bar{x} \\ \phi \bar{x} & \text{if } x \geq \bar{x} \end{cases}$$

Where x is the total amount invested in R&D. This specific form has three interesting properties. First, it assumes that sales are proportional to total R&D, a stylized fact that has been documented by many econometric studies (Cohen and Klepper, 1996). Second, the private return on R&D is represented by one parameter, ϕ, that depends on the imitation rate of products, the enforcement of intellectual property rights or the anticipated market power of the firm. Third, it implies that the optimal amount of R&D has a Tobit shape, where there is no investment if the private return on R&D is too low and a continuum of investments when it is high, since each firm will invest a different amount depending on the value of its firm-level ϕ parameter.[10]

The ministry aims to maximize total R&D expenditures. Its gain function is thus:

$$G(R, A) = R + A$$

An increase of the subsidy has the following impact on the ministry's gain:

$$\frac{\partial G}{\partial A} = 1 + \frac{\partial R}{\partial A}$$

This gain will thus be lower in the presence of substitution between public and private funds ($\partial R / \partial A < 0$), so that the ministry will has an incentive to restrict the allocation of subsidies when substitution takes place.

The objectives of the two players (firm and ministry) lead to maximize the following Nash bargaining criterion:

[10] The tobit form of R&D investment is a standard of the econometrics of research. See Bound et al. (1984) on American data and Crépon, Duguet and Kabla (1996) on French data.

$$N(A,R) = (\Pi(A,R) - \Pi_0)^{1-\theta}(G(A,R) - G_0)^{\theta}$$

Where $\theta \in [0,1]$ represents the bargaining power of the ministry and (Π_0, G_0) the threat points that are defined as the gains of the two players when no subsidy is granted (i.e., a bargaining failure is often represented as the non-cooperative outcome of the game). The bargaining power of the ministry depends on the conditions applying to the grant of subsidies and on the number of firms competing for a fixed subsidy budget. We must consider two cases that involve two different threat points:

- When the private return on R&D is low, the firm would not invest in R&D without a subsidy and the ministry would records no R&D. The threat point is $(\Pi_0, G_0) = (0,0)$.
- When the private return on R&D is high, the firms would invest an amount R_0 and the threat point is $(\Pi_0, G_0) = (\Pi(R_0, 0), R_0)$

1.2 When the private return on R&D is low

In this situation, the firm does not invest without a subsidy. The condition is that there exists no private R&D amount that provides a positive profit: [11]

$$\Pi(R,0) = (\phi - 1)R < 0 \quad \forall R > 0 \Rightarrow \phi < 1$$

In this situation, the optimal amount of private R&D is $R_0 = 0$ so that $\Pi_0 = 0$. The firm has therefore to apply for a subsidy in order to realize its research project. The project must have a part that is privately funded. The Nash criterion is equal to:

$$N = (\Pi - 0)^{1-\theta}(G - 0)^{\theta} = ((\phi - 1)R + \phi A)^{1-\theta}(A + R)^{\theta}, \; 0 < \phi < 1$$

The first result is that the ministry has interest in spending its entire subsidy budget since:

$$\frac{\partial N}{\partial A}(R, A) > 0, \; \forall A > 0$$

The amount invested by the firm under the subsidy grant is given by:

$$\frac{\partial N}{\partial R}(R, A) = 0 \Leftrightarrow R_1 = A\left(\frac{\theta}{1-\phi} - 1\right)$$

and the subsidy will be granted only if the firm invests some money of its own into the project:

$$R_1 > 0 \Leftrightarrow \phi > 1 - \theta$$

[11] It is equivalent to say that the maximum of profit is negative.

If the ministry has a strong bargaining power (θ close to 1) this condition reduces to $\phi > 0$ and all the projects that have a low private return will be undertaken. In this case, we can show that the whole revenue of the firm is invested in R&D. In the other cases, the lower the private return on R&D the less the firm invests.

Overall, the effect of subsidies on private R&D when the private return is low equals:

$$\bar{c} = R_1 - R_0 = A\left(\frac{\theta}{1-\phi} - 1\right) > 0 \quad \forall\, 1-\theta < \phi < 1$$

As expected, it is when the private return on R&D is low that subsidies can have a positive effect on private R&D.

1.3 When the private return on R&D is high

In this situation $\phi > 1$, the private R&D activity is profitable. The optimal investment equals:

$$R_0 = \bar{R}$$

and the corresponding profit equals:

$$\Pi_0 = (\phi - 1)\bar{R} > 0\,.$$

Even though the firm does not need public support, it has an obvious interest to apply for a subsidy anyway. The Nash criterion equals:

$$N = (\Pi - \Pi_0)^{1-\theta}(G - G_0)^{\theta} = \left((\phi-1)R + \phi A - (\phi-1)\bar{R}\right)^{1-\theta}\left(A + R - \bar{R}\right)^{\theta},\ \ \phi > 1$$

This can be simplified to:

$$N = \left((\phi-1)c + \phi A\right)^{1-\theta}(A + c)^{\theta}$$

with $c = R - \bar{R}$ that measures directly the effect of the subsidy at the optimum (in this case: $c = R_1 - R_0$). The shape of function N is the same as the one of the previous section, but with c instead of R. Therefore, we find that the ministry has an incentive to spend all its budget and that the optimal level of private R&D equals:

$$R_1 = \bar{R} - A\left(\frac{\theta}{\phi-1} + 1\right) < \bar{R} \quad \forall\, \phi > 1\,.$$

There is always a substitution effect. Nevertheless, this effect is limited by the following participation condition:

$$R_1 > 0 \Leftrightarrow A < \bar{A} = \left(\frac{\phi - 1}{\phi - 1 + \theta}\right)\bar{R}\,.$$

The ministry has an interest in limiting the subsidy it grants to the firms that participate to the project. Therefore a part of the substitution effect will never take place because the firms need to put some money of their own.

Overall, there is always a decrease of private R&D expenditures when a subsidy is granted to projects with a high private return:

$$\underline{c} = R_1 - R_0 = -A\left(\frac{\theta}{\phi - 1} + 1\right) < 0 \quad \forall \phi < 1$$

This model provides a simple rule to determine whether a research project would have needed public funds or not: firms that decrease their private R&D after a subsidy grant are likely to have a private return that is high enough to fund the research of their own. This property can be linked to our evaluation method.

1.4 Interpretation of econometric results

The data include both projects that need public funds and projects that do not. Therefore, we will observe a weighted average of the effects $\underline{c}$ (substitution case) and $\overline{c}$ (complement case).

Our results imply that the estimated effect of subsidies can vary over time depending on the variations of appropriability conditions and the bargaining power of the ministry. The first parameter depends on many conditions: the enforcement of property rights, the nature of the knowledge that is developed over time, the fact that the ministry prefers to give support to different technologies over time or simply the variations of the expectations of the firm about the private value of the innovative products and processes. The second parameter may vary according to the amount of subsidy available: when a small budget is available the bargaining power of the ministry may be stronger because firms compete for subsidies, when the budget is large the ministry may be tempted to give either higher subsidies to the same projects or to increase the number of recipients. In this case we should expect that beyond some threshold of average subsidies - below some threshold of bargaining power - substitution effects could occur.

According to the model, we will examine two measures of performance. The first one is a dummy variable indicating whether the firm has maintained or increased its privately funded R&D. The second measure is the growth rate of privately funded R&D. The use of growth rates instead of levels in the latter case is further motivated by the need to eliminate fixed effects when performing the evaluation (see Crépon and Iung, 1999, appendix 2).

1.5 Previous studies and the model

The recent studies that we comment use a methodology that is comparable to ours and are natural background for international comparisons.[12] We discuss their result according to the model.

Toivanen and Niininen (1998) examine whether R&D subsidies are substitutes or complements to the private funding of innovative activities. They study this issue on Finish data over the period 1985-1993. The subsidies can represent as much as 50% of the total R&D budget and a loan can be granted up to 60% of the total funding. For the small firms, the thresholds are increased by 10%. Two main results are obtained. On the one hand, for the large firms, the R&D subsidies would have no effect on the private funding of private R&D. On the other hand, for the small firms, the subsidies would increase the private funding of R&D by 5%. Globally, a simple addition effect dominates. Since the small firms are likely to experience less favorable appropriability conditions, these results are compatible with our model.

Busom (1999) studies the effect of the R&D subsidies given by the Spanish ministry of industry in 1998, that funds on average 39% of the total R&D budget. A first look at the data shows that theses subsidies are more often granted to small-sized firms that have applied for a large number of patents over the ten years before the subsidy was granted.[13] The results are mixed. On the one hand, for two firms out of three, the subsidies increase the private funding of R&D by 20%. On the other hand, for the remaining third of the firms, there would be a complete crowding out. The explanation of this result could simply be that a third of the firms benefit from good appropriability condition. It also shows that the global effects of subsidies that are measured in applied studies are the aggregation of positive and negative effects, so that the heterogeneity of the effect of subsidies should be accounted for in the evaluation. Our methodology will therefore account for it.

