LSE RICAFE

Working Paper NO. 014

SELECTION OR SELF-REJECTION

Applications into a Treatment Program: The Case of R&D Subsidies

Tuomas Takalo
Research Department, Bank of Finland and IDEI, Toulouse

Tanja Tanayama
VTT Technology Studies and Helsinki School of Economics

Otto Toivanen
HECER, University of Helsinki

February 2005

RICAFE - Risk Capital and the Financing of European Innovative Firms

A project financed by the European Commission, DG Research
Improving the Human Potential and the Socio- Economic Knowledge Base Programme.
Contract No : HPSE-CT-2002-00140

Financial Markets Group, London School of Economics and Political Sciences
Department of Economics and Finance, Turin University.
Centre for Financial Studies - CFS (Frankfurt)
Haute Etudes Commerciales - HEC (Paris)

SELECTION OR SELF-REJECTION?

APPLICATIONS INTO A TREATMENT PROGRAM: THE CASE OF R&D SUBSIDIES

TUOMAS TAKALO, TANJA TANAYAMA, AND OTTO TOIVANEN*

First draft: January 25, 2004

This draft: December 21, 2004

Abstract

We build a structural model of application and selection into a treatment program to study the effects of R&D subsidies. The investments of the participant and the public agency running the program affect the treatment outcome. Using project level data we find that larger firms have higher marginal profitability of R&D. Rates of return on R&D are high and skewly distributed. Median expected profit from applicants' R&D projects is 500 000 euros, which increased by 12 000 by expected subsidies. Firms with SME status receive higher subsidies. Conditioning on SME status, subsidies are increasing in firm size, and in technical challenge of the project. The median increase in the public agency's utility not appropriated by the applicant is 16 000. Non-applicants have lower expected subsidies and higher application costs. Application costs decrease in the number of previous applications and increase with the profitability shock. Ignoring application costs severely biases the estimated rates of return upwards.

KEYWORDS: applications, effort, investment, rate of return, R&D, selection, subsidies, treatment program, treatment effects, welfare.

---

* Takalo: Research Department, Bank of Finland, and IDEI, University of Toulouse, email: takalo@cict.fr; Tanayama: VTT Technology Studies and Helsinki School of Economics, email: Tanja.Tanayama@vtt.fi; Toivanen: HECER, University of Helsinki, email: otto.toivanen@helsinki.fi.
We would like to thank Tim Bresnahan, David Card, Ken Chay, Stefanie Franzke, Ari Hyytinen, Guido Imbens, Phil Leslie, Jim Powell, Paul Ruud, Chris Shannon, Javier Suarez and seminar audiences at HSE (Helsinki), UC Berkeley, UC Davis, Stanford GSB, the Tor Jakob Klette memorial conference at University of Oslo, and the 2nd RICAFE Conference at CFS, Frankfurt, 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 largely written while Toivanen was visiting UC Berkeley, whose hospitality he gratefully acknowledges. The remaining errors and the views expressed in the paper are those of the authors alone.

Government subsidies to R&D investments have become ubiquitous. They are the second largest[1] and fastest growing form of industrial aid in developed countries (Nevo, 1998); the U.S. spends $1.5 billion a year on one R&D subsidy program alone,[2] 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). Given the importance of R&D subsidies we know surprisingly little about the programs that allocate them. How does the public agency running a program decide subsidy levels? How do potential applicants decide whether or not to apply? What are the public agency's and the applicants' costs and benefits from the program? Our main objective is to answer these questions. Our model allows us to also calculate the rates of return on R&D, uncover its determinants, and to make counterfactual policy experiments.

To reach the main objective, we build and estimate a structural model of the application and selection process into a voluntary treatment program. In contrast to the standard treatment model where all accepted applicants get the same treatment and outcome heterogeneity is restricted from the outset, we characterize a continuous, potentially multidimensional, optimal treatment and allow for outcome heterogeneity. The model frees us from the need to concentrate on a single treatment effect parameter and the need to rely on distributional assumptions or ad hoc exclusion restrictions for identification. We obtain economic interpretation for the application equation and all estimated parameters. We therefore extend the treatment program literature (see e.g. Heckman, LaLonde, and Smith, 1999, and Blundell and Costa-Dias, 2002, for surveys) by i) modeling the treatment outcome as a function of applicant investment, ii) endogenizing the level of investment, iii) making it a function of the received treatment,[3] iii) taking into account application costs, and iv)

---

[1] Largest being regional aid. Pretschker (1998) states that financial support schemes constitute the predominant type of government policy towards industrial R&D.

