Tuomas Takalo – Tanja Tanayama – Otto Toivanen

Evaluating innovation policy: a structural treatment effect model of R&D subsidies

SUOMEN PANKKI • FINLANDS BANK
EUROJÄRJESTELMÄ
EUROSYSTEMET

Bank of Finland Research
Discussion Papers
7 • 2008

Suomen Pankki
Bank of Finland
PO Box 160
FI-00101 HELSINKI
Finland
☎ +358 10 8311

http://www.bof.fi

Bank of Finland Research
Discussion Papers
7 • 2008

Tuomas Takalo* – Tanja Tanayama** – Otto Toivanen***

Evaluating innovation policy: a structural treatment effect model of R&D subsidies

* Corresponding author. E-mail: tuomas.takalo@bof.fi.
** HECER, University of Helsinki, E-mail: tanja.tanayama@helsinki.fi
*** HECER, University of Helsinki, E-mail: otto.toivanen@helsinki.fi

We would like to thank Tim Bresnahan, David Card, Chiara Criscuolo, Ken Chay, Stefanie Franzke, Ari Hyytinen, Guido Imbens, Saul Lach, Phil Leslie, Ariel Pakes, Jim Powell, Petri Rouvinen, Paul Ruud, Chris Shannon, George Siotis, Javier Suarez, John Van Reenen, Reinhilde Veugelers, Frank Wolak and numerous seminar audiences for discussions and suggestions. We are thankful for Tekes (the National Technology Agency of Finland) for funding and data, and for several Tekes employees for helpful comments and thorough discussions on the application and selection process. Takalo and Toivanen also thankfully acknowledge funding from the Yrjö Jahnsson Foundation. This paper was partly written while Takalo was visiting IDEI, University of Toulouse and Toivanen UC Berkeley, whose hospitality they gratefully acknowledge.

http://www.bof.fi

ISBN 978-952-462-430-5
ISSN 0785-3572
(print)

ISBN 978-952-462-431-2
ISSN 1456-6184
(online)

Helsinki 2008

Evaluating innovation policy: a structural treatment effect model of R&D subsidies

Bank of Finland Research
Discussion Papers 7/2008

Tuomas Takalo – Tanja Tanayama – Otto Toivanen
Monetary Policy and Research Department

Abstract

This paper studies the welfare effects of R&D subsidies. We develop a model of continuous optimal treatment with outcome heterogeneity where the treatment outcome depends on applicant investment. The model takes into account heterogeneous application costs and identifies the treatment effect on the public agency running the programme. Under the assumption of a welfare-maximizing agency, we identify general equilibrium treatment effects. Applyiing our model to R&D project-level data we find substantial treatment effect heterogeneity. Agency-specific treatment effects are smaller than private treatment effects. We find that the rate of return on subsidies for the agency is 30–50%.

Keywords: applications, effort, investment, R&D, selection, subsidies, treatment programme, treatment effects, welfare

JEL classification numbers: O38, O31, L53, C31

Kansallisen innovaatiopolitiikan arviointi rakenteellisen mallin avulla

Suomen Pankin keskustelualoitteita 7/2008

Tuomas Takalo – Tanja Tanayama – Otto Toivanen
Rahapolitiikka- ja tutkimusosasto

Tiivistelmä

Tutkimus- ja kehitystoiminta (T&K) on keskeisin talouskasvuun vaikuttava seikka, ja suorat T&K-tuet ovat erityisesti Suomessa tärkein politiikkainstrumentti, jolla T&K-toimintaan voidaan vaikuttaa. Tässä tutkimuksessa arvioidaan suorien T&K-tukiaisten hyvinvointivaikutuksia Suomessa. Tutkimuksessa sovelletaan rakenteellisia menetelmiä, joita on aiemmin käytetty lähinnä työmarkkinapolitiikan tutkimiseen. Työssä käytetyn mallin avulla voidaan yksilöidä T&K-tukien hyödyt sekä T&K-toimintaa harjoittaville yrityksille että tukia myöntävälle julkiselle organisaatiolle. Tässä tutkimuksessamme käytetään harvinaista ja yksityiskohtaista T&K-projektitason aineistoa. Päätulos on, että julkisen organisaation tuotot T&K-tukiohjelmalle ovat 30–50 % mutta julkiset hyödyt T&K-tuista ovat selvästi pienemmät kuin yritysten saamat hyödyt.

Avainsanat: talouskasvu, T&K-toiminta, innovaatiopolitiikka, T&K-tuet, rakenteellinen mallinnus

JEL-luokittelu: O38, O31, L53, C31

Contents

1 Introduction

Direct R&D subsidies are one of the most important innovation policy tools both in theory and in practice.[1] We model the functioning of an R&D subsidy program – the actions of firms applying for subsidies and of the agency deciding the subsidies. Taking our model to rarely available R&D project level data provides a welfare evaluation of the R&D subsidy program.

There is large literature on the determinants of R&D and the effects of R&D subsidies (Cohen, 1995, and David, Hall and Toole, 2000). While one of the main motivations of the latter literature is the welfare effect of R&D subsidies, the question has been addressed indirectly by studying the effects of subsidies on private R&D (referred to as additionality) or productivity. Recent examples of this line of work include Wallsten (2000) on the US SBIR program, Lach (2002) on Israel, Czarnitzki and Licht (2006) on Germany and Criscuolo, Martin, Overman and Van Reenen (2007) on UK. To directly study the level and distribution of welfare effects of an R&D subsidy program, we combine innovation policy analysis with methods from the treatment effects and structural industrial organization literatures. Our model generates an R&D equation that is fairly standard, except for a parameter restriction which we cannot reject. Further, we obtain an economic interpretation of the parameters and the error term in the R&D equation, and information on two hitherto unmeasured but important objects: agency specific treatment effects and application costs of firms.

Methodologically we build on the structural industrial organization (surveyed by Ackerberg, Benkard, Berry and Pakes, 2006, and Reiss and Wolak, 2006) and treatment effects (eg Abbring and Heckman, 2006, and Heckman and Vytlacil, 2006a,b) literatures. While structural methods have been extensively used in many areas of innovation research,[2] their applications to R&D subsidies have been limited, as pointed out by Klette, Møen and Griliches (2000). A notable exception is González, Jaumandreu, and Pazó (2005) who focus on the effectiveness of subsidies in stimulating private R&D. We complement their study by abstracting from set-up costs but emphasizing spillovers and treatment effects and by

[1] R&D subsidies are promoted by theoretical research (eg Howitt, 1999, and Segerstrom, 2000) and they constitute the second largest and fastest growing form of industrial aid in developed countries (Nevo, 1998); the US has had several programs (Lerner, 1999) and currently spends $1.5 billion a year on one R&D subsidy program alone (the SBIR; see http://www.sba.gov/sbir/indexwhatwedo.html, visited on January 21, 2004) and the EU exempts R&D subsidies from its state aid rules. In Finland where our data originates, R&D subsidies are the most important tool of innovation policy (Georghiu et al, 2003).

[2] See, eg Pakes (1986) on patent value, Levin and Reiss (1988) on cost-reducing and demand creating R&D, Lanjouw (1998) on patent value and litigation, Eaton and Kortum (2002) on the role of trade in diffusing the benefits of new technology, Jovanovic and Eeckhout (2002) on the impact of technological spillovers on the firm size distribution, and Petrin (2002) on the welfare effects of new products.

explicitly modelling the application process. There is also a close link from our paper to the methods used in research on firm-regulator interaction (eg Wolak, 1994, and Gagnepain and Ivaldi, 2002).

Due to our institutional environment and data we need to modify the standard structural treatment effect model, first advocated by Heckman and Robb (1985) and Björklund and Moffitt (1987). The treatment effects literature focuses on voluntary social programs where participation is largely determined by self-selection.[3] In R&D subsidy programs the behavior of the agency administering the program also affects participation substantially. Our set-up necessitates building, solving and estimating a game of incomplete information between a potential applicant and the agency. Our model generates a treatment outcome that is a function of the applicant's investment, which in turn is a function of the received treatment. We consider a continuous, optimal treatment with ex ante treatment uncertainty, and outcome and application cost heterogeneity. What we ultimately obtain depends on what one is willing to assume about the objectives of the agency. At a minimum, we identify how the agents expect to benefit from a given treatment. If one assumes a benevolent social planner, we identify general equilibrium treatment effects. In this sense our approach of using the information in agency decisions complements existing work on estimating general equilibrium treatment effects (see Heckman, Lochner and Taber, 1998, Abbring and Heckman, 2006).

Our data contains R&D investment plans instead of outcomes. We are hence concerned with ex ante treatment effects and an ex ante policy evaluation in the sense that the R&D investments are yet be made. In other words, our calculations are informative of the decision makers' preferences and the consequences of their decisions prior to uncertainty about project outcomes unfolding.[4] A policy should (at least) at this stage exhibit benefits that are larger than costs.

We have access to rich R&D project level data from Tekes (the National Technology Agency of Finland), the main source of R&D subsidies in Finland.[5] Finland provides an interesting case because innovation policy has long been a central theme in government policy, the country has particularly rapidly transformed to a technology intensive economy (see eg Trajtenberg, 2001), and subsidies are the principal innovation policy tool.[6] The data contain all the subsidy applications with details of the planned R&D projects, the agency's internal ratings of the applications and its decisions over a two- and half-year period (Jan. 2000 – June 2002). The information on applications is matched to data on over

[3] Heckman, Smith and Taber (1996) study the behavior of bureaucrats running a social program and Heckman and Smith (2004) the determinants of social program participation.
[4] McFadden (1975, 1976) is an early example of the revealed preference approach to public sector decision making. We embed that approach into a treatment model.
[5] Henderson and Cockburn (1996) is an important exception in the existing R&D literature in that they also use R&D project level data.
[6] For example, there are no R&D tax benefits.

14 000 Finnish firms that constitute a large proportion of potential applicants. To get acquainted with the actual decision making process, one of us spent eleven months in Tekes. Among other things she participated in the decision making meetings.

We report four main findings. First, the treatment effects on both the firms and the agency are very heterogenous. Second, application costs vary greatly and shocks to application costs and marginal profitability of R&D are positively correlated. That is, the better the project, the less likely a firm is to apply for a subsidy. This is intuitive once one observes that a major part of the application costs that we uncover comes from opportunity costs. Third, the agency specific treatment effects are smaller than private (firm) treatment effects. Fourth, we find that the rate of return on the subsidy program is of the order of 30–50%. In addition to the main findings, some of our parameter estimates are of independent interest, indicating eg economies of scale in externalities.

In Section 2 we first elaborate on the relationship between our model and a typical estimation equation exploring additionality on the one hand, and the standard treatment effect model on the other hand, and then present our model. We explain the institutional background and data in Section 3 and statistical assumptions, identification and estimation in Section 4. Econometric results are reported in Section 5. In Section 6 we present estimates of various treatment effects and our estimate of the agency's return on the R&D subsidy program, exploring also their sensitivity to distributional assumptions. Conclusions are in Section 7.

2 The model

2.1 An additionality model

The rich literature on the additionality of R&D subsidies provides a variety of ways to estimate the effects of subsidies on R&D investments. One variant can be written as

$$f(R_i) = X_i\beta + \kappa b(s_i) + \varepsilon_i \tag{2.1}$$

where $f(R_i)$ is a (eg log) transformation of the R&D investment R_t of firm i (gross or net of subsidies), X_i is a vector of firm characteristics and β the associated parameter vector, $b(s_i)$ a transformation of the subsidy s_i, κ the additionality parameter at the center of interest, and ε_i the error term. If $b(s_i)$ is the monetary value of the subsidy and $f(R_i)$ the R&D investment, κ measures the marginal effect of an extra subsidy dollar on private R&D. Data used to estimate (2.1) is typically at firm level, and R&D expenditures of firms with and without subsidies are observed. The econometric concern is that selection into the subsidy program is based also on ε_i, rendering $b(s_i)$ endogenous.

In the literature (2.1) is interpreted as a behavioral equation. There is some flexibility in choosing the functions $f(R_i)$ and $b(s_i)$, and a range of values for κ are possible. Values of $\kappa \neq 1$ are motivated by nonlinearities eg due to financial market imperfections or set-up costs. In our case, the R&D equation is a first-order condition, which imposes a form on both $f(R_i)$ and $b(s_i)$, and a value on κ. The main interest is in the other estimated parameters and their implications. The model underlying the first order condition suggests economic interpretations of the parameters and the error term, which generate our heterogenous private (firm) treatment effects. Since we only observe R&D plans for those firms that actually applied for subsidies, we encounter a selection problem.

2.2 The standard selection model

In terms of the treatment effect literature, our model is close to a generalized Roy model with essential heterogeneity (Heckman, Urzua and Vytlacil, 2006, and Heckman and Vytlacil, 2006a,b). There are two outcomes, with (subscript 1) and without treatment (subscript 0)

$$\Pi_{i1} = \Pi_1(X_i, \varepsilon_{i1}), \Pi_{i0} = \Pi_0(X_i, \varepsilon_{i0}) \tag{2.2a}$$

where X_i and ε_{il}, $l \in \{0,1\}$, are observable applicant characteristics and the outcome shocks. In addition, there is a selection equation

$$d_i = 1[E(\Pi_{i1} - \Pi_{i0}) - K(Y_i, \nu_i) \geq 0] \tag{2.2b}$$

where d_i is a binary treatment indicator ($d_i = 1$ means treatment and $d_i = 0$ otherwise). A potential applicant decides whether to take the treatment or not based on the expected gain. The gain consists of the difference between the expected difference in outcomes $E(\Pi_{i1} - \Pi_{i0})$ and the application costs $K(Y_i, \nu_i)$, which are affected both by observable applicant characteristics Y_i and a shock ν_i.

Applicant investment and the agency running the treatment program are seldom explicitly modeled, but typically each applicant who is actually treated is assumed to get the same treatment. The model (2.2a,b) is taken to data on applicant (who often are individuals) characteristics and outcomes. Either Π_{i1} or Π_{i0} is observed for each applicant i.