Wallsten (2000) studies the impact of the Small Business Innovation Research Program (SBIR) on American data between 1990 and 1992. In 1998, one billion dollar has been spent on it. In the first step, the agencies can grant a subsidy of one hundred thousand dollars to a firm in order to examine the possibility to implement an idea (the applied research step). In a second step, the agencies can grant a subsidy up to seven and fifty hundred thousand dollars to implement the project that had been supported in the first step (the development step). In a third step, no subsidy is granted but the firm has to market its product or to use its process. The conclusion of the author is that the R&D investment would have been made even without the subsidies; the reason advocated is that the projects selected by the agencies tend to favor the projects with the highest

[12] Nevertheless, the reader should keep in mind that the type of subsidy granted varies from one country to another.

[13] A part of the subsidies are financed the European Union. When this type of public support is considered, the author finds that the attribution of a subsidy increases with the size of the firm.

probability of commercial success. And this is equivalent to select the projects with the highest private return on R&D. The conclusion of the model therefore applies and a full crowding-out would not be surprising. This paper is also interesting because it shows that the substitution between subsidies and private R&D can originate in the criterion that the agencies use to grant the subsidies.

Lach (2000) studies the effect of R&D subsidies on Israeli data over 1990-1995. The public support is given by the ministry of industry and foreign trade and represented 120 millions of dollars in 1988 and 310 millions of dollars in 1990. The granting criterion is the ability of the firm to export its research output and most projects have a flat subsidy rate of 50%. Therefore, a reduction of private R&D involves a reduction of the subsidy. Such a mechanism should entice firms to maintain their private R&D.[14] The first part of the study uses matching methods while the second part uses regression methods on balanced panel data. The first part concludes that subsidies add to the private funding of R&D (table 10, page 24). This first result is confirmed by the first panel data regressions that uses a subsidy dummy variable (table 11, p. 29). Therefore, the author first finds an addition effect. The second part of the study relies on a very different methodology and estimates a reduced form on the subsidy level. The author finds a different result: one additional dollar of R&D would increase private R&D of 41 cents on the long run. Therefore, there would be a positive effect of R&D subsidies. Since exporting firms face international competition and can have more appropriability problems abroad that on their home market, the positive effect on private R&D was expected. However, the first result, an addition of subsidies and private R&D, suggest that the average effect of subsidies may be an aggregation of positive and negative effects so that the effect of subsidies on private R&D may be heterogeneous.

Finally, Czarnitzki and Fier (2001, 2002) study the effect of subsidies on the German firms operating in services over 1994-1998. In their first study, the authors use standard regression methods and find that on average one Euro of subsidy would increase private R&D by 1.3 to 1.4 Euros. In the second study, the authors implement matching methods and show that the average private R&D to sales ratio is higher (by 8 points) among subsidized firms. However, no analysis of the growth rate of R&D has been made. The results go in the same direction than Lach (2000) but with a stronger impact of subsidies on R&D. This difference can also be explained by our model: the firms in services would have more difficulties to protect their knowledge because the property rights - at the time of the study - allow for less protection in services than in industry (where patents can be used more efficiently). Therefore we should expect - as long as the German legislation is not changed to cover business methods, for instance - that the effect of subsidies should be stronger in services.

[14] A similar mechanism is present under the R&D tax credit, for which a positive effect is found on privately funded R&D (Hall and Van Reenen, 2000).

2 The data

2.1 Sources

The quantitative data come from two sources: the « BRN » (from INSEE), that are the fiscal files for all firms above ten employees, and the R&D survey (from the Ministry of Research).[15] The BRN provide information about the line of business of the firms as well as their main accounting data. The research survey provides information about the R&D investments and the subsidies granted by all the ministries in France. It does not include information about the R&D tax credit.[16] The main sources of funding are the Ministry of Defense, the Ministry of Industry and the Ministry of Research. In order to keep the larger number of firms in our analysis we have made 12 separate samples including to consecutive years over the period 1985-1997. The necessity to include two year of consecutive data comes from the fact that we work on growth rates and that we need lagged control variables for estimating the subsidy granting equation. The details of the dataset are presented in appendix 1. Depending on the year, the samples include between 1032 and 1672 research performers in both industry and services.

2.2 Sample statistics

The total R&D expenditures of the sample are strongly increasing up to 1991, is stable up to 1994 and then decreases (Figure 1). The origin of this decrease mainly comes from the decrease of subsidies. Indeed, private R&D slightly decreases since 1995 while the subsidies decrease since 1992. In order to correct these first results for the variation of the number of firms in the sample, we have computed the average R&D expenditures (Figure 2). Over the same period, the average R&D expenditures are first increasing up to 1991, where they reach ten million Euros and then steadily decrease to reach a little more than 8 millions Euros in 1997. This evolution can be decomposed between the publicly and the privately funded R&D. The average privately funded R&D is increasing up to 1993 (8.5 billions Euros) where it decreases to a stable value of 8 millions Euros. But this decrease cannot explain the decrease of the average total R&D expenditures. An important part of that fall comes from the R&D subsidies. At the beginning of the period (1985-1991), where the average subsidy oscillates between 1.5 and 2 millions Euros, the average subsidy decreases steadily down to 0.6 million Euros in 1997. Globally, the decrease in public funding has not been compensated by an increase in private funding.

These first results are interesting if we consider them in the debate about crowding-out effects. If there was a substitution between the variations of public and private funds, we

[15] The statistical treatment on the BRN files has been made at INSEE « Marchés et Stratégies d'Entreprises » Division.

[16] This is because the R&D tax credit in not managed by the Ministry of Research. This should not be a problem to our study since all firms can apply to the R&D tax credit under the same conditions.

should observe a negative correlation between public and private R&D. The Figure 2 shows that this is the case on two short periods only, in 1987-1989 and 1991-1993 where a decrease of public funds happens at the same time as an increase in private funds. But over the whole period, out of twelve variations, only four do not contradict the substitution hypothesis. The other cases that we see are either an increase of both public and private funds (1985-87 and 1989-91) or a decrease of both (1993-97).

The origin of this decrease of public funds is illustrated in Figure 3. The fall comes from the subsidies granted by the Ministry of Defense. After a relative stability around 1.3 millions Euros up to 1991, the average subsidy diminishes steadily down to 0.4 millions of Euros in 1997, that is three times less.

Looking more in the details of the fall, we should distinguish the variation of the number of recipients from the variation of the average amount granted to each recipient (Table 1). In 1985, one firm out of three benefited from a subsidy, in 1997, the ratio is one firm out of four. Over the same period, the share of publicly funded R&D felt from 24% to 7% of the total R&D expenditures.

Table 2 describes the entry and the exit from the subsidy system. The majority of firms do not have subsidies (65% in 1997) and at the same date 22% had a subsidy for two years in a row. Overall, the stability of the recipients is rather strong. The percentage of entrants in the subsidy system is 5% and the percentage of firms that exit the subsidy system is 7%. The latter figure is higher because the fall in subsidy has been made partly by reducing the number of recipients.

However, this decrease of the number of recipients cannot explain alone the fall of the public funding. Figure 4 shows that the average subsidy by recipient has strongly diminished since 1991, from about 6 millions Euros to about 2 millions Euros. This decrease will have different effects on the private funding on R&D depending on whether there is substitution or not. A first look at this issue can be obtained from Table 3 that compare the average performances of the supported and the non-supported firms. Two performance measures have been retained for our analysis: the fact to have maintained of increased the privately funded R&D (dummy variable) and the growth rate of private R&D.

Overall, the percentage of firms that have maintained or increased their private funding is significantly higher among firms that have not been supported for 5 years out of 12. These first results would rather go for the substitution hypothesis. But the comparison of the growth rates show that the difference is significant for one year only (in 1987), a result that goes against the substitution hypothesis.

The comparison of Table 3 however could suffer from a selection bias, since it sums up both the effects of the subsidies and of all the characteristics of the firms that are correlated with it. For instance, we will see later that the supported firms do not have the same size and do not operate in the same lines of business than the non-supported ones. The comparison could therefore reflect the latter differences rather than the effect of subsidies. In order to fix that problem, we use matching methods.

3 Methodology

3.1 The evaluation problem

The methodology used in this study starts from the causal model of Rubin (1974) designed to evaluate the effect of a treatment from non-experimental data.[17] According to this approach, a treatment (denoted *T*) is applied to a sample of individuals with characteristics *X* in order to improve their performance *y.* Here, the treatment is a dummy variable indicating whether a firm got a subsidy (*T=1)* or not (*T=0)*. We assume that each firm will react in a different way to the subsidy so that it has two potential outcomes:[18]

- $y_i(0)$ if the firm is not supported;
- $y_i(1)$ if the firm gets a subsidy.

What we seek to evaluate, the “causal effect”, defined as the difference between these two potential outcomes, $c_i = y_i(1) - y_i(0)$. More precisely, three quantities are of interest:

- $c = E(y_i(1) - y_i(0))$: the average effect of the subsidies on the whole population of firms;
- $c_1 = E(y_i(1) - y_i(0) | T = 1)$: the average effect on the treated. It is the standard definition of the evaluation;
- $c_0 = E(y_i(1) - y_i(0) | T = 0)$: the average effect on the non-treated, which measures the opportunity to grant subsidies to the firms that do not have it.