[2] The Small Business Innovation Research Program, SBIR: "In FY (fiscal year) 2001, [the SBIR program] produced 3,215 Phase I awards and 1,533 Phase II awards for approximately $1.5 billion dollars". Phase I is the startup phase. Awards of up to $100,000 for approximately 6 months support exploration of the technical merit or feasibility of an idea or technology. Phase II awards of up to $750,000, for as many as 2 years, expand Phase I results. During this time, the R&D work is performed and the developer evaluates commercialization potential. Only Phase I award winners are considered for Phase II. Quotation and information are from http://www.sba.gov/sbir/indexwhatwedo.html, visited on January 21, 2004.

[3] The implicit assumption in the prototypical treatment model with a zero-one treatment and a homogenous treatment effect is that the optimal level of investment conditional on treatment status is the same for everyone, and therefore need not be modeled explicitly.

isolating the effects of the treatment that are specific to the agency. Under the assumption of a benevolent public agency, we identify general equilibrium treatment effects.

We find that larger firms and firms with higher value added current production have higher marginal profitability of R&D. Private rates of return are high and their distribution skew. Non-applicants' projects generate larger returns on investments, but applicants' projects generate larger joint rates of return on subsidies, defined as the sum of the applicant's and agency's benefits divided by the subsidy. The public agency obtains a return of 9% on its subsidy program, and in allocating the subsidies, the agency generally adheres to the publicly announced principles. In addition, larger firms get larger subsidy-percents, implying that moving a given R&D project to a larger firm will create a larger surplus that is not appropriated by the firm.

Our model allows us to identify application costs. We find that neglecting application costs causes a significant upward bias of the order of 70-90 percentage points in estimated joint rates of return. We also identify a potential selection problem: A positive shock to marginal profitability of R&D leads to an even larger positive shock to application costs. Thus firms with more profitable inventions are less likely to apply, ceteris paribus, due to higher opportunity costs of applying, creating a negative selection bias. This contrasts with the received view of a potential upward bias in the estimated effect of subsidies on R&D (e.g., Wallsten, 2000). Finally, there is evidence that previous contacts with the public agency reduce application costs.

As a counterfactual experiment, we studied what would happen if the agency put a larger emphasis on its internal grading of applications in deciding the subsidy levels. Doubling the effect of a one point increase on a Likert-scale in evaluated technical challenge on the granted subsidy generates a 16% median application probability for current non-applicants.

Our paper incorporates ingredients from several literatures. Methodologically we draw on the literature on treatment effects, structural labor supply (surveyed by Blundell and MaCurdy, 1999) and on structural industrial organization (surveyed by Reiss and Wolak, 2004) whereas our empirical application primarily relates to the literatures on R&D and the effects R&D subsidies. The closest papers in the treatment program literature are Heckman and Robb (1985), Maddala (1983, ch. 9), Manski (2000), and Heckman and Smith (2004) who stress the application and selection

decisions and Heckman, Smith and Taber (1996) who study how the objectives of the office holders affect the selection decisions (cf. Heckman, Heinrich and Smith, 1997). Prior to us, continuous treatment effects are theoretically modeled, e.g., by Heckman (1997), and Imbens (2000) and Lechner (2001) generalize the standard discrete zero-one treatment model into models of multiple treatment levels.[4] Heckman, Lochner and Taber (1998) and Davidson and Woodbury (1993) suggest procedures to identify general equilibrium treatment effects. Dehejia (2005), like us, models the selection decisions of the public agency. Our paper has also a link with the literature on revealed bureaucrat preferences (McFadden 1975, 1976).

The literature on the effects of R&D subsidies is extensive but, unfortunately, characterized by subtle points of inconclusive controversy (see David, Hall, and Toole, 2000, and Klette, Møen and Griliches, 2000, for surveys). A reason for conflicting results might be that the question has only recently been cast in terms of the treatment program literature (see e.g. Klette, Møen and Griliches, 2000 and Jaffe, 2002). Advance has also been hampered by lack of sufficient data. For example, the established but unsettled literature on the R&D-size relationship (see e.g. Cohen and Levin, 1989, and Cohen, 1995) relies almost exclusively on firm level data (an exception is Henderson and Cockburn, 1996). The only paper we know that studies the granting and application side of R&D subsidies is Blanes and Busom (2003). They estimate reduced form models of the joint application and granting decision. Their main finding that firms even in the same industry have different application thresholds both within and between the agencies supports our model and results.

We have access to rich data from Tekes (the National Technology Agency of Finland), the sole source of R&D subsidies in Finland. As will be described in detail in Section III, the data contain all the subsidy applications, 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 14 000 Finnish firms that constitute a large proportion of potential applicants.