In our case, (potential) applicants are firms who have ideas for R&D projects that require costly investments. The treatment is an R&D subsidy, reimbursing some fraction of R&D investment cost, ie the treatment is continuous rather than discrete. The treatment program is run by a public agency who screens and evaluates the project proposals and then decides how large a subsidy, if any, to give to each actual applicant. When the applicants decide whether or not to apply for a treatment, they encounter uncertainty about the exact level of treatment. Subsequently, all firms, irrespective of the application and treatment decisions, invest in R&D effort in their project to maximize their expected discounted profits.

In many environments, functions $\Pi_l(X_i, \varepsilon_{il})$ and outcome shocks ε_{il} may be choice specific (ie $\Pi_1(\ ) \neq \Pi_0(\ )$ and $\varepsilon_{i1} \neq \varepsilon_{i0}$), since the treatment may lead to eg a change in profession. However, in our case the actual treatment affects marginal cost of investment rather than the outcome of the investment directly, while the expected effects of the treatment affect the application decision. Hence, assuming $\Pi_1(\ ) = \Pi_0(\ )$ and $\varepsilon_{i1} = \varepsilon_{i0}$ is a plausible point of departure.[7]

2.3 A model with endogenous continuous treatment and investment

We model the subsidy program as a four-stage game of incomplete information between a firm with an R&D project and the agency. In stage zero, the players' types are determined. The agency has a three-dimensional type, $t^A = (\eta, \omega_c, \omega_m) \in \Re^3$, drawn from a common knowledge joint distribution. Each firm is endowed with one R&D project which has a two-dimensional type, $t^F = (\varepsilon, \nu_0) \in \Re^2$, drawn from a common knowledge bivariate distribution. Given our assumption of one project per firm, we talk interchangeably of 'firm' and 'project' type in what follows. Conditional on publicly observed information the shocks are independently distributed. The type of a player contributes to the player's valuation of a project. In stage one, the firm decides whether or not to apply for the subsidy program. The application includes a proposal for an R&D project. In stage two, the agency whose objective function is assumed to include the firm's profits as an argument screens and evaluates the proposed project. It then decides the level of subsidy, s, $s \in [0, \bar{s}]$ and $\bar{s} \leq 1$, which is the share of the

[7] The treatment model we develop could also be applied (with proper modifications) to some other treatment programs than R&D subsidies where similar institutional details are relevant. For example, our set-up is close to what Jaffe (2002) calls a 'canonical' research grant program and the one in Roberts, Maddala and Enholm (1978) who study what determines whether a regulated firm requests a review of its regulated rate of return.

investment cost covered by the agency. In stage three, the firm makes the R&D investment, R, $R \in [0,\infty)$, with or without the subsidy.

Our model builds on the following key assumptions:

A.1 The potential applicant is uncertain about the agency's valuation of the applicant's project.
A.2 A subsidy cannot be misused.
A.3 There are no constraints on the firm's investment.
A.4 The agency's budget constraint does not bind.
A.5 The firm's investment is non-contractible.

A.1 ensures – in line with our data – equilibrium outcomes where a firm applies for a subsidy only to be turned down. It accommodates various informational assumptions concerning the players' types. Due to our functional form assumptions (see equations (2.3) and (2.8) below), the firm can neither signal its type nor does the agency care about it. We only need to assume that the firm, when contemplating application, does not exactly know how the agency values the proposed project. For clarity, we assume that the firm's type is common knowledge and that the agency learns its type exactly after screening.[8]

A.2 excludes moral hazard problems in the use of a treatment.[9] By A.3, the solution to the applicant's maximization problem in the last stage is interior which greatly facilitates the estimation of our model. This assumption rules out credit rationing and other discontinuities that have been emphasized in the innovation policy literature. For example, A.3 amounts to assuming that firms have already made the fixed (R&D) investments. One can defend the assumption of no fixed costs on the grounds that the applicants are existing firms who submit plans for a

[8] That is, symmetric but incomplete information regarding the agency's type prevails in the application stage. Alternatively, we could assume that the applicant has private information about the agency's utility from its project and that the agency receives a noisy signal upon it after screening the project. Since the applicant could not credibly signal its private information in our model, this assumption would yield the same optimal application and subsidy decisions as the (more realistic) assumption we use. In an earlier version (HECER DP no. 76/2005) we develop a treatment program model with general functional forms. There we need to assume that the firm's type is common knowledge to rule out signalling.

[9] In practice, moral hazard temptations are certainly possible with monetary treatments. As a result, Tekes has several safe-guards against expropriation. For example, subsidies are only paid against receipts, there is a euro limit to a subsidy, and a significant number of subsidized R&D projects is annually randomly audited. Because the safe-guards are common knowledge, and the misuses found in the audits or otherwise are rare, we think that the assumption depicts equilibrium behavior.

new project.[10] A.4 is also motivated by simplicity, but we do impose a cost of financing on the agency. A.5 is more realistic, since it prevents the firm and the agency from writing a binding contract specifying the amount the firm invests conditional on the subsidy.

We focus on perfect Bayesian equilibria where, in stage one, a potential applicant correctly anticipates the type-contingent strategies of the agency in stage two, and where the firm's and agency's strategies are sequentially rational. In this extensive form game the firm's posterior belief concerning the agency's type after receiving a subsidy is inconsequential, so we start from the firm's maximization problem in stage three.

2.3.1 Objective function of the firm and stage three of the game

We specify firm i's objective function as

$$\Pi(R_i, s_i, X_i, \varepsilon_i) = \pi_i + \exp(X_i\beta + \varepsilon_i)\ln R_i - (1 - s_i)R_i \qquad (2.3)$$

where s_i is the subsidy, R_i the R&D investment, X_i a vector of observable firm characteristics, and β a vector of parameters to be estimated. The marginal profitability is affected by a random shock, ε_i, (ie, by firm i's type), uncorrelated with the observable firm characteristics, observed by the firm, and unobserved by the econometrician. It may or may not be observable to the agency. The reservation value including other projects is embodied in π_i.[11] $\Pi()$ measures the expected discounted profits, conditional on subsidy s_i.

In stage three, the firm chooses its investment R_i to maximize (2.3). Since the objective function is concave in R_i, the first-order condition

$$R_i = \frac{\exp(X_i\beta + \varepsilon_i)}{1 - s_i} \qquad (2.4)$$

[10] While we make A.3 for simplicity, we note that the revealed motivations for R&D subsidies have increasingly been based on spillovers rather than financial market failures. A study using Finnish data (Hyytinen and Pajarinen, 2003), and an evaluation of Finnish innovation policy (Georghiu et al, 2003) conclude that only small, R&D intensive, growth-oriented firms may face financial constraints. The situation is similar in many other industrialized countries, as the survey by Hall (2002) confirms. The decline of the financial constraint motivation for R&D subsidies is also reflected in our application: although Tekes also grants low-interest loans, most firms were not interested in them.

[11] We could also generalize (2.3) to multiple projects. For each firm with multiple project applications, we could treat each project as a separate observation. If the project-specific unobservables are uncorrelated, this will not materially affect estimation. The interpretation for non-applicants would be that none of their projects resulted in an application.

gives the firm's optimal investment $R_i(s_i)$. Equations (2.3) and (2.4) show the economic interpretation of ε_i: a positive shock to the marginal profitability leads to a larger investment. An optimal investment given by (2.4) could in theory decrease profits but, in such a case, the firm would not invest at all, and consequently would not apply for a subsidy.

Equation (2.3) is our equivalent of the outcome equation (2.2a) in the generalized Roy model with the exception that in our case Π refers to expected discounted profits. Equation (2.4) in turn will produce our R&D investment equation that is close to the additionality equation (2.1). Since we abstract from fixed costs of starting an R&D project, and assume that a subsidy has no impact on the idea behind the project, the treatment in our model only affects the intensive margin without nonlinear effects. To allow a nonlinear effect of the subsidy we could, for example, write the last part of (2.3) as $(1-s_i)^{\kappa}R_i$ where κ would measure additionality. We prefer our formulation for three reasons. First, the formulation $(1-s_i)^{\kappa}R_i$, while useful, is ad hoc.[12] Second, the interpretation of κ would be ambiguous. Using a Box-Cox transformation to model returns to R&D in (2.3) would yield the same estimation equation as the assumptions of logarithmic returns in R&D and $(1-s_i)^{\kappa}R_i$. Third, we cannot reject the Null that $\kappa = 1$.

2.3.2 Agency utility and stage two of the game

The agency's expected utility from applicant i's project is given by

$$\begin{aligned} &U(R_i(s_i), s_i, X_i, Z_i, \varepsilon_i, \eta_i) \\ &= V(Z_i, \eta_i, R_i(s_i)) + \Pi(X_i, R_i(s_i), s_i, \varepsilon_i) - g s_i R_i(s_i) - F_i \end{aligned} \tag{2.5}$$

where F_i captures the fixed costs of applying and processing the application and g is the constant opportunity cost of the agency's resources, eg the opportunity cost of tax funds. The firm's utility enters directly and additively the agency's utility function.

The interpretation of V( ) is fundamental to our analysis. V( ) captures the expected effects of the firm's project on the agency beyond the firm's utility and the direct costs of the subsidy and the application process. V( ) can include externalities from a firm's R&D such as consumer surplus or spillovers to other firms. It could also contain idiosyncratic benefits to the decision maker, through eg bribes or a revolving door mechanism. This agency specific utility can also be decreasing in R&D through negative externalities.

[12] Theoretically more justified ways to introduce non-linear effects of subsidies, eg via fixed start-up costs at the project level, would greatly complicate the estimation of the model.

η_i in V( ) constitutes part of the agency's type, and it is defined as a random shock to the agency specific utility from project i. It is assumed to be uncorrelated with firm characteristics, and unobserved by the econometrician. By A.1, η_i is also unobserved by the potential applicant and observed by the agency (at the latest) after application and screening takes place. In other words, the potential applicant is uncertain about how the agency, after screening the project proposal, sees the project and its potential to generate spillovers, consumer surplus, or private benefits to the agency's civil servants.

The agency specific utility V( ) also includes Z_i, a vector of observable firm characteristics, which contains the same elements as X_i. In our case, Z_i also includes two screening outcomes that are partly determined by ω_{ic} and ω_{im}, which are part of the agency's type. After having received a proposal the agency grades its quality in two dimensions. ω_{ic} and ω_{im} are defined as random shocks to the screening outcome (= grade) of project i in grading dimension c and m respectively (where c and m stand for technical challenge and market risk as will be explained in Section 3). The screening outcomes are two grades on a Likert scale of 5 observed by the agency and by the econometrician but not by the firm. We assume that the grading process and its parameters (excluding the ω's) are common knowledge. That is, conditional on observables, the firm correctly assesses the probability of getting a particular grade in each of the two grading dimensions.

In stage two, the agency chooses a subsidy level s_i, $s \in [0, \bar{s}]$ where $\bar{s} \leq 1$, to maximize (2.5), taking (2.4) into account. To arrive at an estimable model we need to specify the effect of R_i on V( ). We assume that

$$\partial V / \partial R_i = Z_i \delta + \eta_i \tag{2.6}$$

where δ is a vector of parameters to be estimated. An implication of (2.6) is that V( ) is proportional to R&D investment. Similar assumptions are common in the literature on growth and R&D spillovers. We test this assumption below and do not reject it.

Using the envelope theorem, (2.3), (2.4) and (2.6), the first-order condition for the agency's unconstrained problem can be written as

$$s_i = 1 - g + Z_i \delta + \eta_i \tag{2.7}$$

We verify later that (2.7) characterizes the maximum. Equation (2.7) shows how the agency's unconstrained decision rule is decreasing in the shadow cost of public funds, g. It is independent of the firm's type, $t^F = (\varepsilon, \nu_0)$, so even if the agency did not know the private shock to the marginal profitability of R&D, it would not matter. The optimal subsidy depends positively on the shock to the agency specific utility, η_i. The minimum constraint of $s=0$ binds for $\eta_i \leq \underline{\eta_i} \equiv g-1-Z_i\delta$ and the maximum constraint of $\bar{s}$ for $\eta_i \geq \overline{\eta_i} \equiv \bar{s}+g-1-Z_i\delta$.

2.3.3 The firm's beliefs and application costs, and the stage one of the game

In stage one, a profit maximizing firm applies for a subsidy if the expected utility from applying is at least as large as that from not applying. To calculate the expected profit increase from applying, the firm needs to calculate expected profits from submitting an application based on its beliefs about the agency's valuation of its application. As mentioned, the agency's valuation of the project i depends on its type $t_i^A = (\eta_i, \omega_{ic}, \omega_{im})$, which is unknown to the firm prior to the application. Let $\phi(\eta_i)$ define firm i's belief about η_i and let $\Phi(\eta_i)$ be the corresponding cumulative distribution function. Moreover, let $p_{ich}(\omega_{ic})$ and $p_{imh}(\omega_{im})$ denote the firm's beliefs (= probability) that its project i gets grade $h\in\{1,\ldots,5\}$ in grading dimensions c and m respectively.

The firm weights the expected profit increase from applying against its costs. We specify the costs of application as

$$K_i = \exp(Y_i\theta + \nu_i) \tag{2.8}$$

where Y_i is a vector of observable firm characteristics, θ is a vector of parameters to be estimated and ν_i is a random cost shock, distributed by nature, uncorrelated with observable firm characteristics, observed by the firm, and unobserved by the econometrician and the agency (again, the latter is immaterial).

Dropping the subscript i we can now write the applicant's decision rule as

$$\begin{aligned} d = 1\Bigg\{ &\sum_{ch=1}^{5}\sum_{mh=1}^{5} p_{ch}(\omega_c)p_{mh}(\omega_m)\{\Phi(\underline{\eta})\Pi(R(0),0) + \int_{\underline{\eta}}^{\overline{\eta}} \Pi(R(s(\eta)),s(\eta))\phi(\eta)d\eta \\ &+ [1-\Phi(\overline{\eta})]\Pi(R(\bar{s}),\bar{s})\} - \Pi(R(0),0) - K \geq 0\Bigg\} \end{aligned} \tag{2.9}$$

where d_i is an indicator function that takes the value one if firm i finds it profitable to apply for a subsidy and is zero otherwise. Clearly, the application rule (2.9) corresponds to the selection equation of the generalized Roy model (2.2b): there is sorting on the gain. In (2.9) the summation is over the potential screening outcomes. The first term in the curly brackets is the expected profit in case the application is rejected. The rejection occurs when $\eta_i \leq \underline{\eta_i}$, ie with probability

$\Phi(\underline{\eta}_i \equiv g - 1 - Z_i\delta)$. The second term is the expected profit when $\eta_i \in (\underline{\eta}_i, \overline{\eta}_i)$ in which case the firm receives the optimal interior subsidy given by (2.7). The third term is the probability of receiving a maximal subsidy multiplied by the profits with the maximal subsidy. This case occurs with probability $1 - \Phi(\overline{\eta}_i \equiv \bar{s} + g - 1 - Z_i\delta)$. The two last terms capture the costs of applying. Besides the fixed application costs K_i, the firm takes into account that it can execute the project without a subsidy (as implied by A.3), in which case the project yields $\Pi(R_i(0),0)$.