The relationship between these three quantities is given by:

$$c = c_0 \Pr[T = 0] + c_1 \Pr[T = 1].$$

This relationship is important because it implies that a negative global effect (*c*) can originate from two different cases: either $c_1 < 0$, when the subsidy is not efficient among the supported firms, or $c_0 < 0$, when the subsidy would not be efficient on the not-supported firms. This is why one should be cautious when interpreting the results of the standard regression methods. In the standard case, these regressions estimate *c,* not c_1, so that a negative global effect can be found even when the effect on the treated is positive. This problem is more likely to happen when the probability to get a subsidy is

[17] For a comprehensive presentation see Rubin (1997), on the developments of the methodology see Brodaty, Crépon and Fougère (2002).

[18] One reason why the effect may vary from one firm to another is simply that the subsidies come from different ministries. Moreover, the previous empirical studies in the field show a significant amount of heterogeneity in the responses of the firms to R&D subsidies.

low, since this probability is the weight of the effect on the treated in the average effect equation. This problem does occur in our data.

The evaluation problem is that one cannot observe $y_i(1)$ and $y_i(0)$ at the same time. Either the firm is subsidized and we observe $y_i(1)$, or it is not and we observe $y_i(0)$. The observable outcome is therefore equal to:

$$y_i(T_i) = y_i(0)(1 - T_i) + y_i(1)T_i, \quad T_i \in \{0,1\}.$$

What we need to evaluate are thus the expectation of $y_i(0)$ for the firms that have been subsidized, and the expectation of $y_i(1)$ for the firms that have not been. It is the problem of the comparison group. For instance, in order to evaluate c_1, we need to select firms that have not been treated but that would have behaved like the treated.

3.2 The naïve estimator

The simplest method reduces to take the difference of the averages y_{0i} and y_{1i} in order to estimate $E(y_0)$ and $E(y_1)$, like in Table 3. But this comparison implicitly assumes that the conditional and the marginal expectations are equal: $E(y_0) = E(y_0|T=0)$ and $E(y_1) = E(y_1|T=1)$. These equalities can be obtained when one assumes that the treatment is independent of the performance, that is:

$$y \perp T \Rightarrow E(y_k|T=k) = E(y_k), \quad k \in \{0,1\}$$

It is clear that this assumption is too strong for this study because it would imply that the private funding of research without public support is independent from belonging to the group of the subsidized firms. To illustrate the problem further, the difference of averages is the empirical counterpart of:

$$\begin{aligned}\gamma &= E(y_1|T=1) - E(y_1|T=0) \\ &= E(y_1|T=1) - E(y_0|T=1) + E(y_0|T=1) - E(y_0|T=0) \\ &= c_1 + \beta\end{aligned}$$

where β is the selection bias. This bias is equal to zero only if the performance without public support is the same among treated and non-treated firms. This holds under the assumption:

$$y_0 \perp T \Rightarrow E(y_0|T=0) = E(y_0).$$

and we have many reasons to believe that this condition will not be fulfilled in our study.[19] Among them:

- The firms in the two groups do not have the same size and R&D investment is proportional to size ;
- The firms in the two groups do not operate in the same lines of business and the appropriability conditions, the demand and the technological opportunities vary strongly among different lines of business ;
- The firms that have a subsidy often had one in the past and this may influence their private investment in R&D through the projects they have begun.

Therefore we need to generalize the naïve estimation method by relaxing its assumptions.

3.3 The conditional estimator

If we had experimental samples, we could use the difference of the average performances. Here, it is not possible because the subsidies are not attributed at random due to the existence of a policy. The simplest method is thus to find attributes *X* such that, among firms that have similar attributes, the allocation of the treatment can be considered as attributed at random. The identification assumption for the effect on the treated is: [20]

$$y(0) \perp T | X \Rightarrow E(y(0)|T = 1, X) = E(y(0)|T = 0, X)$$

Under this assumption, the causal effect is given by:

$$\begin{aligned} c_1 &= E(y(1) - y(0)|T = 1, X) \\ &= E(y(1)|T = 1, X) - E(y(0)|T = 1, X) \\ &= E(y(1)|T = 1, X) - E(y(0)|T = 0, X) \end{aligned}$$

Integrating over *X,* we get the average causal effect on the treated. Hence, one needs to regroup firms in homogenous *X* classes before to make a comparison. The most popular method is to match each treated firm with its non-treated neighbors (Rubin, 1979). The first applications of this method used discriminant analysis to constitute homogenous groups of firms before to compare their performances.[21] For instance, Rosenbaum and

[19] The arguments that follow rely on the literature on the determinants of R&D. See Cohen et Levin (1989) and Crépon, Duguet and Kabla (1996).

[20] The argument is symmetric for the effect on the non-treated so that we do not develop it here.

[21] On the link between discriminant analysis and the *logit model* see Maddala (1983).

Rubin (1985) compare the Mahalanobis distance method with the *logit* model. Finally the latter method became a standard due to its performance and its simplicity.

The reason why we do not work directly on the whole set of attributes is that it would involve a multidimensional matching that would rapidly become intractable. For instance, if we have 10 lagged size classes, 10 industries, 2 classes of past public support and 10 classes of lagged private R&D ratios, we would need to consider 10x20x2x10= 4000 possible cells and evaluate the effect of the treatment in each of it. The propensity score method reduces this multidimensional problem to matching on a real number, the probability to get the treatment. Moreover, this propensity score method relies on a very interesting intuition given below.

3.4 The propensity score

The propensity score method is based on the following result by Rosenbaum and Rubin (1983):

$$y(0) \perp T|X \Rightarrow y(0) \perp T|\Pr[T = 1|X].$$

Then, one need to compare firms that have the same probability to be treated, called the *propensity score*. The intuition of this method is the following: in an experimental sample we can compare the averages because the allocation of the treatment in random. Now suppose that this is not the case, and consider a group of firms that have the same probability to be treated. In this group, there are firms that have been treated and firms that have not been treated, hence the allocation of the treatment can be considered as random inside this group of firms. This implies that one can evaluate the causal effect inside this group by taking the difference of the average performances between the treated and the non-treated firms.

In practice, we need to match each treated firm with the non-treated firms that have the same probability to be treated and to compute the difference of their performances. Repeating this operation for all the treated firms and taking the average of these differences will provide an estimate of the effect of the treatment.

The only problem is whether we can find firms with the same propensity score when their treatment is different.[22] It is unlikely to be the case for all the values of the propensity score for the following reason: firms that have a strong probability to get the treatment will tend to be much more often treated, and the contrary for firms that have a small probability to be treated. Hence, the comparison will generally be impossible at both ends of the probability distribution. Only a part of the data will be useful for the comparison. Therefore, we will first look at the propensity score distributions among the treated and the non-treated and give the percentage of firms that belong to the common support of

[22] The absence of treatment can be considered as a specific treatment.

these distributions. In practice, it will never be far from 80% in this study so that the comparison is possible.

3.5 Selection of the attributes

We have a last problem to solve: what variables should be included in the attributes? It is clear that the only variables that can influence the evaluation are either the ones that determine the selection (*T*) or the ones that influence the performance (*y*). More precisely, we can restrict our attention to the intersection of these variables for the following reasons:

- The variables that do not influence the treatment (*T*) cannot influence the evaluation by definition since they create no selection and thus no selection bias.

- The variables that do not influence the treatment (*T*) but that do influence the performance (*y)* can also be discarded because they create no selection and therefore cannot create a selection bias as well.

- The variables that influence the treatment (*T*) but that do not influence the performance can also be discarded. They create a selection but since they do not influence the distribution of the performance, there is a selection without a selection bias.

- The attributes (*X*) must be chosen in the common determinants of the treatment (*T*) and of the performance (*y*). These determinants can be found in the theoretical literature on R&D investment and from an analysis of the subsidy granting process.

- It is also clear that the attributes (*X*) may never depend on the treatment (*T*). This is a reason why using lagged variables can be useful.

A simple rule for the selection of the conditioning variables is to start from the determinants of the performance, that we can identify from the theoretical and applied literatures, and then to eliminate the variables that are not significant in the attribution of the treatment. This last step can be made from the estimation of a *logit* model explaining the allocation of the treatment.

3.6 Evaluation method

There are many ways to match firms.[23] In this paper, following Heckman, Ichimura and Todd (1998), we use the Nadaraya-Watson (non-parametric) estimator of the average effect. For each treated firm, we compute the difference between its performance and a

---

[23] See Heckman and Hotz (1989), Heckman, Ichimura and Todd (1997).

local weighted average of the performances of its non-treated neighbors, where the weight decreases with the difference of the propensity score.