A generic model of a treatment program is presented in Section II. We explain the institutional background and data in Section III. The generic model is then tailored to fit the institutional background in Section IV. There we also discuss identification

---

[4] Although Imbens (2000) does not explicitly further generalize his model, it is evident that it could also accommodate continuous treatments.

and estimation. Econometric results are reported in Section V, and implications of the model in Section VI. Our counterfactual experiment is discussed in Section VII and the conclusions are collected in Section VIII.

II. The theoretical model

There is a pool of potential applicants who have projects that require costly investments. The applicants need to decide whether or not to apply to a treatment program. A treatment lowers the marginal (shadow) cost of the investment the project requires, and all agents know the effect of the treatment. The program is run by a public agency whose utility function includes applicants' utility as an argument. The agency decides what treatment to give to each actual applicant, subject to constraints. For the moment, we do not allow for the screening and evaluation of project proposals by the agency but consider them in Section IV.

Our empirical application resembles what Jaffe (2002) calls a 'canonical' research grant program as our applicants are firms, they have R&D projects, the agency is Tekes (the National Technology Agency of Finland) and a treatment is an R&D subsidy. The generic model of this section, however, accommodates various other interpretations. For example, one can think of expected employment (the discounted utility thereof) as a project. According to Heckman and Smith (1995) the same treatments that educational programs such as JTPA offer are often available to those who do not participate in the program. One could thus model the effect of e.g. JTPA as reducing marginal costs of educational investment. The situation we model is also close to 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.

We model the above-described treatment program as a four-stage game of imperfect information between the applicant and the agency. In stage zero, nature draws the types of the players from a common knowledge prior distribution. In stage one, the applicant decides whether or not to apply to the program. In stage two, the agency decides the level of treatment, s, $s \in [0,1)$. As will be specified below, the treatment level is subject to minimum and maximum constraints if the application is accepted. The level is zero if there is no application or the application is rejected. In stage three, after receiving the treatment, the applicant makes the investment, denoted by R, $R \in [0,\infty)$.

Some of the assumptions we make below are made keeping in mind that our objective is to build a structural econometric model instead of a purely theoretical model.

A.1. Both the agency's and the applicant's utility functions are continuous and everywhere differentiable in their decision variables. The applicant's utility function is concave.

A.2. The applicant's type is common knowledge. The agency's type is private information.

A.3. The treatment cannot be misused.

A.4. There are no constraints on applicant's investment.

A.5. The agency's budget constraint does not bind.

A.6. The applicant's investment is non-contractible.

A.7. The level of treatment is $[\underline{s}, \bar{s}]$, $0 \leq \underline{s} < \bar{s} < 1$.

A.8. The level of the treatment is independent of the applied level.

A.9. All potential general equilibrium effects are captured by the agency's utility function.

A.1. ensures that the model behaves nicely. In particular, the applicant's best-reply to a treatment in stage three will be unique and given by a function. Because the agency's utility function is rather complex, we do not directly assume its concavity but use A.1. in the search for the conditions for a unique equilibrium of the game. The informational asymmetry in A.2. regarding the player' types generates equilibrium outcomes where the applicant applies for a treatment only to be turned down. We will posit that the agency's type is a parameter of a specific part of agency's utility function that captures the effects of the applicant's investment that affect the agency's utility, but not the applicant's. In practice, it may be neither desirable nor possible to make the agency fully transparent.[5] A.3. excludes moral hazard problems in use of treatment.[6] By A.4., the unique solution to the applicant's maximization problem in

[5] The alternative and perhaps more standard assumption would reverse the informational asymmetry so that the agency's type is common knowledge but the applicants' type is observed only by themselves. Although this would also generate rejected applications, it would lead to signalling, unnecessarily complicating the analysis. Moreover, the way we model the informational asymmetry is appealing in our empirical application because of the centralized subsidy allocation. A problem with our approach is that the agency could in principle give the optimal treatment without an application. Thus, we should strictly speaking assume that the applicant's type becomes common knowledge only upon the application. In so far the applicant cannot signal her type the assumption is inconsequential.