2.3.4 Equilibrium

We complete the model by showing that there is a unique Perfect Bayesian equilibrium, ensuring a meaningful econometric implementation of the model. Perfect Baysian equilibria in our model consist of four components: 1) A firm's decision whether to apply for a subsidy or not, $d_i \in \{0,1\}$; 2) the firm's belief functions $p_{ijh}(\omega_{ij})$, $h \in \{1,\ldots,5\}$, $j \in \{c, m\}$, and $\phi(\eta_i)$ that describe a (common) assessment of how the agency values the firm's project; 3) the agency's subsidy decision rule $s_i = s_i^* d_i$ which determines the level of subsidy granted to firm i given d_i; and 4) the firm's investment rule $R_i^*(s_i)$ given s_i and d_i.

PROPOSITION. There is a unique Perfect Bayesian equilibrium where d_i is given by (2.9), s_i^* is zero for $\eta_i \le \underline{\eta}_i$, is given by (2.7) for $\eta_i \in (\underline{\eta}_i, \overline{\eta}_i)$ and is $\bar{s}$ for $\eta_i \ge \overline{\eta}_i$, and $R_i^*(s_i)$ is given by (2.4).

Proof: For brevity of notation, we drop the subscript i. In stage three, the firm has a well-defined best-reply function R*(s) given by (2.4). In stage two, the agency maximizes its expected utility conditional on its type $t^A = (\eta, \omega_c, \omega_m)$ and receiving an application $(d = 1)$. There is a unique type-contingent optimal subsidy s*, if the second order condition for the agency's decision problem holds. Since we have linear constrains of minimum and maximum subsidies, it suffices to show that U(R*(s),s) is concave when evaluated at the interior solution given by (2.7). Differentiating (2.5) twice using the fact that $\partial \Pi / \partial R = 0$ shows that U(R*(s),s) is concave if

$$\frac{\partial^2 V}{\partial R^2}\left(\frac{dR}{ds}\right)^2 + \frac{dR}{ds}\left(\frac{\partial^2 \Pi}{\partial R \partial s} - 2g\right) + \frac{d^2 R}{ds^2}\left(\frac{\partial V}{\partial R} - gs\right) + \frac{\partial^2 \Pi}{\partial s^2} < 0 \tag{2.10}$$

Since from (2.3) and (2.6) we see that $\partial^2\Pi/\partial s^2$ and $\partial^2 V/\partial R^2$ are zero, (2.10) simplifies to $\frac{dR}{ds}\left(\frac{\partial^2\Pi}{\partial R\partial s}-2g\right)+\frac{d^2R}{ds^2}\left(\frac{\partial V}{\partial R}-gs\right)<0$. Using (2.3), (2.4) and (2.6) we get $\frac{R}{1-s}(1-2g)+\frac{2R}{(1-s)^2}(Z\delta+\eta-gs)<0$, which is equivalent to $1-2g+\frac{2(Z\delta+\eta-gs)}{1-s}<0$. Evaluating this inequality at the interior solution given by (2.7) yields –1 < 0. Consequently, there is a unique maximum that solves the agency's decision problem. Because the optimal unconstrained subsidy (2.7) is increasing in η, s* = 0 for $\eta \le \underline{\eta}$, s* is given by (2.7) for $\eta \in (\underline{\eta},\overline{\eta})$ and $s^* = \bar{s}$ for $\eta \ge \overline{\eta}$, and this s* determines s given d = 1. If the agency does not receive an application (d = 0), s = 0 irrespective of the agency's type. Thus, conditional on d, the type-contingent action of the agency in stage two is unique. In stage one the firm decides whether to apply or not given s* and $p_{jh}(\omega_j)$ and $\phi(\eta)$. Since in a Perfect Bayesian Equilibrium the firm's choice must maximize the profits and the firm's beliefs must be consistent with the agency's strategy, d = 1 only if (2.9) holds and d = 0 otherwise. Clearly, the agency's best response to d = 1 is s = s* so we have found a Perfect Bayesian equilibrium. Since the utility maximizing action in each stage of the game is unique, the equilibrium is also unique.■

2.4 Treatment effects

In the language of the treatment effects literature, we concentrate on internal validity (eg Heckman and Vytlacil (2006a), ie, the effects of actual interventions on outcomes, and we estimate ex ante treatment effects (eg Abbring and Heckman, 2006). The traditionally measured treatment effect on the treated (for a given s_i) is given by $\Pi(R_i,s_i)-\Pi(R_i,0)=-\exp(X_i\beta+\varepsilon_i)\ln(1-s_i)$. We label this the 'gross private treatment effect'. This depends on firm-specific observables and unobservables and is by construction heterogenous. We measure the 'net private treatment' effect by subtracting the heterogenous costs of application K_i.

In our model a subsidy has a separate effect on the agency. We name the subsidy-induced change in the agency specific utility V( ) 'the agency treatment effect'. The (ex ante) gross agency treatment effect conditional on s_i is $V_i[R_i(s_i)-R_i(0)]=-\frac{(Z_i\delta+\eta_i)\exp(X_i\beta+\varepsilon_i)s_i}{1-s_i}$.[13] The joint treatment effect is defined as the sum of private and agency treatment effects.

We identify the private and agency treatment effects irrespective of the interpretation of V( ). If one furthermore assumes that the agency is a benevolent

[13] The fixed cost of screening applications is ignored (ie we assume that $F_i = K_i$).

social planner, V( ) will capture all general equilibrium effects of a treatment outside those appropriated by the applicant and the joint treatment effect is the general equilibrium treatment effect. In this sense our approach of using the information in agency decisions complements existing work on estimating general equilibrium treatment effects (see Heckman, Lochner and Taber, 1998, Abbring and Heckman, 2006). To assess the expected welfare effects of the program one can then compare the estimated rate of return on subsides to the opportunity cost of public funds.

3 Finnish innovation policy, Tekes and data[14]

3.1 Innovation policy and Tekes' subsidy program

In 2001 Finland invested 3.6 per cent of GDP – 5 billion€ – on R&D. Tekes is the principal public financier of private R&D in Finland.[15] The primary objective of Tekes is to promote the competitiveness of Finnish industry and the service sector by providing funding and advice to both business and public R&D. To this end Tekes strives to increase Finnish firms' R&D and risk-taking. Tekes is also responsible for allocating funding from European Regional Development Funds (ERDF), which is meant for the less-favored regions. Finnish regions are heterogenous: eg some 20% of the population lives in the capital region in Southern Finland, where also a large part of the economic activity and most of R&D takes place.

Besides funding business R&D, Tekes finances feasibility studies, and R&D by public sector including scientific research. In 2001 Tekes funding amounted to 387 million€, and it received 2948 applications of which almost exactly 2/3 were accepted. The number of applications by the business sector for R&D funding was 1357 and, again, 2/3 of them were accepted. In monetary terms, the business sector applied for 526 million€ while 211 million€ were granted to it.

Business R&D funding consists of grants, low-interest loans and capital loans. Tekes' low-interest loans not only have an interest rate below the market rate but they are also soft. If the project turns out to be a commercial failure, the loan may

[14] As our application data is from Jan. 2000 – June 2002, we use 2001 figures to describe the environment. Public information about Tekes can be found at http://www.tekes.fi/eng/, accessed December 20th, 2004. Public information is supplemented by knowledge we acquired when one of us spent nine months in Tekes to participate in the actual decision making process.

[15] Main public funding organizations in the Finnish innovation system in addition to Tekes are the Academy of Finland, Employment and Economic Development Centers (T&E Centers), Finnvera, Industry Investment and Sitra. Also the Foundation of Finnish Inventions (Innofin) provides financial support for innovation. See Georghiu et al (2003) for a description of the Finnish innovation policy institutions.

not have to be paid back. A capital loan granted by Tekes differs from the standard private sector debt contract in various ways: it is included in fixed assets in the balance sheet, it can be paid off only when unrestricted shareholders' equity is positive and the debtor cannot give collateral for the loan. The share of each instrument of the total funding allocated to business R&D in 2001 was 69%, 18% and 13%. Subsidy applications covered 83% of the amount applied whereas in terms of granted amount subsidies' share was 67%. The overlook of loans by applicants suggests that they may not encounter significant financial constrains, supporting our assumption A.4 (cf. footnote 10).

The application process from the submission to the final decision, which to our understanding is well known among potential applicants, proceeds along the lines of the theory model of Section 2. In practice, Tekes screens the application and grades it in several dimensions, not two, as we assume for simplicity. The two dimensions concerning the technological challenge of the project and its market risk that we use are, however, in practice the most important ones.[16] Tekes' public decision criteria are: the project's effect on the competitiveness of the applicant, the technology to be developed, the resources reserved for the project, the collaboration with other firms within the project, societal benefits, and the effect of Tekes' funding. Tekes takes into account whether the application comes from an SME and, as mentioned above, the funding also has a regional dimension through ERDF. Putting the regional aspect aside, the funding from ERDF is subject to the same general criteria as other Tekes' funding.

An application has to include the purpose and the budget of the R&D project for which Tekes funding is needed, and the applied amount of funding in euros. Tekes' final decision is based on the proposed budget of the project before the R&D investments are made and a subsidy is granted as a share of to-be-incurred R&D costs.[17] Decision making is constrained by the rules preventing negative subsidies and very large subsidies both in relative and absolute terms. If the firm fulfils the EU SME criterion, the upper bound for the share of covered R&D costs is 0.6, otherwise 0.5. We use this variation in identification, imposing the exclusion restriction that the SME status has an effect only on Tekes' decision and application costs of the firms. This exclusion restriction comes from the

[16] A loose translation of grades of technological challenge is 0 = 'no technical challenge', 1 = 'technological novelty only for the applicant', 2 = 'technological novelty for the network or the region', 3 = 'national state-of-the-art', 4 = 'demanding international level', and 5 = 'international state-of-the-art'. For market risk, it is 0 = 'no identifiable risk', 1 = 'small risk', 2 = 'considerable risk', 3 = 'big risk', 4 = 'very big risk', and 5 = 'unbearable risk'. Since only five grades are used in practice, we, too, use a 5-grade Likert scale.

[17] Tekes sometimes adjusts a proposed budget downwards when an applicant, eg applies subsidies for costs that Tekes cannot cover. An upward adjustment is also possible in principle but rare in practice, occurring virtually only if a project significantly changes character during the application process. Such upgrades can thus be taken as exogenous events that cannot be manipulated by Tekes to overcome the institutional limits on its subsidy allocation.

institutional environment and could also be used in an analysis of additionality.[18] Actual funding is only given after the R&D investments are made, covering the promised share of incurred costs up to a specified euro limit. The limit should allow the promised reimbursement of investment costs up to the profit maximizing level but prevent Tekes from covering costs extraneous to the project proposal.[19] In terms of our model, the rules governing feasible subsidies amount to $\underline{s} = 0$, $\overline{s} \in \{0.5, 0.6\}$ and a goal of setting the euro limit at sR(s).

3.2 Data

Our data come from two sources. The project level data come from Tekes, containing all applications to Tekes from January 1st 2000 to June 30th 2002. They consist of detailed information on the project proposals and Tekes' decisions. The firm level data covering originally 14 657 Finnish firms come from Asiakastieto Ltd, which is a for-profit company collecting, standardizing, and selling firm specific quantitative information.[20] Asiakastieto's data are based on public registers and on information collected by Asiakastieto itself. The data contain for example, firms' official profit sheet and balance sheet statements, and include all the firms who file their data in the public register or submit the information to Asiakastieto. We use the 1999 cross section, ie all firm characteristics are recorded earlier than the application data. The sample was drawn from Asiakastieto's registers in 2002 according to three criteria: i) the most recent financial statement of the firm in the register is either from 2000 or 2001; ii) the firm is a corporation; and iii) the industrial classification of the firm is manufacturing, ICT, research and development, architectural and engineering and related technical consultancy, or technical testing and analysis. Firms in these industries are the most likely to apply for funding from Tekes. After cleaning the

---

[18] Given our data, it is unlikely that firms deliberately keep themselves below the EU SME boundary requiring that a firm has less than 250 employees and has either sales less than 40 million euros or the balance sheet less than 27 million euros. Most of the firms in our data are well below the boundary, as 95% them have less than 110 employees, less than 14 million euros in sales, and a balance sheet of less than 11 million. As the SME criterion also maintains that large firms can hold at most 25% of a SME's equity and votes, it is unlikely that many of the SMEs are subsidiaries of large firms. We thus consider the SME status of a firm exogenous.

[19] As mentioned in footnote 9, the euro limit alleviates the moral hazard problem. There are also other reasons for the limit. Because Tekes has an annual operating budget, a practical decision rule is to cap the euro amount using the proposed budget, as it is the best available information at the time the subsidy decision is made. Tekes is also monitored both by the press and politicians. Tekes civil servants may want avoid the accusations of granting larger subsidies than originally planned. At the same time, however, there may be a desire to make the limit high enough to allow profit maximizing behavior of applicants.

[20] More information about Asiakastieto can be found at http://www.asiakastieto.fi/en/, accessed June 20th, 2005.

data of firms with missing values, we are left with 10 944 firms. These firms constitute our sample of potential applicants.

The firms in our sample account for roughly half of all applications. There are three principal reasons for the exclusion of an applicant from our sample: 1) the firm did not exist in 1999; 2) the firm did not operate in the industries from which the sample was formed; and 3) the firm was so small that it was not obliged by law to send its balance and profit sheets to the official registry.