The kernel estimator of $E[y_i(0)|T=1]$ is defined as:

$$\hat{E}[y_i(0)|T=1] = \sum_{j \in I_0} \omega_j \times y_j, \quad i \in I_0 \text{ with } \omega_j = \frac{K[(p_i - p_j)/h]}{\sum_{j \in I_1} K[(p_i - p_j)/h]}, \quad j \in I_0$$

where p_i is the propensity score of the (treated) firm *i,* p_j the propensity score of the (non-treated) firm *j*, K(*x*) is a gaussian kernel, *h* the window and I_0 the set of firms that got no subsidy. [24]

The average causal effect on the treated is obtained by:

$$\hat{c}_1 = \frac{1}{N_1} \sum_{i \in I_1} \left\{ y_i(1) - \hat{E}[y_i(0)|T=1] \right\},$$

similarly the effects on the non-treated and on the whole population are estimated respectively by:

$$\hat{c}_0 = \frac{1}{N_0} \sum_{i \in I_0} \left\{ \hat{E}[y_i(1)|T=0] - y_i(0) \right\}$$

and $$\hat{c} = \frac{1}{N} \left(\sum_{i \in I_0} \left\{ \hat{E}[y_i(1)|T=0] - y_i(0) \right\} + \sum_{i \in I_1} \left\{ y_i(1) - \hat{E}[y_i(0)|T=1] \right\} \right).$$

These estimators are asymptotically normal and their variances are obtained by the bootstrap. We use the method of Andrews and Buchinsky (2000) in order to determine the number of bootstrap repetitions (see appendix 3).[25] Notice that the logit model is re-estimated at each simulation. The common support of the propensity score is also determined at each simulation, and is defined by the intersection of the probabilities intervals of the treated and of the non-treated firms, defined by their first and 99th percentiles. The estimation was performed using SAS-Logistic procedure and SAS-IML.

---

[24] In order to make this estimation we took a Silverman window. For more information about this non-parametric estimation method, see Härdle (1990).

[25] On the bootstrap, see Efron and Tibshirani (1993).

4 Results

4.1 The logit model

The first step of the evaluation method is to choose the attributes. We take the variables that influence both the probability to get public support and the investment in private R&D. Five determinants are significant.[26]

1. *Line of business.* On the one hand, this variable summarizes the conditions of demand, technology and appropriation that determine R&D investments. On the other hand, it influences the attribution of public support because the policy technology objectives are correlated to the lines of business that use these technologies.

2. *Size (lagged).* On the one hand, the proportionality of R&D investments to size is a stylized fact with theoretical support (Cohen and Klepper, 1996). On the other hand, it is likely that large size firms more often get subsidies than the small ones.

3. *The private R&D to sales ratio (lagged).* On the one hand, it influences current R&D investment because the research programs last over several years before completion. On the other hand, it influences the granting of subsidies because the firms that do the more R&D are the ones that are the most likely to apply for subsidies. An additional argument is that we study the growth rate of private R&D so that one needs to include its lagged value among the regressors in order to obtain estimates that are robust to fixed effects (see Crépon and Iung, 1999, appendix 2).

4. *The Debt to Sales ratio (lagged).* On the one hand, it influences R&D investment through the financial constraint and, on the other hand, it influences subsidy applications by the firms or the granting of the subsidies by the ministry.

5. *Past public support.* First, it influences the private R&D investments since the research is made on several years. Second, the public support can be granted on several years. Last, it summarizes unobservable characteristics linked to the ability of the firm to get public support.[27] We use two variables for the past public support: first, a dummy variable that equals one when the firm got a subsidy at least once in the past. Second, among the firms that had a subsidy, we have computed the

[26] For a justification and regressions about the determinants of R&D and innovation on French data, see Crépon, Duguet and Kabla (1996).

[27] Using the past values of public support is sometimes advocated to account for firm-level fixed effects (see Geroski, Van Reenen and Walters, 2002, p. 37). Here, we have few entry and exit so that the estimation of a fixed effect logit model is not possible. Moreover, with that method we would need to impose that the coefficients are stable over time.

average subsidy rate. These two variables therefore control for both the fact to get a subsidy and for its importance in the total R&D budget.[28]

All the explanative variables are lagged in order to avoid simultaneity issues. The results are presented in the Tables 4 and 5. Overall we find that the probability to get a subsidy:

1. Increases with size measured by sales;

2. Increases with the private R&D to Sales ratio;

3. Increases with the Debt to Sales ratio, for most years;

4. Increases with the existence and importance of a past public support;

5. Varies with the line of business.

Globally, the predictions of the logit model are good, from 85% to 90% of concordant predictions. The latter figures strongly depend on the introduction of the past public support among the regressors. Without it, the percentage of concordant prediction falls to 65%. Therefore, including past public support is likely to influence the quality of the matching. This conclusion is close to the analysis of Lach (2000).

Our results are interesting at the light of the theoretical model. The fact that the debt to sales ratio increases the probability to get a subsidy suggests that the firms that have a stronger financial constraint are more likely to be supported. This could explain the result of the comparison between the supported and the non-supported firms (Table 3), where only one difference out of twelve is significant.

The second interesting result is on the relationship between the past private R&D investment and the probability to get a subsidy. We find that the higher the private R&D ratio the higher the probability to get a subsidy. This validates our assumption that the ministries account for the part of the R&D project that is privately funded when they make their decisions. This result goes in the same direction than the first one and could also explain why our sample statistics reveal few substitution effects.

Finally, we also find that size significantly increases the access to public support. This shows that the access to subsidies would be less open than the access to fiscal measures (like the R&D tax credit).[29] It is also likely that the large firms face better appropriability conditions so that this type of selection would in fact let substitution effects appear. However, in the case of the large firms, it is possible that international competition reduces their market power.

---

[28] We have also examined what happens when we use one lag only. It did not affect our results.

[29] One could reply that when the tax credit is attributed on the basis of the growth of R&D expenditures, the large firms may be at a disadvantage.

Overall, we find that the allocation of subsidies includes some mechanism that tends to reduce the substitution effects but that there remains some room for crowding-out effects.

4.2 The effect of subsidies on privately funded R&D

The estimations by matching are presented in Table 6. We estimate three quantities: the global effect of subsidies c, that is the effect the subsidies would have if they were given to all the firms in the sample; the effect on the treated c_1 which is the evaluation strictly speaking and the effect on the non treated c_0 that represents the effect that the subsidies would have if they had been given to the firms that did not have public support.

The effect of subsidies is indeed heterogeneous, as the previous studies in the field already suggested. In our case, it means that the effect on the treated is sometimes different from the effect on the non-treated. For the twelve couples of years under study, the subsidies have no effect on the proportion of firms that have maintained or increased their private R&D expenditures (at the 5% level). Only one effect is negative at the 10% level in 1987. Therefore we find an addition effect. The picture is different for the effect on the firms that did not have a subsidy: it is negative for five years (four year at the 5% level) therefore suggesting that the subsidized firms have been well selected. As we saw, the determinants of the subsidy grant may explain this results: firm with high debt to sales ratio and high private R&D ratios are more often selected that the others, so that the effect of the subsidies may be stronger on the treated.

We can therefore reach the following conclusion: the global effect is negative for four years out of twelve but this is not because the subsidies would have crowd-out private R&D, it is because the global effect is a weighted average of the effect on the subsidized firms and on the unsubsidized firms and that the latter is negative. In other words, except for 1987, the negative sign of the global effect always comes from a potential effect on the firms that did not get a subsidy.

Our second performance measure is the growth rate of private R&D. Here there is less heterogeneity between the treated and the non-treated. We find that the subsidies have in general no effect on the growth rate of private R&D. The only exception, at the 5% level, is in 1987 were we find again a negative effect of subsidies on the private funding of R&D.

Overall we find that the subsidies add to private R&D. We should then ask why there is a substitution effect in 1987? The answer may well be provided by Figure 4. Year 1987 is the one where the average subsidy was the highest over the whole period 1985-1997, therefore some crowding-out is likely to have occurred from too much “generosity”.

4.3 Extensions

Three extensions of this work have been made. The first is to examine the growth rate of the R&D to sales ratio. This ratio allows correcting for the variations of R&D that would be linked to size. However, in this study we match firms on size so that we should expect that nothing changes. This is indeed what we find in appendix 2 (table A.3).

The two other extensions are presented in Duguet (2002b). The first one is methodological: we implement the weighted estimator of the causal effect proposed in Crépon and Iung (1999). We find that its variance is higher so that no effect is significant at the 5% level. However we reach the same conclusion if we perform the significance tests at the 10% level.

Last, we have compared the subsidies of the Ministry of Defense with the other subsidies. This needs an extension of our methodology to three treatments (see Imbens, 1999; Lechner 2001). On the one hand, nearly all firms that have a subsidy from the Ministry of Defense have a subsidy from another ministry, so that we can only study the incremental effect of the Ministry of Defense subsidies. On the other hand, we find no difference between the firms that benefit from Defense subsidies and the other firms, except in 1994. We thus face the same problem as in 1987, so that we have examined the average Defense subsidy in 1994: without surprise, it was the highest. This confirms that above some threshold, there would be a substitution between public and private R&D.

Conclusion

The subsidies to private R&D have been steadily decreasing in France over the 1990's, which raises the issue of the effect of this decrease. Studying this issue can be made by studying the effect that R&D subsidies have on the private funding of research, because it is the basis of any subsequent effect of innovation policies on firm performance. If there were substitution between subsidies and private funds, the decrease of subsidies would not affect firm performance and conversely, if subsidies and private funds were complementary, the reduction of subsidies would decrease the total R&D effort and therefore the number of innovations and firms performance.

We first examine the granting of the subsidies. We find that a firm as higher chances to get a subsidy if it has a high debt to sales ratio, a high private R&D to sales ratio and a higher size. The access to subsidies is therefore more restricted than the access to R&D tax credit, however the accent that is put on the financial situation of the firm and on its involvement in private R&D suggest that there exist in the subsidy system some mechanism that try to prevent the substitution between the public and the private funds.