[6] In practice, moral hazard temptations are certainly pervasive with monetary treatments, as in our application. As a result, Tekes has several safe-guards against expropriation. For example, subsidies are

stage three is interior. This rules out credit rationing.[7] A.5 is motivated by simplicity.[8] A.6. is more realistic as it prevents the applicant and the agency from writing a binding contract specifying the amount the applicant invests conditional on the treatment it receives. A.7. corresponds with our application, and leads to a rather general way of modeling the treatment, and A.8. is imposed to avoid recursivity.[9] A.9. is a weaker form of the standard, heavily criticized (e.g. Heckman, Lochner and Taber, 1998), assumption in the treatment literature that allows no general equilibrium effects. In principle the agency should be a benevolent social planner that takes into account all effects of the treatment. If this were the case, our model would identify general equilibrium treatment effects.

We focus on perfect Bayesian equilibria where, in stage one, the applicant correctly anticipates the type contingent strategies of the agency in stage two, and where the applicant's and agency's strategies are sequentially rational. In this extensive form game the applicant's posterior belief concerning the agency's type is immaterial so there is no need to specify the belief formation. As a result, we can solve the game by backward induction, starting from the applicant's maximization problem in stage three. The applicant chooses the level of investment, $R(s)$, $R \in [0, \infty)$ to maximize

(1) $\pi(R(s),s)$.

$\pi(R(s),s)$ captures the expected discounted private utility from the project, with $\pi(0,.) = 0$. The first-order condition $\partial\pi(R(s),s)/\partial R = 0$ gives us $R^*(s)$, the applicant's best-reply function.

In stage two, the agency chooses the treatment s, $s \in [\underline{s}, \bar{s}]$, to maximize its expected discounted utility conditional on its type

---

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.

[7] This may sound like a strong assumption in the context of our application, as traditionally financial market failure has been a key rationale for R&D subsidies. However, a recent survey (Hall, 2002), study using Finnish data (Pajarinen and Hyytinen, 2003), and an evaluation of Finnish innovation policy (Georghiu et al., 2003) concluded that only small, R&D intensive, growth-oriented firms may face financial constraints. This is also reflected in our application: although Tekes also grants low-interest loans, most firms were not interested in them. Recently, the revealed motivations for R&D subsidies have become increasingly spillover-oriented.

[8] This is admittedly a strong assumption and we plan to take the agency budget constraint into account in future work.

[9] In particular, the assumption rules out the possibility of the actual treatment being a function of the proposed project size. We have tested the assumption, and our data do not reject it.

(2) $U(R(s),s) = V(R^*(s), \eta) + \pi(R^*(s),s) - gsR^*(s) - F$.

In (2), g ($g>1$) is the constant opportunity cost of agency resources, e.g., the opportunity cost of tax funds, and F is the sum of the applicant's fixed cost of applying and the agency's fixed cost of processing the application. The agency's type is given by η, which has a common knowledge probability density function $\phi(\eta)$ and cumulative probability density function $\Phi(\eta)$.

Equation (2) shows how the treatment s equals the share of the investment cost covered by the agency and how the applicant's utility directly enters the agency's utility function. The applicant's utility has an equal weight to $V()$ that captures the effects of the applicant's investment on the agency beyond the applicant's utility and the direct costs of treatment and the application process. Examples are externalities from firm R&D or from individual investments in human capital, and program-mandated payments to the agency (such as in JTPA). At the level of individual decision makers $V()$ can include idiosyncratic benefits from giving a certain treatment such as direct bribes or indirect ones, e.g., through a revolving door mechanism (see Laffont and Tirole, 1991, or Che, 1995). Such effects of the applicant's investment can also be negative or decreasing in the investment level. In what follows, we call $V()$ agency specific utility. As can be seen from (2), the informational asymmetry stems from this part of the agency's utility function. It is quite realistic to assume that the applicant does not know some characteristic of $V()$.

Because the agency is assumed to be subject to minimum and maximum constraints in choosing the treatment level (A.7), we form the Lagrangean

(3) $L(s,\lambda_1,\lambda_2) \equiv U(R^*(s),s) - \lambda_1(\underline{s} - s) - \lambda_2(s - \bar{s})$

and by using (2) write out the first order conditions:

(4a) $\frac{\partial L}{\partial s} = \frac{\partial V}{\partial R}\frac{dR^*}{ds} - gR^*(s) - gs\frac{dR^*}{ds} + \frac{\partial \pi}{\partial R}\frac{dR^*}{ds} + \frac{\partial \pi}{\partial s} + \lambda_1 - \lambda_2 \leq 0, s\frac{\partial L}{\partial s} = 0$

(4b) $\frac{\partial L}{\partial \lambda_1} = s - \underline{s} \geq 0, \lambda_1 \frac{\partial L}{\partial \lambda_1} = 0$

(4c) $\frac{\partial L}{\partial \lambda_2} = \bar{s} - s \geq 0, \lambda_2 \frac{\partial L}{\partial \lambda_2} = 0$.