The data we use in the estimations comprises 915 applications, where we use the first application in case a firm had multiple applications within our observation period. 722 of these applications were accepted, ie received a positive subsidy share. Table 3.1 displays summary statistics of our explanatory variables for potential applicants, and Table 3.2 conditions the statistics on the application decision and success. As Table 3.1 shows, potential applicants are heterogenous. They are on average 12 years old with 35 employees. A very high proportion of firms are SMEs according to the official EU standard (cf. footnote 18). Sales per employee, a measure of value added, is 165 000€. Some 22% are exporters.

Table 3.1 **Descriptive statistics**

	Mean	S.d.	Min.	Max.
Age, years	12	9.3	1	97
# Employees	35	257	1	13451
Sales/employee, 1000€	165	2157	0	206875.5
Exporter	0.22	0.42	0	1
SME	0.98	0.16	0	1
CEO is chairman of board	0.14	0.35	0	1
Board size	4.35	2.00	1	10
# past Tekes applications	0.58	3.49	0	146
Applicant	0.08	0.28	0	1

Notes: There are 10945 observations.
Data sources: Asiakastieto LTd. Otherwise; for data on applications, Tekes.

We also have information on two corporate governance variables. In some 14% of potential applicants, the CEO is also the chairman of the board. Such an arrangement can, on the one hand, improve the information flow between the board and the executive but, on the other hand, weakens the board's independence. The board of an average potential applicant has four to five members. A larger board is costlier but is more likely to include members with outside knowledge that may be useful either in conducting R&D (eg choosing among competing projects, organizing management of current projects, monitoring), or in the application process itself.

From Table 3.2 we see that applicants are larger than non-applicants and successful applicants larger than rejected ones. The median number of employees for non-applicants is 5, for applicants 26, and for rejected applicants 21. The applicants also tend to have larger boards. Quite naturally, applicants have more previous applications on average than non-applicants. The difference in both means and medians is 4.

Table 3.3 reports information about applications and Tekes' decisions (see Appendix A2 for more details). Some 21% of applications are rejected. The proposed projects involve on average an investment of 630 000€, the rejected proposals being smaller with a mean of 386 000€. According to Tekes' rating, the projects have on average a technical challenge of 2 (scale 0–5), and rejected proposals have on average a lower score of 1.5. The mean risk score is also 2, but it is the same for successful and rejected applications.

Table 3.2 **Conditional descriptive statistics**

	Non-Applicants	Applicants	Rejected Applicants	Successful Applicants
Age	12	12	12	12
	(9)	(10)	(10)	(9)
	[10]	[10]	[9]	[10]
# Employees	21	189	101	212
	(122)	(776)	(188)	(867)
	[5]	[26]	[21]	[27]
Sales/employee	169	122	105	126
	(2253)	(55)	(94)	(167)
	[76]	[90]	[83]	[92]
Exporter	0.19	0.57	0.52	0.59
	(0.39)	(0.50)	(0.50)	(0.49)
SME	0.99	0.85	0.86	0.85
	(0.12)	(0.36)	(0.35)	(0.36)
CEO is chairman of board	0.14	0.15	0.18	0.14
	(0.35)	(0.36)	(0.38)	(0.35)
Board size	4.2	6.2	5.9	6.3
	(1.9)	(2.4)	(2.3)	(2.5)
	[4]	[6]	[5]	[6]
# past Tekes applications	0.25	4.16	3.23	4.41
	(1.28)	(10.66)	(10.93)	(10.58)
	[0]	[2]	[1]	[2]
Nobs.	10030	915	193	722

Notes: Number reported are mean, (standard deviation), and for other than [0,1] variables, [median].
Data sources: Asiakastieto Ltd. Otherwise; for data on applications, Tekes.

As explained, Tekes grants low-interest and capital loans besides subsidies. Because it is hard to calculate the value of such non-standard loans to the applicants, we pool the instruments. We thus define the subsidy per cent as the sum of all three forms of financing, divided by 'accepted proposed' investment.[21]

[21] As mentioned in footnote 17, Tekes sometimes adjusts a proposed budget, eg when an applicant applies for subsidies for costs that Tekes cannot cover. We use the unadjusted number ('proposed investment') in our estimations.

As some 60% of applicants only apply for a subsidy, and over 80% are only granted a subsidy, this seems a reasonable simplification. Measuring a subsidy in this way, 0.4% of applicants get the maximum subsidy.[22] Successful applicants receive on average a subsidy that covers 32% of the R&D investment costs. We test the robustness of our results to the definition of a subsidy by using only pure subsidies.

Table 3.3 **Descriptive statistics of Tekes and application variables**

	All Applicants	Sucdessful Applicants	Rejected Applicants
Applied amount, €	634294	700378	385790
	(1254977)	(1363460)	(657540)
Applied for subsidy only	0.59	0.48	1.00
	(0.49)	(0.50)	(0.00)
Technical challenge	2.1	2.3	1.5
	(0.98)	(0.87)	(1.00)
	{582}	{426}	{156}
Risk	2.2	2.2	2.3
	(0.94)	(0.93)	(0.94)
	{422}	{326}	{96}
Granted subsidy rate	–	0.32	–
		(0.13)	
Granted subsidy only	–	0.84	–
		(0.60)	
Nobs.	915	722	193

Notes: Datasource: Tekes. Reported numbers are mean, standard deviation, and {nobs}, the last in case it deviates from that reported on the last row.

4 The econometric model

4.1 The model

We now operationalize the model presented in Section 2. We assume that the agency gives each application i a grade $h\in\{1,\ldots,5\}$ in dimension $j\in\{c, m\}$ by using a latent regression framework. Denoting the latent value of grading dimension $j\in\{c, m\}$ for application i by w^*_{ij} and the observed value by w_{ij}, we get

[22] There is a cluster of firms right below the maximum subsidy: 1.9% of applicants get a subsidy which is less than one percentage point below the maximum subsidy, and 2.5% get a subsidy less than 5 percentage points below the maximum. At the lower end there is no such clustering: on the contrary, no firm gets a subsidy that is less than 2.9%: however, 2.6% of applicants get a subsidy that is greater than 2.9% and less than 5%.

$$
\begin{aligned}
&w_{ij} = h \text{ if } \mu_{h-1} < w^*_{ij} = T_i\zeta_j + \omega_{ij} \le \mu_h \\
&h = 1,\ldots,5, \mu_0 \to -\infty, \mu_1 = 1, \mu_2 = 2,\ldots,\mu_5 \to \infty \\
&\omega_{ij} \sim N(0,1),\ j \in \{c,m\},\ \text{cov}(\omega_{ic},\omega_{im}) = 0
\end{aligned} \tag{4.1}
$$

where T_i is a vector of observable firm characteristics and ζ_j is a parameter vector to be estimated. The unobservables ω_{ij}, which are part of the agency's type, are assumed to be normally distributed and uncorrelated both with each other and other unobservables of the model. We further assume that firms know this grading process and its parameters (excluding the ω_{ij}), using (4.1) to generate the probabilities $p_{ijh}(\omega_{ij})$ of getting grade h in dimension j.

By substituting (2.4) for R_i in (2.3) and using the resulting expression in (2.9) the application decision rule (2.9) can be simplified to

$$d_i = 1\{\exp(X_i\beta + \varepsilon_i)[-E(\ln(1-s_i)) - K_i] \ge 0\} \tag{4.2}$$

Using (2.8), and taking logarithms on both sides, the application rule can be derived from (4.2) as

$$d_i = 1\{X_i\beta - Y_i\theta + \ln[-E(\ln(1-s_i))] \ge \nu_i - \varepsilon_i\} \tag{4.3}$$

The investment equation can be rewritten, upon taking logarithms of (2.4), as

$$\ln R_i^*(s_i) = X_i\beta - \ln(1-s_i) + \varepsilon_i \tag{4.4}$$

with observation $\ln R_i = d_i \ln R_i^*$. As it stands, (4.4) is close to the prototypical additionality equation of Section 2.1. In our model, $f(R_i) = \ln R_i^*(s_i)$, $b(s_i) = -\ln(1-s_i)$ (as in González, Jaumandreu, and Pazó, 2005), and $\kappa = 1$. The traditional approach would concentrate on estimating and interpreting κ, and would interpret (4.4) as a behavioral rule. Here we test and do not reject the restriction that $\kappa = 1$, which arises from the underlying economic model. Another difference to earlier work is that in our model (4.4) is a first order condition. We can hence use equation (4.4) to estimate the parameters of the underlying profit function to calculate treatment effects on expected discounted profits We could calculate the treatment effects for arbitrary values of κ.

The agency decision rule is

$$s_i^* = 1 - g + Z_i\delta + \eta_i \tag{4.5}$$

with observations $s_i = s_i^* d_i$ for $s_i^* \in (0,\bar{s})$, $s_i = \bar{s} d_i$ if $s_i^* \ge \bar{s}$ and $s_i = 0$ if $s_i^* \le 0$.

Our econometric model can thus be summarized by *the screening equations (4.1), the application equation (4.3), the investment equation (4.4) and the Tekes decision rule (4.5).*

4.2 Statistical assumptions, identification and estimation

We now explain our statistical assumptions, how identification takes place, and how we estimate the model. We drop the i subscript whenever no confusion arises. The five unobservables (ω_j, ε, η and ν) are assumed uncorrelated with observed applicant characteristics. Assuming that η is uncorrelated with ε and ν_0 yields a large reduction in computational cost, as then the Tekes decision rule (4.5) is no longer subject to a selection problem and estimation can be broken into three steps. Since we cannot reject the Null hypothesis of no correlation between $\varepsilon - \nu$ and η, in estimating the model by ML, we impose

A.6 a) $\nu = (1 + \rho)\varepsilon + \nu_0$, b) $\eta \perp \varepsilon$, c) $\eta \perp \nu_0$, d) $\varepsilon \perp \nu_0$, e) $\omega_j \perp \varepsilon$, f) $\omega_j \perp \eta$, g) $\omega_j \perp \nu$, h) $\eta \sim N(0, \sigma_\eta^2)$, i) $\varepsilon \sim N(0, \sigma_\varepsilon^2)$, j) $\nu \sim N(0, \sigma_{\nu_0}^2)$.

In words, the unobservable η affecting the agency specific utility is uncorrelated both with the unobservable ε affecting the marginal profitability of the applicant's investment and with the unobservable ν affecting the application cost. Note that the assumptions do *not* mean that the agency specific utility V is uncorrelated with the shock to the marginal profitability of R&D (ε). The agency specific utility is given by $V = (Z\delta + \eta)R = (Z\delta + \eta)\exp(X\beta + \varepsilon)/(1 - s)$ and thus affected by ε. The screening equation unobservables ω_j are uncorrelated with all other shocks. As A.6a shows, there is no restriction on the correlation between ν and ε. A.6h-j are relaxed when we use semi-parametric estimation methods.

The first step is the estimation of the ordered probit screening equations (4.1). Using the estimates we can calculate the firms' belief that a submitted application gets a particular grade in the two evaluation dimensions. Our assumption that the unobservables are normally distributed allows us to identify the coefficients up to scale.

The second step is to estimate the Tekes decision rule (4.5). It identifies δ, ie the effect of observed applicant and project characteristics on the agency specific utility. This is enough to identify V( ) since a project generates utility only with positive R&D and therefore the constant of the integration of (2.6) is zero. We impose (and test) A.6b and A.6c, and estimate (4.5) using a two-limit Tobit model without correcting for selection. We also estimate (4.5) non-parametrically by a two-limit version of Powell's (1984) CLAD estimator.

We then calculate an estimate of E(ln(1 – s)) using the estimated screening equations and the Tekes decision rule. In step three we then estimate the application and investment equations ((4.3) and (4.4)) by replacing E(ln(1 – s)) with its estimated counterpart. Estimation is done using both ML and a semi-parametric variant of the approach suggested by Das, Newey, and Vella (2003, henceforth DNV).[23]

Our data contains information on the proposed R&D investment, not the realized one. The model implies that an applicant strictly prefers proposing a budget based on a maximum subsidy over proposing any smaller amount, and is indifferent between proposing that budget and any larger amount.[24] We estimate β on our data by inserting $\bar{s}$ into (4.4) and correcting for selection bias using the application equation (4.3).

The application equation (4.3) allows us to identify how observed applicant characteristics affect the fixed costs of application. These costs are crucial for welfare analysis and any counterfactual analyses, and they could not be identified without a theoretical model. Our theoretical model imposes a form for the error term in the application equation and, as a result, we identify the correlation between ν and ε when using ML and assuming normal distributions for the shocks. Moreover, we can then identify the variance of the error term in (4.3) since following theory the coefficient of the term E(ln(1 – s)) is constrained to unity.[25] To calculate the treatment effects and application costs we also estimate (4.2) using the semi-nonparametric estimator of Gallant and Nychka (1987) as it allows to recover the distribution of shock term ($\rho\varepsilon + \nu_0$) in (4.3) without imposing a distributional assumption on either ε or ν_0.

To obtain consistent standard errors in the application and investment equations ((4.3) and (4.4)), we bootstrap the whole model (4.1), (4.3)–(4.5) both when using ML with normality assumptions and when using the semi-parametric DNV estimator.

Note also what we cannot identify. In (2.3) we are unable to identify π, the applicant's reservation value. Our cross section estimates are however not affected by unobserved differences in the reservation value. Similarly, we cannot

[23] Manski (1989) compares merits of the parametric and non-parametric approaches. Manski argues that, although the nonparametric approach appears to be more flexible, it involves arbitrary exclusion restrictions. Therefore it is not necessarily preferable over the parametric one.

[24] Too see this, recall first that the applicant does not know Tekes' type (A.1) and the subsidy share is bounded above at $\bar{s}$. As mentioned in Section 3.1, there is also an euro limit to the ex post reimbursements which is based on the proposed budget. Then, since $\partial\Pi/\partial s > 0$ by (2.3), the applicant wants as high a subsidy as possible. Therefore it proposes an optimal project based on the maximum subsidy share, $R^*(\bar{s})$. Proposing anything less risks foregoing profits in case where the actual subsidy turns out to be larger and the applicant subsequently reoptimizes because of the euro limit. On the other hand, the applicant would never want to implement a project larger than $R^*(\bar{s})$, and it is indifferent between announcing $R^*(\bar{s})$ and any larger budget, given the assumption that it cannot misappropriate the funds.