In a second step, we examine the effect of subsidies on private R&D with matching methods. We find that, at the exception of 1987, no significant substitution effect appears. The most probable reason why 1987 exhibits a substitution is that this is the year where the average subsidy was the highest (over 1985-1997). This result can however be obtained only if we compute the effect of subsidies on the firms that got it (i.e., on the treated). We global effect (over the whole sample) is sometimes found to be negative, but this result comes from a potential effects of subsidies on the firms that did not get one (i.e., on the non-treated). The methodology is therefore important.

Overall, we find that R&D subsidies adds to private R&D. But such an effect implies that the reduction of subsidies is unlikely to have been compensated by inverse movements of private funds, so that a lower innovation-related performance cannot be ignored. Of course, this conclusion would hold only if other sources of funding, private or public, would not be promoted.

References

Andrews D. and M. Buchinsky (2000). A three-step method for choosing the number of bootstrap repetitions. Econometrica, 67, 23-51.

Andrews D. and M. Buchinsky (2001). Evaluation of a three-step method for choosing the number of bootstrap repetitions. Journal of Econometrics, 103, 345-386.

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In R. Nelson ed., *The rate and direction of economic activity*, New York: Princeton University Press.

Arvanitis S. and H. Hollenstein (2002). The impact of spillovers and knowledge heterogeneity of firm performance: evidence from Swiss manufacturing. In Kleinknecht A. and P. Mohnen eds, *Innovation and Firm Performance : Econometric Explorations of Survey Data.* Palgrave (ISBN 0-333-96109-9), 225-252.

Brodaty T., B. Crépon and D. Fougère (2002). « Développements récents dans l'évaluation des politiques économiques ». *Mimeo* CREST. Available on the web page of B. Crépon (www.crest.fr).

Busom I. (1999). An empirical evaluation of the effects of R&D subsidies. Working paper, University of California at Berkeley, Burch Working Paper n°B99-05.

Cohen W. and S. Klepper, [1996], "A reprise of size and R&D", Economic Journal, 106, 925-951.

Crampes C. (1986). Les inconvénients d'un dépôt de brevet pour une entreprise innovatrice. *L'Actualité Economique*, vol. 62, 521-534.

Crépon B. and E. Duguet (1996). Innovation : measurement, returns and competition. *Insee studies in Economics and Statistics,* 1, 82-95.

Crépon B., Duguet E. and I . Kabla I. (1996). A moderate support to Schumpeterian conjectures from various innovation measures. In A. Kleinknecht éd*., Determinants of Innovation: the message from new indicators* (chapitre 3), Mc Millan Press, Londres.

Crépon B., E. Duguet and J. Mairesse (1998). Research and development, innovation and productivity : an econometric analysis at the firm level*. Economics of Innovation and New Technology, 7(2).*

Crépon B. and N. Iung (1999). Innovation, emploi et performances. INSEE Working paper G9904.

Czarnitzki D. and A. Fier (2001). Do R&D subsidies matter ? Evidence for the German service sector. Document de travail du ZEW, n°01-19.

Czarnitzki D. and A. Fier (2002). Substitutive or Complementary ? Innovation subidies in the German service sector. Document de travail du ZEW, n°02-04.

David P, B. Hall and A. Toole (2000). Is public R&D a complement or a substitute for private R&D ? A review of the econometric evidence. *Research Policy,* vol. 29, 497-529.

Duguet E. (2002a). Innovation height, spillovers and TFP growth at the firm level : Evidence from French manufacturing . To appear in *Economics of Innovation and New Technology*, vol. 13(1-2). Working paper, Université de Paris I, Cahiers de la MSE, EUREQua 2002.73.

Duguet E. (2002b). Les subventions à la recherche et développement sont-elles substituables ou complémentaires à leur financement privé? Une étude microéconométrique. Working paper, Université de Paris I, Cahiers de la MSE, EUREQua, forthcoming in December.

Duguet E. and I. Kabla (1998). Appropriation strategy and the motivations to use the patent system in France : an econometric analysis at the firm level. *Annales d'Economie et de Statistique*, 49-50.

Efron B. and J. Tibshirani (1993). *An introduction to the bootstrap*. Monographs on Statistics and Applied Probability n°57. Chapman & Hall (ISBN 0-412-04231-2).

Gouriéroux C. (2000). Econometrics of Qualitative Dependent Variables. Cambridge University Press (ISBN 0521589851).

Geroski P., J. van Reenen and C. Walters (2002). Innovations, patents and cash flow. In Kleinknecht A. and P. Mohnen eds, *Innovation and Firm Performance : Econometric Explorations of Survey Data.* Palgrave (ISBN 0-333-96109-9), 31-55.

Griliches Z. (1998). R&D and productivity : the econometric evidence. NBER monographs.

Hall B. H. and J. Mairesse (1995). Exploring the relationship between R&D and productivity in French manufacturing firms, 65(1), 263-293.

Hall B. and J. Van Reenen (2000). How effective are fiscal incentives for R&D ? A review of the evidence. *Research Policy,* 29, 449-469.

Härdle W. (1990). Applied non-parametric regression. Econometric Society Monograph 19, Cambridge University Press (ISBN 0-521-38248-3).

Heckman J. and J. Hotz (1989). Choosing among alternative nonexperimental methods for estimating the impact of social programs : the case of manpower training. *Journal of the American Statistical Society,* vol. 84(408), 862-874.

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

Heckman J., J. Lalonde and J. Smith (1999). The economics and econometrics of active labor market programs. In Ashenfelter and Card eds, *Handbook of Labor Economics*, vol. 3, Elsevier.

Imbens G. (1999). The role of the propensity score in estimating dose-response functions. *NBER Technical Working Paper,* n°237.

Jefferson G., Huamao B., Xiaojing G. and Y. Xiayung (2001). R&D performance in Chinese industry. Working paper, proceedings of the NBER Productivity workshop.

Lach S. (2000). Do R&D subsidies stimulate or displace private R&D ? Evidence from Israël. NBER, Working paper 7943.

Lechner M. (2001). Identification and estimation of causal effets of multiple treatment under the conditional independance asumption. *In* Lechner M. and F. Pfeiffer eds, *Econometric Evaluation of Labour Market Policies.* Physica-Verlag (ISBN 3-7908-1372-9).

Levin R., A. Klevorick, R. Nelson and S. Winter, 1987, Appropriating the returns from industrial research and development. *Brookings Papers on Economic Activity* 3, 783-831.

Lööf H. and A. Heshmati (2001). On the relationship between innovation and performance: a sensitivity analysis. Swedish working paper series in economics and finance, 2001:446, Stockholm school in economics.

Lööf H., Heshmati A., Asplund R. and S. Nääs (2001). Innovation and performance in manufacturing industries: a comparison of the Nordic countries. Swedish working paper series in economics and finance, 2001:457, Stockholm school in economics.

Maddala G. S. (1983). Limited dependent and qualitative variables in econometrics. Econometric Society Monograph n°3, Cambridge University Press (ISBN 0-521-33825-5).

Mansfield E., M. Schwartz and S. Wagner, 1981, Imitation costs and patents : an empirical study. *Economic Journal* 91, 907-918.

Nash J., 1953. Two-person cooperative games. *Econometrica*, vol. 21, 128-140.

Nelson R. (1959), The Simple Economics of Basic Scientific Research, Journal of Political Economy, 67: 297-306.

Rosenbaum P. and D. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*; vol. 70, 41-55.

Rosenbaum P. and D. Rubin (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*; vol. 39, n°1, 33-38.

Rubin D. (1974). Estimating causal effects of treatements in randomized and non-randomized studies. *Journal of Educational Psychology,* 66, 688-701.

Rubin D. (1979). Using multivariate matched sampling and regression adjustment to control bias in observational studies. *Journal of the American Statistical Association,* 74(366), 318-328.

Rubin D. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine,* Part 2, n°127, 757-763.

Van Leeuwen G. and L. Klomp (2001). On the contribution of innovation to muti-factor productivity growth. Working paper, Statistics Netherlands.

Wallsten S. (2000). The effect of governement-industry R&D programs on private R&D : the case of Small Business Innovation Research program. *RAND Journal of Economics,* vol. 31(1), 82-100.

Appendix 1: The data

We work on two-year datasets. The reason is that we want to avoid one problem associated with mergers or when a firm sells one of its departments. The capital per capita plays an important role on the discarding of firms when cleaning a dataset. The reason may be that buyouts or selling a part of the firm more often affect R&D firms than the others. However, our method drops these firms only during the capital modifications because our two-year samples are made independently for each couple year. For instance, a firm that would make a buyout on year *t* would be discarded from the sample at times *t –1* and t but would be reintegrated in the next sample on years *t* from *t+1*. This is why we have much more firms than in a balanced panel. If we had a balanced panel, we would have about 300 firms only on the period 1985-97. With our method we end up with three to four more times firms depending on the couple of years studied.