[25] This implication of our theoretical model cannot be tested.

identify g, the opportunity cost of government funds. We are also unable to identify the agency's screening costs (F–K), which will result in an upward bias in the welfare calculations.

5 Estimation results

We include into all estimation equations firm age, the log of the number of employees, sales per employee, a dummy for a parent company, the number of previous applications, a dummy indicating if the CEO acts as the chairman of the board, board size, and a dummy for exporters. We also include industry and region dummies.[26] The SME dummy is only included in the Tekes decision rule (4.5) and the application equation (4.3). We include it in (4.3) to allow for the possibility that SMEs' opportunity costs are different eg because of different access to other types of subsidies. Inclusion of the SME dummy in the application equation and exclusion of it from the R&D equation is sufficient for (nonparametric) identification. Our model yields additional identification through the expectation term in (4.3).

In the reported specifications, we use a slightly different set of explanatory variables in the screening equations (4.1) and the Tekes decision rule (4.5) on the one hand, and the application and investment equations ((4.3) and (4.4)) on the other. For example, we include the squares of the continuous variables in application and investment equations ((4.3) and (4.4)).[27] The results from the estimation of the screening equations (4.1) are reported in the Appendix.

Based on our semi- and non-parametric estimations we find no evidence that the distributional assumptions of shocks are driving our parameter estimates. Our cross validation results (see Appendix A5) however reject the double normality assumptions A.6i,j on the investment and application cost shocks. To explore the robustness of our treatment effect estimates we therefore also employ the Gallant and Nychka semi-nonparametric estimator when estimating the application decision. A detailed description of these robustness checks is presented in the Appendix. We have also estimated the model (by ML with normality assumptions) by excluding the observations in the 99th size (sales) percentile, with essentially identical results to those reported. Other robustness checks will be taken up in the context of the appropriate estimation.

[26] We divide Finland into five regions: Southern, Western, Eastern, Northern and Central Finland. Of these, Eastern and Northern Finland are the least developed. We did try interactions between firm characteristics and industry and region dummies.

[27] To speed up the computation of the bootstrap we used LR-tests to narrow the set of explanatory variables in each equation. The second order terms were excluded from the screening equations (4.1) and the Tekes decision rule (4.5) based on the LR-tests.

5.1 The Tekes decision rule and agency specific utility

In Table 5.1 we report the results concerning the Tekes decision rule. The coefficients can be interpreted as the marginal effects of R&D on agency specific utility. We find that the more challenging a project is technically, the higher is its subsidy rate. A one point increase on the 5-point Likert scale leads to a 10 percentage point increase in the subsidy rate. Market risk carries a negative but insignificant coefficient. Firm size obtains a positive and significant (at 10% level) coefficient. Moving an otherwise identical R&D project into a larger firm creates larger positive externalities, eg through higher employee rents. As against Tekes' stated preference that allows a 10 percentage points higher level of maximum subsidy for SMEs, it is unsurprising that SMEs are granted a higher subsidy, everything else equal: the difference is 8.3 percentage points. The corporate governance variables and the number of previous applications have no effect. We relegate the industry and regional dummy-results to the Appendix.

Table 5.1 **Tekes decision rule results**

Variable	Dep. var. subsidy-intensity (all finance)
Risk	-.020*
	[-.043 .003]
Technical challenge	.100***
	[.076 .124]
Age	-.001
	[-.003 .002]
Log employment	.019*
	[-.001 .039]
Sales/employment	.00005
	[-.0001 .0002]
SME	.083*
	[-.003 .169]
Parent company	.006
	[-.041 .052]
# previous applications	-.001
	[-.007 .004]
CEO also chairman	.001
	[-.054 .055]
Board size	-.007
	[-.017 .003]
Exporter	-.021
	[-.079 .038]
Constant	-.054
	[-.215 .107]
σ_η	.190***
	[.173 .206]
Nobs.	379
LogL.	-19.216
Wald	0.000
Linearity 1	0.659
Linearity 2	0.197
Sample sel.	.030
	(.027)

Notes: Reported numbers are coefficient and [95% confidence interval]. Wald is the p-value of a Wald test of joint significance of all RHS variables. All specifications include industry and region dummies.
Linearity 1 = the p-value of a LR-test of including the proposed R&D investment into the equation.
Linearity 2 = the p-value of a LR-test of including the proposed R&D investment into the equation, plus interactions between it and age, log employment, and sales/employee.
Sample sel. = coeff. and (s.e.) of the Mills ratio term when the 1(apply) specification same as in Table 5.2.
***, **, and * denote significance at 1, 5, and 10% level.

The above results are obtained under the assumptions A.6b and A.6c, which maintain that the error in the Tekes decision rule is uncorrelated with the errors in the investment and application equations. To test these assumptions, we first estimated a probit application equation[28] and then re-estimated the Tekes decision rule by inserting the Mills ratio into it. The Mills ratio obtained imprecisely estimated coefficients with values close to zero in all of our several specifications, validating our assumptions of no correlation. Recall that this does not imply that agency specific utility is independent of profitability shocks, but rather that profitability shocks are transmitted to agency specific utility entirely through R&D.

We also tested our assumption that V( ), the agency specific utility, is linear in the applicant's investment as implied by (2.7). Were V( ) non-linear in the applicant's investment, the Tekes decision rule would contain an investment term (R) or its interactions with observable applicant characteristics. We included these and could not reject the Null of (joint) insignificance of them. The agency specific utility from a project seem thus to be linear in R&D.

In Appendix A4 we report results from estimating the Tekes decision rule using CLAD, and an alternative dependent variable. The results are in line with those reported here.

5.2 Cost of application function

In Table 5.2 we report the estimates of the application cost function (equation (2.8)). Age, CEO being chairman, and parent company status have no statistically significant effect, but firm size has a non-linear decreasing effect on application costs. Sales per employee increase application costs. One interpretation is that firms producing high value added products and services have complicated R&D projects based on soft information that are laborious to write down. Another is that because the opportunity costs of the effort of making and promoting an application are probably far greater than the direct monetary costs of filling in and filing it, firms with high value current production have higher opportunity costs of applying. The size of the board has a decreasing effect on application costs. This may reflect the role of external knowledge in lowering application costs. Exporters have lower costs, maybe because they are relatively more experienced in dealing with government bureaucracy than non-exporting firms.

[28] Naturally, the probit was run without the expected subsidy term, but both with and without added interactions to improve identification.

Table 5.2 **Application cost function results**

Variable	Coefficient [95% confidence interval]
Age	.019 [-.016 .709]
Age sq.	-.0001 [-.007 .0003]
Log of employment	-.423** [-10.856 -.043]
Ln(emp) sq.	.069*** [.022 1.382]
Sales/employee	.002*** [.0007 .022]
Sales/emp. Sq.	-7.97e-0.8 [-8.53e-07 1.76e-06]
SME	.591 [-.581 6.939]
Parent company	-.188 [-4.164 .119]
# Previous applications	-.236*** [-5.383 -.077]
# Prev appl. sq.	.002*** [.0005 .037]
CEO is chairman	-.243 [-1.575 .388]
Board size	-.098* [-2.486 .006]
Exporter	-.866*** [-16.604 -.181]
Constant	13.449*** [11.156 100.589]
Nobs	10944

Notes: Confidence intervals are estimated using a bootstrap with 400 repetitions. The specification includes industry and regional dummies.
Wald is the p-value of the joint significance of all explanatory variables in the probit 1st stage regression.
***, **, *, and a denote that the whole 99%, 95%, 90% and 85% confidence interval has the same sign as the coefficient estimate.

The number of past applications has a nonlinear effect, first decreasing and then, after 118 applications, increasing application costs. Increasing the number of past applications from non-applicants' median of zero to applicants' median of two decreases application costs by 37%. One prior application decreases costs by 21% and four by 60%. It seems that learning by doing is going on. Given that our data is cross sectional it is however possible that the results are generated by unobserved heterogeneity.

5.3 Investment equation

This equation is often estimated in existing work on R&D subsidies: Our investment equation (4.4) identifies the effects of exogenous variables on marginal profitability of R&D investment. In view of the received R&D literature, it is likely that unobserved heterogeneity accounts for a substantial part of the marginal profitability of R&D. This is also what we find, as Table 5.3 shows. Firms with higher value-added current production have higher marginal profitability of R&D whereas it appears to be lower in firms with CEOs as chairmen. Other findings are not robust over specifications (see Appendix A5).

The coefficient of $\ln(1-\bar{s})$ is a double test of our specification, though it is admittedly weakened by having variation only through firms' SME status. First, the coefficient could measure additionality (as in González, Jaumandreu, and Pazó, 2005). In that interpretation, with a point estimate of 0.765 and a wide confidence interval, we cannot reject the Null of 1:1 additionality which our theoretical model assumes. Second, if one assumes that the effect of R&D on profits is modeled using a Box-Cox transformation *and* that the additionality parameter κ equals unity, the test becomes a test of the Box-Cox parameter. The interpretation then is that we cannot reject the Null of profits being logarithmic in R&D. Both interpretations provide support for our modeling assumptions.

Table 5.3 **R&D investment function results**

Variable	Coefficient [95% confidence interval]
Age	-.005
	[-.024 .011]
Age sq.	.0001
	[-.0001 .0004]
Log of employment	-.106
	[-.259 .069]
Ln(emp) sq.	.024**
	[.003 .046]
Sales/empl.	.001**
	[.0001 .002]
Sales/emp. sq.	-7.42e-08
	[-5.59e-07 1.74e-06]
Parent company	-.023
	[-.184 .149]
# Previous applications	-.043**
	[-.073 -.008]
# Prev appl. sq.	.0002**
	[-7.26e-06 .0006]
CEO is chairman	-.097
	[-.274 .097]
Board size	.008
	[-.028 .050]
Exporter	-.190*
	[-.383 .043]
Constant	12.840***
	[11.638 13.674]
Nobs.	914
Wald (d.f. X)	0.000
$\ln(1-\bar{s})$	-0.765
	(0.780)

Notes: Confidence intervals are based on a bootstrap with 400 repetitions.
Wald is the p-value of joint significance of RHS variables.
$\ln(1-\bar{s})$ coefficient reports the coefficient and the (p-value) of a χ^2-test of difference from unity.
***, **, *, and a denote that the whole 99%, 95%, 90% and 85% confidence interval has the same sign as the coefficient estimate.

Table 5.4 **Covariance structure results**

Variable	Coefficient [95% confidence interval]
σ_ε	1.212***
Standard deviation of the investment equation shock	[1.010 1.351]
σ_η	.190***
Standard deviation of the Tekes specific utility (= V( )) shock	[.173 .206]
$\sigma_{\nu 0}$	.791***
Standard deviation of the uncorrelated part of the application cost function shock	[.234 20.917]
$1+\rho$	1.673***
Measure of the variance share of ε in ν	[1.174 17.304]
$\rho_{\varepsilon\nu}$	-.718***
Correlation between ε and the application equation error term	[-.832 -.462]

Notes: For all but σ_η, values are based on a bootstrap with 400 repetitions. For σ_η, it is based on the estimated covariance matrix.
***, **, and * denote significance at 1, 5, and 10% level.

5.4 Covariance structure

We are able to identify the variances of all error terms, and the covariance between the unobservables in the application and investment equations (Table 5.4). The coefficient determining the variance share of investment shock in the application cost shock (assumption A.6a) obtains a value of 1.7. Ceteris paribus, the higher the unobserved marginal profitability of the R&D project of a firm, the less likely it is that the firm will submit an application. It could be that, similar to projects with higher sales per employee, projects with higher marginal profitability of R&D are more complicated involving tacit knowledge and are therefore more difficult to describe in an application. Or it could be that projects with higher marginal profitability of R&D have higher opportunity costs, which constitute a major part of application costs.

6 Treatment effects

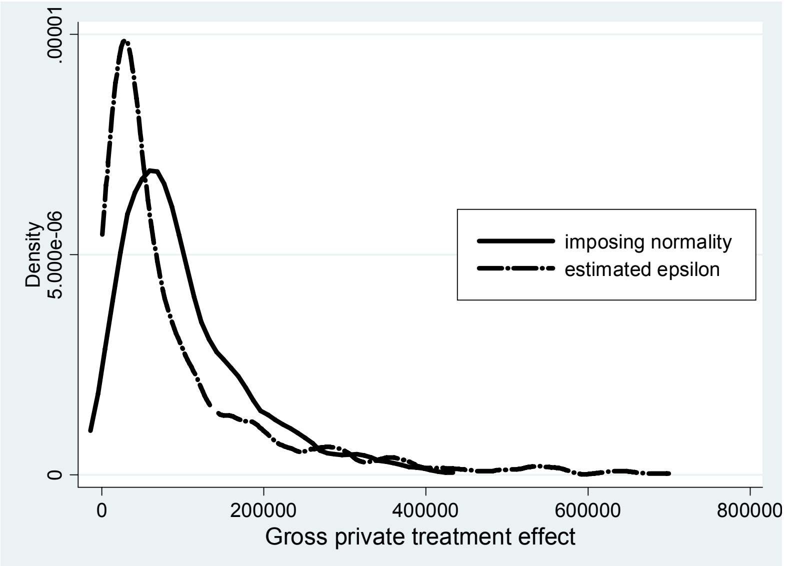
We report medians and means (in parenthesis) of private (firm) and agency treatment effects, relegating the details of the calculations into Appendix A7. Recall that all treatment effects are ex ante in the sense that they are measured prior to the launch of the R&D projects (but after the subsidy decisions).[29]

Table 6.1 **Gross private treatment effects**

Distributional assumptions underlying estimation and integration	Calculation of shocks	Gross private treatment effect on treated, actual subsidy
Double normal	No shocks	149 442 (170 984)
	Integration over ε, ν_0	81 871 (107 461)
	Actual ε	49 706 (108 902)
Double free	Actual ε	49 706 (108 902)

Note: Reported numbers are median and (mean).
Double normal = both ε and ν_0 are assumed to be normally distributed.
Double free = both shocks' distributions are determined (semi-)nonparametrically.

Figure 6.1 **Distribution of the gross private treatment effect for accepted applicants**



[29] The results using subsidies expected by the firms prior to the application decision are very similar to those reported.

In Table 6.1 we report the gross private treatment effects that ignore the costs of application. Comparing the first row with other rows shows that ignoring the investment shock ε and application cost shock ν_0 leads to a large upward bias. By conditioning the value of the shocks on the firm being an applicant (by integrating them over the relevant regions of the shock distributions), and imposing normality on both ε and ν_0 greatly reduces the treatment effects (row 2). This is not surprising given the finding of Section 5.4 that the applicants have smaller values of ε than non-applicants. Rows 2 and 3 reveal that using the estimated value of ε instead of integrating over its (imposed) normal distribution further lowers the median treatment effects by some 40–45%, but mean treatment effects are close to each other. Rows 3 and 4 in turn show that the distributional assumptions make no difference for gross private treatment effects when we use the estimated value of ε.[30] The median (mean) gross private treatment effect is of the order of 50 000€ (100 000€).

Figure 6.1 displays the substantial heterogeneity of gross private treatment effects on the treated applicants (ie the applicants that received a subsidy), calculated under the assumption that both ε and ν_0 are normally distributed and using the estimated ε.

Table 6.2 **Net private treatment effects**

Distributional assumptions underlying estimation and integration	Calculation of shocks	Net private treatment effect on applicants, actual subsidy	Net private treatment effect on treated, actual subsidy
Double normal	No shocks	-121 926 (-384 093)	-74 512 (-385 961)
	Integration over ε, ν_0	30 228 (42 966)	43 916 (64 896)
	Actual ε, integration over $\rho\varepsilon+\nu_0$	31 496 (81 263)	46 253 (103 689)
Double free	Actual ε, integration over $\rho\varepsilon+\nu_0$	33 846 (84 965)	49 463 (107 758)

Note: Reported numbers are median and (mean).
Double normal = both ε and ν_0 are assumed to be normally distributed.
Double free = both shocks' distributions are determined (semi-)nonparametrically

Table 6.2 presents the net private treatment effects that take into account the costs of application. They have been calculated both for all applicants and the treated

[30] This is because $\hat{\varepsilon} = R(\bar{s}) + \ln(1-\bar{s}) - X\hat{\beta}$ and irrespective of the estimation method $X\hat{\beta} + \hat{\varepsilon}$ amounts to the same.

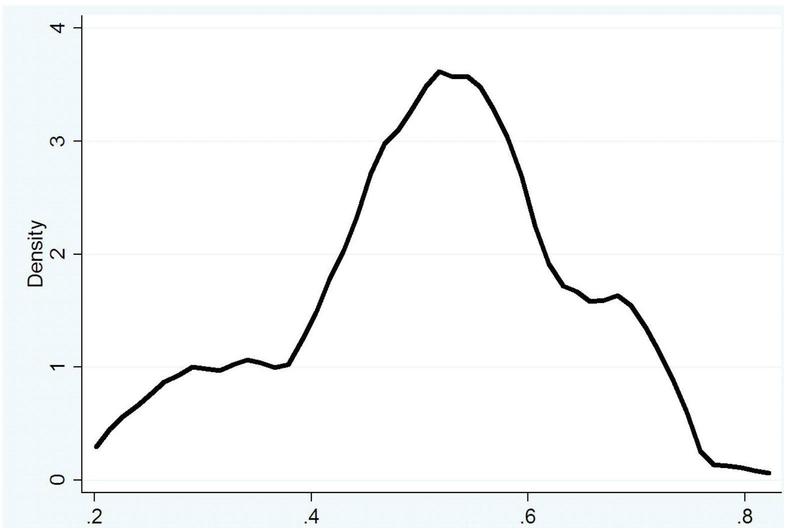
ones. The net private treatment effects ignoring shocks are negative contrary to what we should observe but are positive once we take into account the shocks. Integration over the distributions of ε and ν_0 yields clearly lower estimates of the mean treatment effect than using the estimated value of ε while the medians are relatively close to each other. As we relax the distributional assumptions the treatment effects slightly increase. Net private treatment effects are lower than gross treatment effects, but not substantially for the most reliable estimates.

Table 6.3 **Gross agency specific treatment effects**

Distributional assumptions underlying estimation and integration	Calculation of shocks	Gross agency specific treatment effect, treated, actual subsidy
Double normal	No shocks	97 717 (130 802)
	Integration over ε, ν_0	56 331 (79 990)
	Actual ε	33 565 (75 720)
Double free	Actual ε	33 565 (75 720)

Note: Reported numbers are median and (mean).
Double normal = both ε and ν_0 are assumed to be normally distributed.
Double free = both shocks' distributions are determined (semi-)nonparametrically.

Figure 6.2 **Marginal effect of R&D on agency specific utlity**



We then turn to gross agency treatment effects (Table 6.3). If the agency is a benevolent social planner, they should reflect the change in R&D spillovers and consumer surplus due to the treatment. Agency treatment effects are also calculated with and without the shocks. As in the case of gross private treatment effects, taking the shocks into account lowers the median treatment effects substantially. Using the estimated value of ε again yields lower estimates of the treatment effects than integrating over its (assumed) distribution. Comparison of Tables 6.2 and 6.3 suggests that firms appropriate some 60% of the treatment effects.

Figure 6.2 shows the distribution of $Z\hat{\delta}+\hat{\eta}$, which is the marginal effect of R&D on expected agency specific utility (recall that $V=(Z\delta+\eta)R$). The expected agency specific utility is increasing in R&D investments and hence in the subsidy rate for the most of the projects in our data. The expected increase in agency specific utility is typically between 0.25 and 0.5 per one euro of R&D and for 99% of firms, a one euro increase in R&D leads to a less than 0.85 euro increase in agency specific utility.[31]

Table 6.4 **Application Costs**

Distributional assumptions underlying estimation and integration	Calculation of shocks	Application cost, applicants
Double normal	No shocks	242 986 (519 012)
	Integration over ε, ν_0	35 533 (41 827)
	Actual ε, integration over $\rho\varepsilon+\nu_0$	2 431 (4 762)
Double free	Actual ε, integration over $\rho\varepsilon+\nu_0$	503 (1 061)

Note: Reported numbers are median and (mean).
Double normal = both ε and ν_0 are assumed to be normally distributed.
Double free = both shocks' distributions are determined (semi-)nonparametrically

Table 6.4 reports our estimates of application costs. Ignoring shocks leads to very high application cost estimates, explaining the negative net private treatment effects. Taking the shocks into account by integrating over the distributions of ε and ν_0 reduces the estimated application costs by 85%. Using the estimated value of the investment shock ε reduces the median application cost by another 33 000€, or more than 90%. Relaxing the distributional assumptions further lowers the estimate of the median (mean) application cost shock for applicants. It thus ultimately seems that for an evaluation of the actual policy, application costs may not be of first order importance. Nonetheless, our results from Section 5.4 suggests that any counterfactual policy analysis will critically depend on

[31] We trimmed the sample used in Figure 6.2 at the 99th percentile.

application costs since non-applicants have higher investment shocks and investment and application cost shocks are positively correlated.

Table 6.5 **Rate of return on the subsidy program**

Distributional assumptions underlying estimation and integration	Calculation of shocks	Using the subsidy amount predicted by the model	Based on actual accepted costs
Double normal	No shocks	-1.62	-3.42
	Integration over ε, ν_0	0.98	1.45
	Actual ε, integration over $\rho\varepsilon+\nu_0$	1.31	1.51
Double free	Actual ε, integration over $\rho\varepsilon+\nu_0$	1.34	1.55

Note: Double normal = both ε and ν_0 are assumed to be normally distributed.
Double free = both shocks' distributions are determined (semi-)nonparametrically

Finally, Table 6.5 shows our calculations of the rate of return on the subsidy program.[32] Under the assumption of a benevolent social planner the rate of return can be compared to the opportunity cost of public funds (g) to evaluate the program.[33] We have used g = 1.2. Figures in Table 6.5 show that once shocks are taken into account, the estimated rate of return on the subsidy policy exceeds the opportunity cost of public funds.

7 Conclusions

We analyze one of the mostly widely used innovation policy tools: R&D subsidies. We complement the existing literature by building a structural model of the R&D subsidy process. Our model generates an R&D equation through firms' first order condition that is close to those estimated in existing work. Our model also yields a testable restriction on the key additionality parameter of the traditional estimation equation that we cannot reject. We show how selection of treatment by the agency and 'self-rejection' by the firms – the decision whether to

[32] The joint rate of return on the subsidy program is the overall benefits due to subsidies (net private treatment effect plus gross agency treatment effect) divided by the overall cost of subsidies (granted subsidy share multiplied by the investment given by the granted subsidy), ignoring the shadow cost of taxes, and taking all applicants into account. The difference between the two columns is that the first one is based on the R&D investment predicted by our model and the second one uses the minimum of this and the R&D investment accepted by Tekes.
[33] Kuismanen (2000) estimates the dead-weight loss of existing Finnish taxation to be 15% using labor supply models.

apply or not – provide information on hitherto unmeasured objects: agency specific treatment effects and application costs.

Taking the model to project level data from Finland we find that large firms produce larger agency specific utility, as do technically more challenging projects. Firms with higher value added current production have higher marginal returns to R&D and higher application costs. Profitability and application cost shocks are positively related, implying that firms do not apply for subsidies for the privately most profitable projects.

We estimate ex ante treatment effects that reflect the revealed preferences of the key decision makers at the time they make their decisions – the firms on applying or not, and the agency on the level of subsidy. They thus embody the perceived benefits and costs of the program prior to the actual R&D investments taking place. We are able to extend the number of identified treatment effects and find considerable heterogeneity in all of them. Our estimate of the median net private treatment effect on the treated is close to 50 000€ and the median gross agency specific treatment is approximately 33 000€. These numbers suggest that treated firms internalize 60% of the total treatment effect.

To produce a welfare analysis we use strong but standard assumptions. Our agency specific treatment effects can be interpreted as externalities and our calculated rate of return on subsidies as a social rate of return if one is willing to assume that the agency giving subsidies is a benevolent social planner. In that case our estimates suggest that the program benefits cover the opportunity cost of public funds.

References

Abbring – Heckman, J (2006) **Econometric evaluation of social programs Part III: Distributional treatment effects, dynamic treatment effects, dynamic discrete choice, and general equilibrium policy evaluation.** In Heckman, J and Leamer, E (eds.), Handbook of Econometrics, Vol. 6, North-Holland.

Ackerberg, D – Benkard, L – Berry, S – Pakes, A (2005) **Econometric tools for analyzing market outcomes.** Forthcoming in Heckman, J and Leamer, E (eds.), Handbook of Econometrics, Vol. 6, Amsterdam: North-Holland.

Björklund, A – Moffitt, R (1987) **The estimation of wage gains and welfare gains in self-selection models.** The Review of Economics and Statistics, 69, 42–49.

Cohen, W (1995) **Empirical studies of innovative activity.** In Stoneman, P (ed.), Handbook of the Economics of Innovation and Technological Change, Basil Blackwell.

Criscuolo, M – Overman – Van Reenen (2007) **The effect of industrial policy on corporate performance: Evidence from panel data.** Mimeo, LSE.

Czarnitzki, D – Licht, G (2006) **Additionality of public R&D grants in a transition economy: the case of Eastern Germany.** Economics of Transition, 14, 101–131.

Das, M – Newey, W – Vella, F (2003) **Nonparametric estimation of sample selection models.** Review of Economic Studies, 70, 33–58.

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

Eaton, J – Kortum, S (2002) **Technology, geography, and trade.** Econometrica, 70, 1741–1779.

Gagnepain, P – Ivaldi, M (2002) **Incentive regulatory policies: The case of public transit systems in France.** RAND Journal of Economics, 33, 1–25.

Gallant, A – Nychka, D (1987) **Semi-nonparametric maximum likelihood estimation,** Econometrica, 55, 363–390.

Georghiu, L – Smith, K – Toivanen, O – Ylä-Anttila, P (2003) **Evaluation of the Finnish innovation support system.** Finland's Ministry of Trade and Industry, publications 5/2003.

González, X – Jaumandreu, J – Pazó, C (2005) **Barriers to innovation and subsidy effectiveness.** RAND Journal of Economics, 36, 930–950.

Hall, B (2002) **The financing of research and development.** Oxford Review of Economic Policy, 18, 35–51.

Heckman, J – Lochner, L – Taber, C (1998) **General-equilibrium treatment effects: A study of tuition policy.** American Economic Review, Papers and Proceedings, 88, 381–386.

Heckman, J – Robb, R (1985) **Alternative methods for evaluating the impact of interventions.** In Heckman, J and Singer, B (eds.), Longitudinal Analysis of Labor Market data, Econometric Society Monographs, 10, Cambridge.

Heckman, J – Smith, J (2004) **The determinants of participation in a social program: Evidence from a prototypical job training program.** Journal of Labor Economics, 22, 243–298.

Heckman, J – Smith, J – Taber, C (1996) **What do bureaucrats do? The effects of performance standards and bureaucratic preferences on acceptance into the JTPA program.** In Liebcap, G (ed.), Advances in the Study of Entrepreneurship, Innovation and Growth, 7, 191–217, JAI Press.

Heckman, J – Urzua, S – Vytlacil, E (2006) **Understanding instrumental variables in models with essential heterogeneity.** Review of Economics and Statistics, 88, 3, 389–432.

Heckman, J – Vytlacil, E (2006a) **Econometric evaluation of social programs Part I: Causal models, Structural models, and econometric policy evaluation.** In Heckman, J and Leamer, E (eds.), Handbook of Econometrics, Vol. 6, North-Holland.

Heckman, J – Vytlacil, E (2006b) **Econometric evaluation of social programs Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New Environments.** In Heckman, J and Leamer, E (eds.), Handbook of Econometrics, Vol. 6, North-Holland.

Henderson, R – Cockburn, I (1996) **Scale, scope and spillovers: The determinants of research productivity in drug discovery.** RAND Journal of Economics, 27, 1, 32–59.

Howitt, P (1999) **Steady endogenous growth with population and R&D inputs growing.** Journal of Political Economy, 107, 715–730.

Hyytinen, A – Pajarinen, M (2003) **Financial systems and firm performance: Theoretical and empirical perspectives.** Taloustieto Ltd., Helsinki.