<u>Cleaning on the annual datasets</u> (Table A.1)

The samples are obtained from the merger between the accounting data (BRN, INSEE) and the Research Survey (Ministère de la Recherche). We first eliminate the outlyiers on the logarithms of the following per capita variables and then on the growth rates of the same variables:

- Production / Personnel (BRN)
- Value Added / Personnel (BRN)
- Capital / Personnel (BRN)
- Total R&D expenditures / Personnel (R&D survey)
- R&D Personnel / Personnel (R&D survey)

Our cleaning method can be decomposed in two steps. In a first step, we sort the data according to the variable on which we make the cleaning, in a second step we compute the gap between each value and its closest neighbor located after it. If the gap is too important, all the point located beyond the value are discarded. The process starts from the median first in increasing order and then in decreasing order. This method has the advantage to be efficient when the distribution studied is asymmetric because it does not imply to put the same limit at the left and at the right of the mean or the median. In a second step, we check the effect of the cleaning on the distribution graphically in order to avoid any truncation due to the method. After these graphical checks we have decided to discard the points when the difference between to subsequent value is 0.2 which means 20% since our variables are in logarithms. Even though this condition looks tolerant, it leads to keep between the two thirds and the three fourths of our sample depending on the year.

Cleaning on the two-year datasets

We apply the same methodology to the growth rates (difference of logarithms) of the same variables. Here two types of losses are possible: First, the ones that come from the mergers themselves and, second, the ones that come from the cleaning. The loss from the merger mostly comes from entry and exit and the fact that some firms in the R&D survey are surveyed every two years only. We have examined the latter case by constructing databases with a two years lag instead of a one-year lag and it very slightly affects the number of firms in the sample. This is because this different convention imposes three years of presence in the dataset for the majority of firms instead of two, so that the gains are compensated by the losses. Therefore we kept the convention to keep firms with two consecutive years of presence in the dataset.

The cleaning threshold for the growth rates is 0.1, that is 10% between two consecutive growth rates. The losses linked to the merger are between 30% to 40%. The losses from the cleaning of the growth rates, that intervenes after the cleaning on logarithm of the ratios and a merger, allows keeping almost all the firms in the sample. The latter result clearly comes from the first cleaning.

Notice that during the merger no loss is linked to the identification of the firms, since in France all firms have a national identification number (the SIREN code) and its use is compulsory for all the relationships with the French administration.

Table A.1: Cleaning on the logarithms of the per capita variables (annual datasets)

Year	Source[a]	After cleaning	% remaining
1985	1730	1340	77.5
1986	1867	1442	77.2
1987	2034	1577	77.5
1988	2404	1785	74.3
1989	2698	1905	70.6
1990	2769	1913	69.1
1991	2805	1990	70.9
1992	2677	2060	77.0
1993	3262	2526	77.4
1994	3468	2610	75.3
1995	3347	2565	76.6
1996	3362	2654	78.9
1997	3091	2409	77.9

a. Merger of the R&D survey with the BRN accounting data.

Table A.2 : Merger and cleaning on the growth rates (two-year datasets)

This cleaning is made from the one-year datasets.

Year	Merger with the previous year	% remaining	Cleaning	% remaining	% after mergers and cleaning
1986	1114	77.3	1032	92.6	55.3
1987	1197	75.9	1126	94.1	55.4
1988	1299	72.8	1236	95.2	51.4
1989	1304	68.5	1230	94.3	45.6
1990	1399	73.1	1316	94.1	47.5
1991	1484	74.6	1407	94.8	50.2
1992	1459	70.8	1396	95.7	52.1
1993	1522	60.3	1445	94.9	44.3
1994	1692	64.8	1633	96.5	47.1
1995	1753	68.3	1665	95.0	49.7
1996	1752	66.0	1672	95.4	49.7
1997	1676	69.6	1618	96.5	52.3

Appendix 2: Effect of subsidies on the R&D to sales ratio

Table A.3 : Estimates on the R&D to sales ratio

Nadaraya-Watson estimates of the causal effect. The estimation uses a gaussian kernel with a Silverman window. The standard errors are computed by the bootstrap with a number of repetitions following Andrews and Buchinsky (see appendix 3).
* significant at the 10% level; ** significant at the 5% level.

Output variables	Dummy = 1 for firms that have increased their R&D to sales ratio (0 otherwise)			Growth rate of the R&D to sales ratio		
Estimates	Effect on the treated (c_1)	Effect on the non-treated (c_0)	Global effect (c)	Effect on the treated (c_1)	Effect on the non-treated (c_0)	Global effect (c)
1986	-0.036 (0.068)	-0.091 (0.062)	-0.076 (0.055)	-0.036 (0.053)	-0.047 (0.046)	-0.044 (0.041)
1987	-0.097* (0.051)	-0.059 (0.067)	-0.068 (0.057)	-0.113** (0.038)	-0.062 (0.047)	-0.075* (0.041)
1988	0.015 (0.041)	0.028 (0.048)	0.025 (0.042)	-0.032 (0.028)	-0.016 (0.028)	-0.020 (0.025)
1989	-0.056 (0.045)	-0.040 (0.058)	-0.045 (0.048)	-0.037 (0.036)	-0.090* (0.051)	-0.073* (0.042)
1990	0.046 (0.038)	0.068 (0.046)	0.061 (0.038)	0.025 (0.029)	0.004 (0.039)	0.011 (0.032)
1991	-0.001 (0.038)	0.017 (0.046)	0.011 (0.040)	0.027 (0.028)	0.019 (0.038)	0.022 (0.032)
1992	-0.036 (0.035)	-0.046 (0.053)	-0.043 (0.042)	-0.031 (0.029)	-0.018 (0.050)	-0.022 (0.039)
1993	-0.027 (0.033)	-0.073* (0.043)	-0.058 (0.035)	-0.007 (0.031)	-0.040 (0.042)	-0.029 (0.035)
1994	0.050 (0.032)	0.065 (0.046)	0.060 (0.037)	0.041 (0.025)	0.060 (0.042)	0.054 (0.033)
1995	0.034 (0.034)	-0.002 (0.057)	0.008 (0.047)	0.013 (0.024)	-0.004 (0.046)	0.001 (0.037)
1996	0.005 (0.034)	-0.060 (0.044)	-0.042 (0.038)	-0.028 (0.024)	-0.107** (0.036)	-0.085** (0.031)
1997	0.026 (0.034)	-0.002 (0.053)	0.006 (0.044)	0.035 (0.026)	-0.018 (0.041)	-0.004 (0.034)

Appendix 3: Number of bootstrap repetitions

We determine the number of bootstrap repetitions by the method of Andrews and Buchinsky (2000, 2001) with, in the authors' notation, $(\text{pdb}, \tau) = (10, 0.05)$. Since we have 12 quantities to evaluate on each year, we take the maximum number of simulations of each year, denoted B*. The following table gives the detail.

Table A.4: Selecting the number of bootstrap repetitions

Year	Effect	Dummy=1 if private R&D increased	Growth rate of private R&D	Dummy=1 if private R&D / sales increased	Growth rate of the private R&D to sales ratio
1986	C	145	99	136	113
B* = 170	C0	170	95	143	124
	C1	122	91	106	92
1987	C	149	109	137	125
B* = 156	C0	137	120	131	134
	C1	156	106	112	132
1988	C	152	132	140	163
B* = 163	C0	145	135	148	162
	C1	90	131	117	163
1989	C	272	264	198	248
B* = 276	C0	276	274	204	233
	C1	122	155	122	168
1990	C	150	140	92	123
B* = 180	C0	157	139	108	117
	C1	165	141	98	180
1991	C	127	116	95	98
B* = 143	C0	143	115	94	104
	C1	113	106	110	104
1992	C	153	197	163	139
B* = 197	C0	166	181	169	143
	C1	123	168	102	123
1993	C	109	157	159	145
B* = 220	C0	121	152	144	123
	C1	96	130	212	220
1994	C	202	196	163	153
B* = 232	C0	232	193	173	142
	C1	118	143	136	153
1995	C	122	127	168	164
B* = 191	C0	124	119	179	152
	C1	143	191	96	173
1996	C	125	127	195	139
B* = 195	C0	133	137	195	146
	C1	110	121	97	92
1997	C	152	148	108	112
B* = 162	C0	162	149	103	115
	C1	105	155	147	127

Figure 1 : Total R&D expenditures of the sample (millions of Euros)

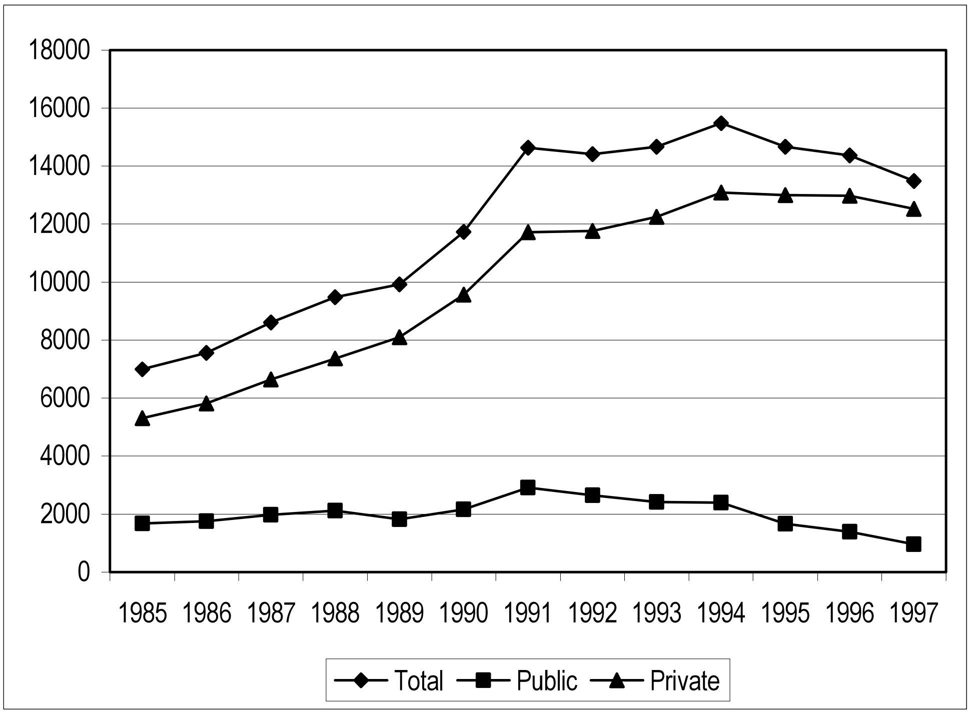

Figure 2 : Average R&D expenditures of the sample (millions of Euros)

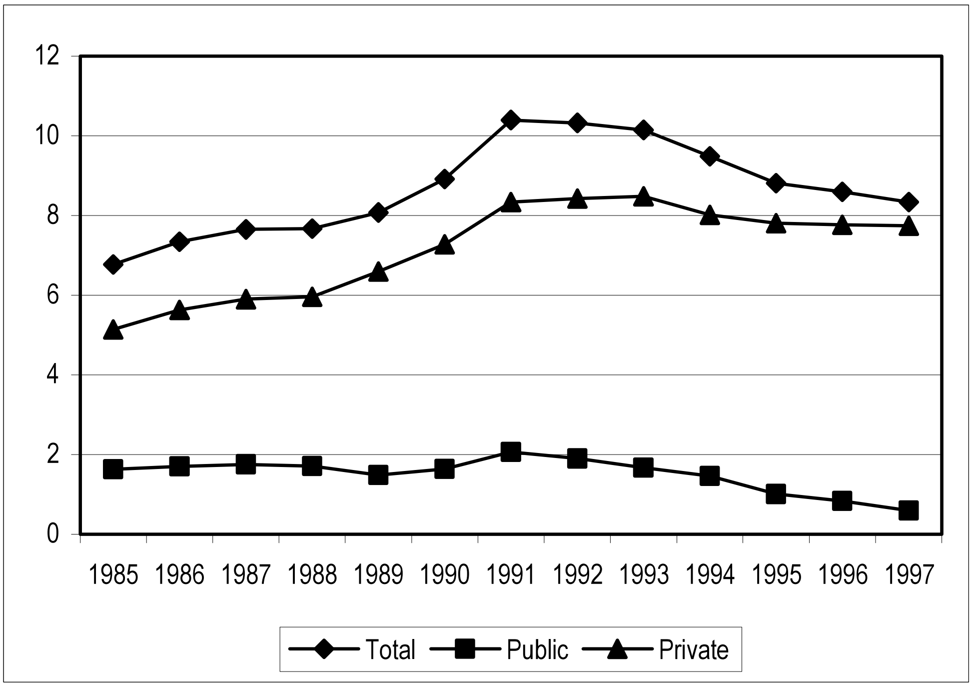

Figure 3: Average R&D subsidy of the sample (thousands of Euros)

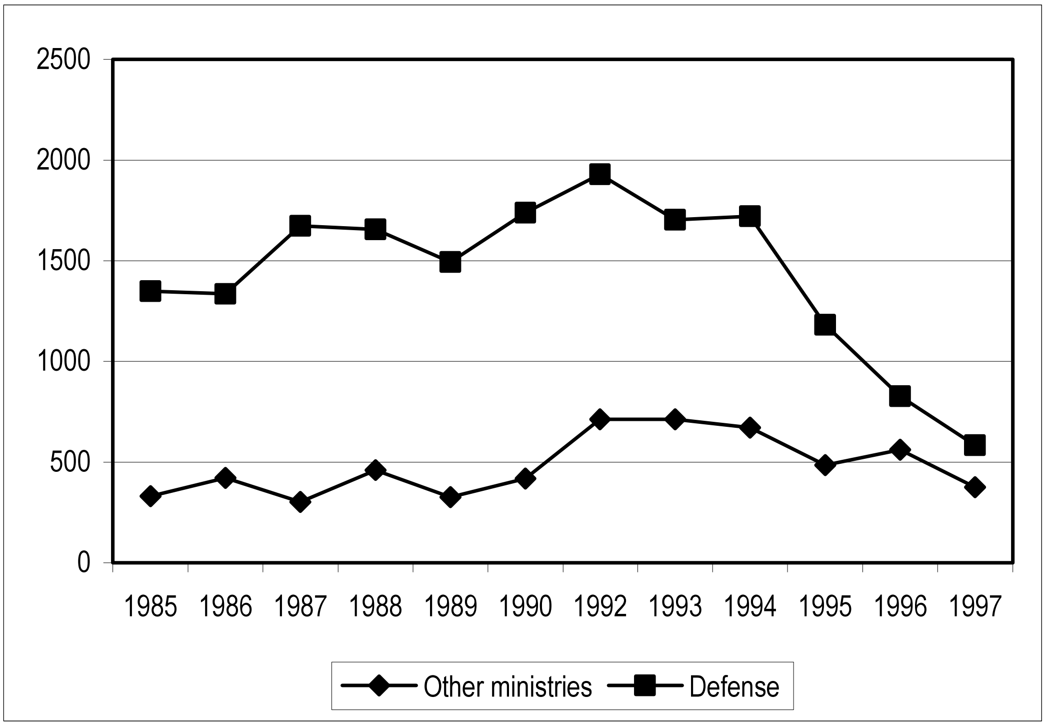

Figure 4 : Average R&D subsidy by recipient (millions of Euros)

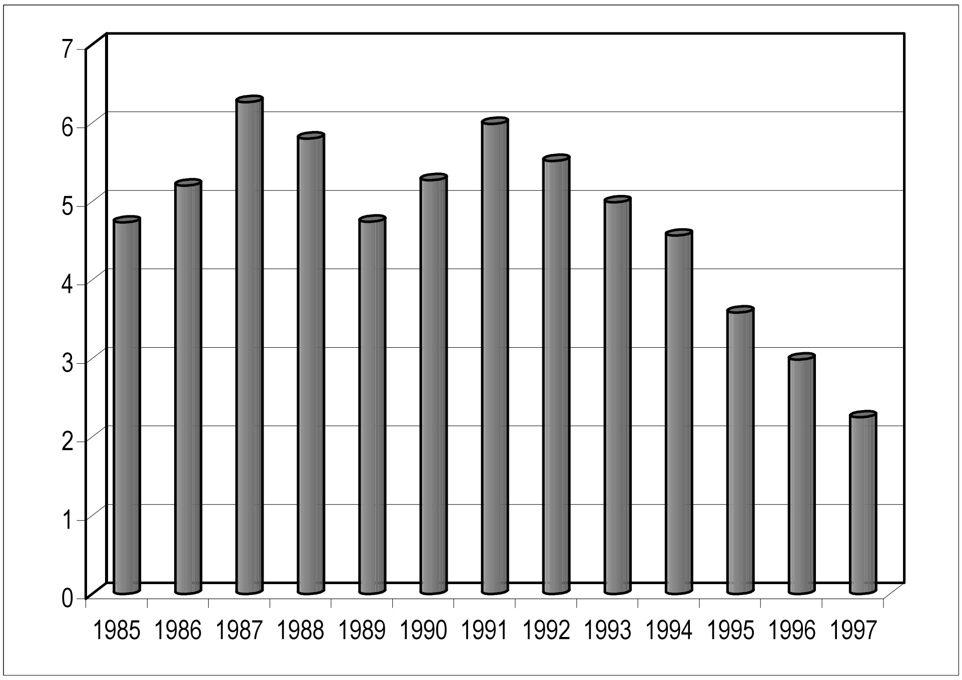

Table 1: Sample Totals
(in Euros)

Variable	Number of firms	Sales	Total R&D	R&D subsidies	Firms with a subsidy	R&D / Sales	R&D subsidies/ Total R&D
Unit	×	Billions	Millions	Millions	%	%	%
1985	1032	158	6985	1680	34.4	4.4	24.0
1986	1032	159	7567	1756	32.7	4.7	23.2
1987	1126	168	8610	1975	28.0	5.1	22.9
1988	1236	174	9483	2115	29.5	5.4	22.3
1989	1230	190	9923	1820	31.2	5.2	18.3
1990	1316	214	11727	2157	31.1	5.5	18.4
1991	1407	254	14623	2907	34.5	5.8	19.9
1992	1396	267	14404	2641	34.3	5.4	18.3
1993	1445	258	14663	2415	33.5	5.7	16.5
1994	1633	281	15477	2393	32.1	5.5	15.5
1995	1665	292	14659	1666	27.9	5.0	11.4
1996	1672	293	14361	1388	27.8	4.9	9.7
1997	1618	297	13476	958	26.3	4.5	7.1

Table 2: Entry and Exit from subsidies

Each line sums up to 100%
Entry: without subsidy on the previous year and with a subsidy on the current year.
Exit: with a subsidy on the previous year and without a subsidy on the current year.