Jaffe, A (2002) **Building programme evaluation into the design of public research-support programmes.**Oxford Review of Economic Policy, 18, 22–34.

Jovanovic, B – Eeckhout, J (2002) **Knowledge spillovers and inequality.** American Economic Review, 92, 5, 1290–1307.

Klette, T – Møen, J – Griliches, Z (2000) **Do subsidies to commercial R&D reduce market failures?** Microeconomic evaluation studies, Research Policy, 29, 471–495.

Kuismanen, M (2000) **Labour supply and income tax changes: a simulation study for Finland.** Bank of Finland Discussion paper, 5/2000.

Lach, S (2002) **Do R&D subsidies stimulate or displace private R&D? evidence from Israel.** Journal of Industrial Economics, L, 369–390.

Lanjouw, J (1998) **Patent protection in the shadow of infringement: Simulation estimations of patent value.** Review of Economic Studies, 65, 4, 671–710.

Lerner, J (1999) **The government as a venture capitalist: The long-run impact of the SBIR program.** Journal of Business, 72, 285–318.

Levin – Reiss, P (1988) **Cost-reducing and Demand-creating R&D with spillovers.** RAND Journal of Economics, 19, 4, 538–556.

Manski, C (1989) **The anatomy of the selection problem.** Journal of Human Resources, 24, 343–360.

McFadden, D (1975) **The revealed preferences of a government bureaucracy: Theory.** Bell Journal of Economics, 6, 401–416.

McFadden, D (1976) **The revealed preferences of a government bureaucracy: Empirical evidence.** Bell Journal of Economics, 7, 55–72.

Nevo, R (1998) **Trends and Patterns of Public Support to Industry in the OECD Area.** STI Review, 21, 12–24.

Pakes, A (1986) **Patents as options: Some estimates of the value of holding European patent stocks.** Econometrica, 57, 755–784.

Petrin, A (2002) **Quantifying the benefits of new products: The case of the Minivan.** Journal of Political Economy, 110, 705–729.

Powell, J (1984) **Least absolute deviations estimation for the censored regression model.** Journal of Econometrics, 25, 303–325.

Reiss, P – Wolak, F (2006) **Structural econometric modeling: Rationales and examples from Industrial Organization.** Forthcoming in Heckman, J and Leamer, E (eds.), Handbook of Econometrics, 6, Amsterdam: North-Holland.

Roberts, B – Maddala, G – Enholm, G (1978) **Determinants of the requested rate of return and the rate of return granted in a formal regulatory process.** Bell Journal of Economics, 9, 611–621.

Segerstrom, P (2000) **The Long-Run Growth Effects of R&D Subsidies.** Journal of Economic Growth, 5, 3, 277–305.

Stewart, M (2004) **Semi-nonparametric Estimation of Extended Ordered Probit Models.** Stata Journal, 4, 27–39.

Trajtenberg, M (2001) **Innovation in Israel 1968–1997: A comparative analysis using patent data.** Research Policy, 30, 363–389.

Wallsten, S (2000) **The effects of government-industry R&D programs on private R&D: The case of the Small Business Innovation Research Program.** RAND Journal of Economics, 31, 82–100.

Wolak, F (1994) **An econometric analysis of the asymmetric information regulator-utility interaction.** Annales d'Economie et de Statistique, 34, 13–69.

Yatchew, A (1998) **Nonparametric regression techniques in economics.** Journal of Economic Literature, 36, 669–721.

Appendix

In this Appendix, we report the ordered probit estimation of the Tekes grading process (A1); descriptive statistics of a) the whole application sample b) the application sample who have strictly positive accepted proposed investments, and c) the application sample for which we observe grades in both evaluation dimensions (A2); industry and region dummy descriptive statistics (A3); robustness checks of the Tekes decision rule (A4) and the investment equation (A5); coefficients of the industry and region dummies for the estimated equations (A6); details of how we have calculated the treatment effects (A7); and point estimates of the application cost function obtained using the semi-nonparametric estimator of Gallant and Nychka (1987) in the application equation and DNV-estimator in the investment equation (A8)

A1 The screening/grading equations

We have different applicant samples in the estimations of the two grading dimensions, because sometimes we only observe one or the other grade for an application. During our observation period, Tekes did not uniformly store grading data in their central database, from which our data has been collected. We use the estimation results to create the probabilities of getting a particular grade for all the 10751 (10944) observations in the estimation sample.

In the technical challenge estimation, sales per employee, number of previous applications, board size, and industry dummies (chemical, industry, electric engineering, data processing, and R&D services) increase the probability of getting a high grade in evaluation of technical challenge. Having a CEO as chairman and being in the food or paper industry decreases the probability of getting a high grade.

In the market risk estimation, sales per employee and a number of industry dummies have a negative effect on the probability of obtaining a high risk rating (high meaning higher risk). The industry dummies that carry significant negative coefficients are paper, other manufacturing, and telecoms. Being located in Western Finland also decreases the probability of being classified as high risk.

Table A1 **Estimation of the screening equations**

Variable	Technical challenge	Risk
Age	.002	-.003
	[-.008 .012]	[-.015 .009]
Log employees	-.006	-.047
	[-.080 .068]	[-.133 .040]
Sales/employee	.001***	-.001*
	[.0001 .002]	[-.002 .0002]
Parent company	-.019	-.118
	[-.223 .185]	[-.357 .120]
# previous applications	.023*	-.020
	[-.0001 .046]	[-.047 .006]
CEO in chairman	-.247**	-.014
	[-.488 -.007]	[-.296 .268]
Board size	.080***	.033
	[.036 .123]	[-.017 .082]
Exporter	.251**	-.319**
	[.005 .498]	[-.619 -.019]
Nobs.	582	422
LogL.	-752.711	-527.563
Joint Significance	0.000	0.0000

Notes: reported numbers are coefficient and [95% confidence interval]. Joint Significance is the p-value of a LR test of joint significance of all explanatory variables. Both specifications include industry and region dummies.
***, **, and * denote significance at 1, 5, and 10% level.

A2 Descriptive statistics of the applicant samples

Table A2 presents the descriptive statistics for the three samples of applicants mentioned above. As can be seen, the differences are minor; judging on observables, we are unlikely to have a selection problem among applicants in the subsidy equation. The only potentially worrisome difference is that in the smallest sample, the mean number of previous application is lower (2.8) than in the other two (4.2 and 4.4). The standard error also declines. Also, the proportion of telecom firms and firms in Eastern Finland are somewhat lower. As we report in the main text, we found no evidence for sample selection after testing it against the whole sample.

Table A2 **Descriptive statistics of different applicant samples**

Variable	All applicants	Applicants with strictly positive proposed accepted investment	Applicants for whom grades in both evaluation dimensions are observed
Age	12	12	11
	(9.6)	(9.5)	(9.0)
Log employees	3.4	3.5	3.2
	(1.8)	(1.8)	(1.7)
Sales/employee	122	126	120
	(155)	(167)	(128)
SME	.85	.85	.88
	(.36)	(.36)	(.33)
Parent company	.51	.53	.48
	(.50)	(.50)	(.50)
# previous applications	4.2	4.4	2.8
	(10.7)	(10.6)	(4.5)
CEO is chairman	.15	.14	.17
	(.36)	(.35)	(.38)
Board size	6.2	6.3	6.1
	(2.4)	(2.5)	(2.4)
Exporter	.57	.59	.58
	(.50)	(.49)	(.50)
Food	.04	.04	.03
	(.18)	(19)	(.18)
Paper	.05	.05	.04
	(.22)	(.22)	(.19)
Chemicals	.03	.04	.03
	(.18)	(.18)	(.16)
Rubber	.06	.06	.06
	(.24)	(.24)	(.24)
Metals	.08	.08	.07
	(.27)	(.27)	(.25)
Electric	.10	.11	.11
	(.30)	(.31)	(.31)
Radio and TV	.04	.04	.05
	(.20)	(.19)	(.21)
Other manufacturing	.09	.09	.09
	(.29)	(.29)	(.28)
Telecoms	.01	.01	.003
	(.09)	(.10)	(.05)
Data processing	.21	.20	.26
	(.41)	(.40)	(.44)
R&D	.15	.15	.13
	(.36)	(.35)	(.34)
Western Finland	.32	.32	.35
	(.47)	(.47)	(.48)
Eastern Finland	.12	.13	.06
	(.32)	(.33)	(.23)
Central Finland / Oulu region	.09	.08	.09
	(.28)	(.27)	(.28)
Northern Finland / Lapland region	.02	.02	.03
	(.15)	(.14)	(.17)
Nobs.	915	722	379

A3 Descriptive statistics of the industry and region dummies for the whole sample

Table A3 **Descriptive statistics of the industry and region dummies for the whole sample**

Variable	Mean (s.d.)
Agriculture	.0001
	(.010)
Food	.045
	(.207)
Paper	.061
	(.239)
Chemicals	.015
	(.120)
Rubber	.056
	(.229)
Metals	.139
	(.346)
Electric	.046
	(.209)
Radio and TV	.015
	(.120)
Other manufacturing	.188
	(.391)
Telecoms	.009
	(.095)
Data processing	.105
	(.307)
R&D	.196
	(.397)
Southern Finland	.453
	(.498)
Western Finland	.386
	(.487)
Eastern Finland	.078
	(.268)
Central Finland / Oulu region	.061
	(.240)
Northern Finland / Lapland	.023
	(.149)

Notes: there are 10945 observations.

A4 Robustness checks of the Tekes decision rule

We also estimated the Tekes decision rule by a two-limit version of Powell's (1984) CLAD estimator.[34] As column two of Table A4 shows, the results are relatively close to those obtained using Tobit ML. The only noteworthy differences are that with CLAD, the rubber industry obtains a significant positive coefficient (approximately 0.008 in value, compared with 0.012 for Tobit), and the coefficient of Central Finland is no more significant. There are some relatively large differences between the insignificant coefficients, though.

To test whether measuring the subsidy per cent by summing subsidies, low-interest loans and capital loans affect the results, we estimated the two-limit Tobit using only subsidies, excluding the loans. Column three in Table A4 reveals that our results are not driven by our definition of the dependent variable. We also checked whether the definition of the dependent variable in the Tekes decision rule affects our parameter estimates in the sample selection model (application and R&D investment). The R&D investment equations' parameters are virtually identical, as are most of the parameters of the application equation. All parameters in the application equation are within one standard deviation of each other.

[34] The two-limit CLAD was estimated by using the following algorithm: we first estimated a LAD using all 379 observations, then excluded all observations with predicted values less than the minimum or more than the maximum allowed, and re-estimated the LAD. This was repeated until convergence.

Table A4 **Tekes decision rule results**

Variable	(1) ML Dep. var. subsidy-intensity (all finance)	(2) CLAD Dep. var. subsidy-intensity (all finance)	(3) ML Dep. var. subsidy-intensity (subsidies only)
Risk	-.020*	-.020	-.024**
	[-.043 .003]	[-.046 .006]	[-.048 -.00005]
Technical	.100***	.092***	.104***
challenge	[.076 .124]	[.065 .119]	[.079 .129]
Age	-.001	-.0001	-.001
	[-.003 .002]	[-.0017 .0023]	[-.004 .001]
Log employment	.019*	.025**	.025**
	[-.001 .039]	[.008 .040]	[.004 .046]
Sales /	.00005	.00005	.00007
employment	[-.0001 .0002]	[-.000083 .000151]	[-.0001 .0002]
SME	.083*	.070	.069
	[-.003 .169]	[-.003 .138]	[-.020 .157]
Parent company	.006	.015	.008
	[-.041 .052]	[-.023 .055]	[-.040 .056]
# previous	-.001	-.002	-.002
applications	[-.007 .004]	[-.006 .002]	[-.007 .003]
CEO also	.001	-.018	-.0002
chairman	[-.054 .055]	[-.064 .028]	[-.057 .056]
Board size	-.007	-.0003	-.008
	[-.017 .003]	[-.0084 .0082]	[-.018 .003]
Exporter	-.021	-.016	-.037
	[-.079 .038]	[-.069 .038]	[-.098 .024]
Constant	-.054	-.083	-.079
	[-.215 .107]	[-.233 .028]	[-.246 .088]
σ_η	.190***	-	.196***
	[.173 .206]		[.179 .213]
Nobs.	379	379	379
LogL.	-19.216	-	-21.542
Wald	0.000	-	0.000
Linearity 1	0.659	-	-
Linearity 2	0.197	-	-
Sample sel.	.030	-	-
	(.027)		

Notes: Reported numbers are coefficient and [95% confidence interval]. Wald is the p-value of a Wald test of joint significance of all RHS variables. All specifications include industry and region dummies.
Linearity 1 = the p-value of a LR-test of including the proposed R&D investment into the equation.
Linearity 2 = the p-value of a LR-test of including the proposed R&D investment into the equation, plus interactions between it and age, log employment, and sales/employee.
Sample sel. = coeff. and (s.e.) of the Mills ratio term when the 1(apply) specification same as in Table 5.2.
***, **, and * denote significance at 1, 5, and 10% level.
In columns (1) and (2), the dependent variable is the proportion of expenses that the Agency covers, defined as the sum of all three types of financing the Agency grants (in €, see main text) divided by accepted proposed investment. In column (3), the dependent variable is the subsidy (in €) divided by the accepted proposed investment.

A5 Robustness checks of the investment equation

We estimated the model both by ML, dropping the second order terms, and using DNV's semi-parametric sample selection estimator. We imposed otherwise the structure of the ML specification, but allowed the additively separable error terms to have an unknown distributions. The results are in line with the main ML estimates (reproduced in column 1): most coefficients are within the ML 95% confidence intervals. This suggests that our ML distributional assumptions are not biasing the parameter estimates.

Cross-validation (see Table A5b) suggests that the (double) normality assumption does not hold in the data. We used the same trimming and transformation DNV. The transformation gives exact sample selection correction for Gaussian disturbances. The trimming explains the difference in the sample size compared to ML estimations. We tried up to the 4th order terms for the variable capturing the effect of subsidies on expected discounted profits in the 1st stage, and started from the ML specification. Cross-validation indicated that we should include the subsidy-terms up to the 3rd order, but should not include interactions of the other explanatory variables. In the 2nd stage, we kept the same specification as in ML, and experimented with including up to the 4th order transformation of the propensity score (without interactions with explanatory variables). We used a Gram-Schmidt ortho-normalization for the 3rd and 4th order terms in both stages.