Firms	Without subsidies for two consecutive years	Entry	Exit	With subsidies for two consecutive years
1986	59.6	5.6	7.2	27.6
1987	62.5	5.5	8.7	23.3
1988	63.4	7.7	7.0	21.9
1989	61.8	8.0	6.4	23.7
1990	61.3	7.2	8.0	23.5
1991	59.1	8.7	6.2	26.0
1992	56.6	7.0	9.0	27.4
1993	56.8	7.5	9.4	26.3
1994	58.6	7.1	8.1	26.2
1995	62.3	5.0	8.0	24.7
1996	63.5	6.2	7.3	23.0
1997	65.2	5.3	7.2	22.3

Table 3 : Averages Comparison

The *t* statistics are computed by the bootstrap on 100 repetitions.

Variables	Percentage of firms that have maintained or increased their private R&D			Average growth rate of private R&D expenditures		
Firms	Unsubsidized	Subsidized	Student	Unsubsidized	Subsidized	Student
1986	72.4	66.8	1.70	7.5	7.6	0.05
1987	74.4	62.9	3.77	10.6	5.2	2.67
1988	74.9	67.9	2.38	9.7	7.3	1.26
1989	74.2	66.4	2.96	10.1	8.3	1.11
1990	71.2	68.2	1.00	7.8	7.9	0.04
1991	71.0	66.4	1.65	7.0	9.1	1.05
1992	61.8	54.3	2.81	-0.1	-2.5	0.48
1993	56.9	50.6	2.25	2.6	-0.2	1.17
1994	58.4	54.7	1.24	1.7	2.1	0.20
1995	61.8	60.8	0.44	5.5	6.3	0.44
1996	60.0	56.0	1.29	2.2	-0.7	1.78
1997	64.0	61.0	0.33	2.8	3.7	0.50

Table 4 : Determinants of subsidy grants 1986-1992

Dependent variable : Subsidy dummy on the current year (0/1)
Maximum likelihood estimation of the logit model (asymptotic standard errors between parentheses). The industry classification is the NAP 1973. It changed between 1992 and 1993. The significant industry dummies only have been included in the final regressions.

Lagged explanative variables	1986	1987	1988	1989	1990	1991	1992
Intercept	-3.54 (0.85)	-3.10 (0.78)	-2.36 (0.62)	-2.36 (0.64)	-2.86 (0.57)	-3.34 (0.57)	-3.45 (0.59)
Past subsidy dummy (0/1)	4.41 (0.33)	3.41 (0.29)	2.85 (0.25)	3.05 (0.25)	2.74 (0.24)	2.85 (0.24)	2.87 (0.24)
Average of the logarithms of past subsidy rates	0.33 (0.09)	0.21 (0.08)	0.17 (0.07)	0.11 (0.07)	0.13 (0.06)	0.13 (0.06)	0.13 (0.06)
Log Private R&D / Sales	0.42 (0.09)	0.56 (0.09)	0.52 (0.07)	0.44 (0.07)	0.39 (0.07)	0.51 (0.07)	0.42 (0.06)
Log Debt / Sales	0.24 (0.23)	0.41 (0.20)	0.39 (0.18)	0.42 (0.17)	0.06 (0.15)	0.26 (0.15)	0.39 (0.14)
Log Sales	0.25 (0.07)	0.27 (0.06)	0.19 (0.05)	0.16 (0.06)	0.17 (0.05)	0.26 (0.05)	0.24 (0.05)
T11 : Chemicals	×	-1.23 (0.42)	×	×	×	×	×
T12 : Pharmaceuticals	-0.53 (0.30)	-1.11 (0.31)	-0.76 (0.26)	-0.73 (0.27)	-0.50 (0.24)	-0.60 (0.24)	-0.56 (0.24)
T13 : Metalworking	×	-0.74 (0.44)	×	×	×	0.60 (0.32)	×
T14 : non-electrical machinery	×	-0.50 (0.26)	×	-0.49 (0.22)	×	×	×
T15A : electrical machinery	×	×	×	×	×	-0.32 (0.18)	×
T16 : car industry	×	-1.20 (0.59)	×	×	×	×	×
T23 : rubber	-1.10 (0.58)	×	×	×	×	×	×
T33 : services to firms	×	-0.71 (0.37)	×	×	×	×	×
% Concordant predictions	89.9%	88.3%	85.1%	86.5%	81.8%	83.6%	83.0%
Common support bounds in %	4.4-87.6	2.9-78.8	3.0-76.5	4.5-79.6	5.2-74.3	4.9-77.0	4.3-77.9
Firms remaining for the matching %	80.6	81.2	89.8	81.3	91.0	86.5	89.5

Table 5 : Determinants of subsidy grants 1993-1997

Dependent variable : Subsidy dummy on the current year (0/1)
Maximum likelihood estimation of the logit model (asymptotic standard errors between parentheses). The industry classification is the NAF 1993. It allows for international comparisons. The significant industry dummies only have been included in the regression.

Lagged explanative variables	1993	1994	1995	1996	1997
Intercept	-2.74 (0.57)	-2.50 (0.52)	-3.60 (0.56)	-3.22 (0.56)	-2.66 (0.57)
Past subsidy dummy (0/1)	2.73 (0.23)	2.31 (0.22)	3.16 (0.24)	2.91 (0.24)	2.94 (0.25)
Average of the logarithms of past subsidy rates	0.15 (0.05)	0.02 (0.05)	0.15 (0.05)	0.16 (0.05)	0.17 (0.05)
Log Private R&D / Sales	0.43 (0.06)	0.44 (0.06)	0.35 (0.06)	0.37 (0.06)	0.25 (0.06)
Log Debt / Sales	0.49 (0.14)	0.37 (0.12)	0.44 (0.12)	0.62 (0.12)	0.57 (0.12)
Log Sales	0.19 (0.05)	0.16 (0.05)	0.19 (0.05)	0.18 (0.05)	0.11 (0.05)
B0 : food industry	×	×	×	0.80 (0.31)	×
C3 : pharmaceuticals	×	-0.73 (0.26)	×	×	×
E2 : non-electrical machinery	-0.40 (0.19)	×	-0.48 (0.19)	-0.37 (0.20)	-0.56 (0.20)
F4 : chemicals	×	×	×	×	-0.34 (0.21)
F6 : electronic components	×	×	×	0.44 (0.27)	×
Concordant predictions %	81.7%	81.3%	82.1%	81.5%	79.9%
Common support bounds in %	4.7-76.9	4.0-75.4	2.6-67.4	3.2-69.5	2.8-66.4
Firms remaining for the matching %	87.5	91.5	89.4	88.9	92.9

Table 6 : Matching evaluation of the effect of R&D Subsidies

Nadaraya-Watson estimates of the causal effect. The estimation uses a gaussian kernel with a Silverman window. The standard errors are computed by the bootstrap with a number of repetitions following Andrews and Buchinsky (see appendix 3).
* significant at the 10% level; ** significant at the 5% level.

Output variables	Dummy = 1 for firms that have increased their R&D to sales ratio (0 otherwise)			Growth rate of the R&D to sales ratio		
Estimates	Effect on the treated (c_1)	Effect on the non-treated (c_0)	Global effect (c)	Effect on the treated (c_1)	Effect on the non-treated (c_0)	Global effect (c)
1986	-0.052 (0.065)	-0.192** (0.058)	-0.153** (0.050)	-0.036 (0.053)	-0.052 (0.039)	-0.056 (0.036)
1987	-0.091* (0.052)	-0.042 (0.065)	-0.054 (0.056)	-0.101** (0.038)	-0.059 (0.047)	-0.069* (0.040)
1988	-0.048 (0.038)	-0.121** (0.051)	-0.102** (0.044)	-0.021 (0.025)	-0.032 (0.031)	-0.029 (0.027)
1989	-0.063 (0.041)	-0.151** (0.060)	-0.123** (0.049)	-0.019 (0.035)	-0.062 (0.049)	-0.048 (0.040)
1990	-0.004 (0.038)	-0.026 (0.047)	-0.019 (0.040)	0.009 (0.027)	0.005 (0.038)	0.007 (0.031)
1991	0.008 (0.034)	-0.064 (0.048)	-0.040 (0.040)	0.041 (0.027)	0.035 (0.043)	0.037 (0.035)
1992	-0.051 (0.037)	-0.122** (0.049)	-0.098** (0.040)	-0.031 (0.028)	-0.031 (0.051)	-0.031 (0.039)
1993	-0.045 (0.036)	-0.065 (0.045)	-0.058 (0.037)	-0.001 (0.030)	-0.032 (0.041)	-0.021 (0.033)
1994	-0.010 (0.034)	-0.006 (0.047)	-0.007 (0.038)	0.028 (0.023)	0.036 (0.034)	0.033 (0.028)
1995	0.003 (0.033)	-0.009 (0.055)	-0.006 (0.045)	0.006 (0.021)	-0.003 (0.042)	-0.001 (0.034)
1996	-0.029 (0.034)	-0.060 (0.046)	-0.051 (0.039)	-0.030 (0.023)	-0.076* (0.039)	-0.063** (0.032)
1997	0.005 (0.030)	-0.083 (0.052)	-0.060 (0.043)	0.026 (0.023)	-0.020 (0.038)	-0.008 (0.031)