Table A5a **R&D investment function results**

Variable	(1) ML Dep. var. proposed investment	(2) ML Dep. var. proposed investment	(3) DNV Dep. var. proposed investment
Age	-.005	-.002	-.013
	[-.024 .011]	[-.008 .005]	[-.092 .089]
Age sq.	.0001	-	.0002
	[-.0001 .0004]		[-.0003 .0008]
Log of	-.106	.068[a]	.052
employment	[-.259 .069]	[-.013 .128]	[-.497 .736]
Ln(emp) sq.	.024**	-	.003
	[.003 .046]		[-.047 .034]
Sales / empl.	.001**	0.001***	.001[a]
	[.0001 .002]	[.0007 .002]	[-.0004 .004]
Sales / emp. sq.	-7.42e-08	-	-1.73e-07
	[-5.59e-07 1.74e-06]		[-1.19e-06 1.25e-06]
Parent company	-.023	-.002	-.015
	[-.184 .149]	[-.143 .167]	[-.843 .888]
# previous	-.043**	-.009	-.090
applications	[-.073 -.008]	[-.019 .004]	[-1.924 1.253]
# prev appl sq.	.0002**	-	.0006
	[-7.26e-06 .0006]		[-.011 .015]
CEO is	-.097	-.100[a]	-.054[a]
chairman	[-.274 .097]	[-.285 .079]	[-.396 .105]
Board size	.008	.022	.013
	[-.028 .050]	[-.020 .058]	[-.402 .439]
Exporter	-.190*	-.072	-.061
	[-.383 .043]	[-.329 .139]	[-2.678 2.236]
Propensity score	-	-	3.257
			[-121.150 112.261]
Propensity			-7.347
score2			[-127.516 77.826]
Propensity			31.505
score3			[-37.036 66.101]
Constant	12.840***	12.008***	-
	[11.638 13.674]	[11.115 12.956]	
Nobs.	914	914	876
Wald (d.f.X)	0.000	0.000	0.000
$\ln(1-\bar{s})$	-0.765	-0.108	
	(0.780)	(0.165)	

Notes: Reported numbers are coefficient and [95% confidence interval]. Confidence intervals are based on a bootstrap with 400 repetitions. In columns (1)–(3) the dependent variable is the log of accepted proposed investment: in column (4) it is the log of proposed investment.

Wald is the p-value of joint significance of RHS variables. The constant is not identified when using DNV.

$\ln(1-\bar{s})$ coefficient reports the coefficient and the (p-value) of a χ^2-test of difference from unity.

***, **, *, and [a] denote that the whole 99%, 95%, 90% and 85% confidence interval has the same sign as the coefficient estimate.

Table A5b **Cross-validation of the application and R&D investment equations**

Specification	Application equation	R&D investment equation
Linear term	0.0595	1.0246
+2nd power	0.0602	1.0227
+2nd and 3rd power	0.0586	1.0217
+2nd–4th power	0.0635	1.0234
+2nd and 3rd powers and 1st order interactions between continuous variables	0.0982	-

Notes: the linear term is the effect of expected subsidies on expected discounted profits in the application equation, and the propensity score transformation that DNV use (Mills ratio) in the R&D investment equation. The base specification is the same as in the ML estimations

Cross-validation figures were calculated using equation (2.22) in Yatchew (1998).

A6 Coefficients of industry and region dummies

The only industry dummies with significant coefficients are food (p-value .000) and data processing (p-value .081). Using metal manufacturing firms as a reference group, firms in the food industry received a substantially higher subsidy, of the order of 25 percentage points, whereas data processing firms obtained subsidies that were 6.5 percentage points lower. During our observation period, Tekes was actively seeking applications from the food industry, which at least partially explains the findings concerning the industry.

Regional aspects affect Tekes decisions: firms in Eastern and Central Finland obtain subsidies that are 7–10 percentage points higher than those in Southern Finland. That regional policy matters is, however, debatable, as the city of Oulu, which is located in Central Finland is one of the R&D centers in Finland. Moreover, we find that firms in the depressed and sparsely populated Northern Finland do not get higher subsidies.[35]

[35] This finding is perhaps not robust as only 2% of our sample firms come from Northern Finland.

Table A6 **Estimated industry and region dummy parameters**

Variable	Tekes Decision Rule Table A4			Application Cost Function Table 5.2	R&D Investment Function Table A5a		
Column	(1)	(2)	(3)		(1)	(2)	(3)
Food	.242***	.224***	.262***	.222	-.524**	-.480**	-.560*
	[.115 .368]	[.091 .357]	[.132 .392]	[-1.515 2.720]	[-.881 -.151]	[-1.00 -.269]	[-1.184 .219]
Paper	-.028	.016	-.028	.354	.191	.184	.122
	[-.151 .094]	[-.116 .148]	[-.156 .099]	[-0.507 10.445]	[-.140 .550]	[-.350 .343]	[-1.452 1.120]
Chemicals	.096	.060	.114	.901	.219	.233	.239
	[-.038 .230]	[-.092 .212]	[.024 .252]	[-3.292 3.257]	[-.352 .731]	[-.162 .752]	[-.663 .903]
Rubber	.011	.082	.011	.269	.111	.106	.089
	[-.086 .107]	[-.029 .193]	[-.089 .111]	[-.381 3.970]	[-.211 .458]	[-.213 .407]	[-.391 .820]
Metals	.006	.016	-.0008	.555[a]	.370***	.340**	.275
	[-.087 .098]	[-.087 .119]	[-.096 .094]	[-.005 5.738]	[.091 .634]	[-.067 .472]	[-.492 .923]
Electric	-.041	-.006	-.034	.019	.286**	.330**	.293
	[-.124 .041]	[-.101 .088]	[-.119 .051]	[-8.945 .595]	[.044 .575]	[-.030 .540]	[-1.129 1.956]
Radio and TV	-.021	.006	-.009	.531	.649**	.652**	.659*
	[-.129 .086]	[-.113 .126]	[-.120 .102]	[-3.192 1.807]	[.125 1.201]	[.247 1.183]	[-.279 1.600]
Other manufacturing	-.017	.002	-.008	.536	.195	.150	.071
	[-.105 .071]	[-.098 .103]	[-.099 .083]	[-.122 10.378]	[-.078 .470]	[-.379 .217]	[-.760 .811]
Telecoms	-	-	-	.831*	.491[a]	.547*	.457
				[-.295 10.181]	[-.180 1.225]	[-.084 1.08]	[-.910 2.878]
Data processing	-.072*	-.040	-.066	-.562	.200	.327	.314
	[-.154 .010]	[-.135 .055]	[-.151 .019]	[-18.026 .372]	[-.091 .521]	[-.029 .484]	[-2.229 2.690]
R&D	.002	.035	.003	.088	.071	.090	.114
	[-.083 .087]	[-.060 .131]	[-.085 .090]	[-4.200 .576]	[-.215 .353]	[-.286 .226]	[-.359 .374]
Western Finland	.017	.023	.015	.399	.242***	.231**	.237[a]
	[-.029 .064]	[-.030 .076]	[-.033 .063]	[-.427 1.124]	[.084 .414]	[.012 .328]	[-.089 .379]
Eastern Finland	.094**	.090**	.118*	-.429[a]	-.450***	-.399***	-.370
	[.005 .184]	[.013 .193]	[.026 .210]	[-9.837 .053]	[-.675 -.196]	[-.548 -.059]	[-1.724 .891]
Central Finland / Oulu region	.063*	.030*	.071*	-.015	.048	.071	.078
	[-.012 .139]	[-.052 .112]	[-.007 .149]	[-5.404 .453]	[-.225 .355]	[-.246 .255]	[-.772 1.146]
Northern Finland / Lapland	-.031	-.039	-.019	-.024	.095	.140a	.136
	[-.159 .096]	[-.174 .097]	[-.151 .113]	[-2.497 1.770]	[-.262 .593]	[-.027 .715]	[-.243 .717]

Notes: in the Tekes decision rule equations, we excluded the telecommunications dummy because of problems in the bootstrap that were due to the low proportion of telecommunications firms in our sample of firms with both Tekes evaluation grades. ***, **, *, and [a] denote significance at 1, 5, 10, and 15% level. Southern Finland is our base region.

A7 Calculation of treatment effects

We have calculated the treatment effects in three ways. Below we show the calculation of net private treatment effect (NPT) in each of the three ways. All the other treatment effects are calculated in similar manner. First, and solely to demonstrate the importance of shocks, we ignore them. This gives us

$$NPT_i^1 = -\exp(X_i\beta)\ln(1-s_i) - \exp(Y_i\theta)$$

Second, when assuming normal distributions for ε and ν_0, we can calculate expected treatment effects conditional on applying by integrating over the relevant part of the distribution. We thus take into account the fact that for the firms that applied, the following condition is satisfied (see equation (4.2))

$$X_i\beta - Y_i\theta + \ln E[-\ln(1-s_i)] \geq \nu_i - \varepsilon_i = \rho\varepsilon_i + \nu_{0i}$$

Rearranging we get

$$\varepsilon_i \leq \frac{1}{\rho}(X_i\beta - Y_i\theta + \ln E[-\ln(1-s_i)] - \nu_{0i}) = \bar{\varepsilon}_i$$

Using this information about the investment shock, using actual granted subsidies and imposing normality, NPT for applicants can be written as

$$NPT_i^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\bar{\varepsilon}_i} \frac{[-\exp(X_i\beta + \varepsilon_i)\ln(1-s_i) - \exp(Y_i\theta + (1+\rho)\varepsilon_i + \nu_{0i})]f(\varepsilon)d\varepsilon}{F(\bar{\varepsilon})} g(\nu_0)d\nu_0$$

where $f(\ )$ and g( ) are the probability density functions of ε and ν_0 respectively (both assumed to be normal), and F( ) is the cumulative distribution function of ε.

Third, we can recover the investment equation shock ε from the investment equation (4.3) and insert it both in the firms' profit function (2.3) and the application cost function (2.9) where $\nu_i = (1+\rho)\varepsilon_i + \nu_{0i} = \varepsilon_i + (\rho\varepsilon_i + \nu_{0i})$. The last term in parenthesis is the error term in the application equation (4.2). We then integrate over $(\rho\varepsilon_i + \nu_{0i}) = \xi_i$ when calculating the application costs. Only the third method is available when we estimate the application equation semi-nonparametrically. Using this third method NPT can be written as

$$\mathrm{NPT}_i^3 = -\exp(X_i\beta + \hat{\varepsilon}_i)\ln(1-s_i) - \left[\frac{\int_{-\infty}^{\bar{\xi}_i} \exp(Y_i\theta + \hat{\varepsilon}_i + \xi_i)h(\xi)d\xi}{H(\bar{\xi})} \right]$$

where $\bar{\xi}_i = X_i\hat{\beta} - Y_i\hat{\theta} + \ln[-\ln(1-s_i)]$, $\hat{\varepsilon} = \ln R + \ln(1-\bar{s}) - X\hat{\beta}$ and h( ) is the probability density function of ξ (either normal or the estimate provided by Gallant-Nychka) and H( ) the corresponding cumulative distribution function.

A8 Point estimates of the application equation based on semi-parametric estimation

In producing these estimates, we used the semi-nonparametric estimator of Gallant and Nychka (1987) in the application equation and the DNV-estimator in the investment equation. The Gallant and Nychka estimation is based on the code written by Stewart (2004). Because estimation is very slow we have not calculated (via bootstrap) the confidence intervals. The point estimates are within the confidence interval of the point estimates produced using the double normal assumption and reported in Table 5.2.

Table A8 **Point estimates of the application cost function based on semi-parametric estimation**

Variable	Coefficient
Age	.006
Age sq.	.00003
Log of employment	-.099
Ln(emp) sq.	.026
Sales / employee	.002
SME	.425
Parent company	-.089
# previous applications	-.532
# prev appl. sq.	.004
CEO is chairman	-.164
Board size	-.058
Exporter	-.522
Const.	13.479
Food	.111
Paper	.180
Chemicals	.911
Rubber	.224
Metals	.355
Electric	.146
Radio and TV	.589
Other manufacturing	.290
Telecoms	.545
Data processing	-.325
R&D	.213
Western Finland	.257
Eastern Finland	-.257
Central Finland / Oulu region	.096
Northern Finland / Lapland	.168

BANK OF FINLAND RESEARCH
DISCUSSION PAPERS

ISSN 0785-3572, print; ISSN 1456-6184, online

1/2008 Peik Granlund **Regulatory choices in global financial markets – restoring the role of aggregate utility in the shaping of market supervision.** 2008. 36 p. ISBN 978-952-462-416-9, print; ISBN 978-952-462-417-6, online.

2/2008 Ari Hyytinen – Tuomas Takalo **Consumer awareness and the use of payment media: evidence from Young Finnish consumers.** 2008. 34 p. ISBN 978-952-462-418-3, print; ISBN 978-952-462-419-0, online.

3/2008 Patrick M Crowley **One money, several cycles? Evaluation of European business cycles using model-based cluster analysis.** 2008. 47 p. ISBN 978-952-462-420-6, print; ISBN 978-952-462-421-3, online.

4/2008 József Molnár **Market power and merger simulation in retail banking.** 2008. 26 p. ISBN 978-952-462-422-0, print; ISBN 978-952-462-423-7, online.

5/2008 Heli Huhtala **Along but beyond mean-variance: Utility maximization in a semimartingale model.** 2008. 29 p. ISBN 978-952-462-426-8, print; ISBN 978-952-462-427-5, online.

6/2008 Mikael Juselius **Cointegration implications of linear rational expectation models.** 2008. 25 p. ISBN 978-952-462-428-2, print; ISBN 978-952-462-429-9, online.

7/2008 Tuomas Takalo – Tanja Tanayama – Otto Toivanen **Evaluating innovation policy: a structural treatment effect model of R&D subsidies.** 2008. 59 p. ISBN 978-952-462-430-5, print; ISBN 978-952-462-431-2, online.

Suomen Pankki
Bank of Finland
P.O.Box 160
FI-00101 HELSINKI
Finland

. 2343