Since the agency understands the applicant's reaction in stage two, the term $\frac{\partial \pi}{\partial R}\frac{dR^*}{ds}$ in (4a) is zero by the envelope theorem. As a result, the agency only needs to know

the direct effect of treatment on the applicant and the applicant's reaction function.[10] Although the informational requirements can sound challenging, we explain later how they are often less demanding than seems from the outset.

We are left with four effects of the treatment on the agency. The first effect is indirect through *V*(), the second and the third are the direct and indirect effect on the costs of the agency, and the fourth is the direct effect of the treatment on the applicant.[11] An interior solution to (4a) is thus given by $s^*(\eta)$ that solves $\frac{dR^*}{ds}\left(\frac{\partial V}{\partial R} - gs\right) + \frac{\partial \pi}{\partial s} - gR^*(s) = 0$. Noting that the sign of dR^*/ds is determined by the sign of $\partial^2\pi/\partial R\partial s$, we can loosely speaking distinguish four cases depending on whether the applicant's and the agency's investments are substitutes ($\partial^2\pi/\partial R\partial s < 0$) or complements, ($\partial^2\pi/\partial R\partial s > 0$), and whether the treatment is primarily motivated by the agency specific benefits ($\partial V/\partial R - gs > 0$) and ($\partial\pi/\partial s - gR^*(s) < 0$) or the applicant specific benefits ($\partial V/\partial R - gs < 0$) and ($\partial\pi/\partial s - gR^*(s) > 0$). For example, if the treatment is primarily motivated by positive externalities, we can have an interior solution only if the treatment and applicant's investments are complements. Similarly, if the treatment crowds out the applicant's investments, there can be an interior solution only for the applicant-motivated treatments.

To characterize the application decision, we assume that $\partial^2 V/\partial R\partial\eta > 0$ and that this is common knowledge. The applicant can then calculate the expected treatment to be

$$(5) \qquad E[s] = \Phi(\underline{\eta})\underline{s} + \int_{\underline{\eta}}^{\overline{\eta}} s^*(\eta)\phi(\eta)d\eta + [1 - \Phi(\overline{\eta})]\overline{s},$$

where $\underline{\eta}$ and $\overline{\eta}$ denote the values of η at which the minimum and maximum treatment constraints begin to bind and where $s^*(\eta)$ is an interior solution to (4a).

In stage one, the applicant decides to apply if the expected utility from applying is at least as large as from not applying. Because the expected utility from not applying is $\pi(R(0),0)$ the applicant's decision rule can be written as

---

[10] Given this knowledge the agency can infer the applicant's utility function up to the constant of integration. In our application, the constant equals the discounted expected profits from all other activities of the applicant, bar the R&D project for which the firm sought subsidies.

[11] In a social program such as JTPA, the direct effect on 'the applicant' could come from the reduced costs of attending an educational course. For example, Heckman and Smith (1995) explain how JTPA directs the participants of the program to take courses that are available also to non-participants with higher tuition fees. By lowering the cost of taking in a course, the program participants may choose more courses than identical non-participants.

(6)

$$\Phi(\underline{\eta})\pi(R(\underline{s}),\underline{s})+\int_{\underline{\eta}}^{\overline{\eta}}\pi(R(s*(\eta)),s*(\eta))\phi(\eta)d\eta+[1-\Phi(\overline{\eta})]\pi(R(\bar{s}),\bar{s})-K\geq\pi(R(0),0),$$

where K is the strictly positive cost of applying.

We now specify the conditions for the agency's utility function that guarantee a unique equilibrium. Let us define

$$f(s)\equiv\left(\frac{dR}{ds}\right)^2\left(\frac{\partial^2V}{\partial R^2}+\frac{\partial^2\pi}{\partial R^2}\right)+2\frac{dR}{ds}\left(\frac{\partial^2\pi}{\partial R\partial s}-g\right)+\frac{d^2R}{ds^2}\left(\frac{\partial V}{\partial R}-gs\right)+\frac{\partial^2\pi}{\partial s^2}.$$

PROPOSITION. If $f(s^*(\eta))<0$, there is a unique equilibrium characterized by the application rule (6), the optimal treatment level $\underline{s}$ for $\eta\leq\underline{\eta}$, $s^*(\eta)$ for $\eta\in(\underline{\eta},\overline{\eta})$ and $\bar{s}$ for $\eta\geq\overline{\eta}$, and the applicant's investment rule R^*(s).

Proof: In stage three, the applicant has a well-defined best-reply function R^*(s) because of A.1. In stage two, the agency maximizes its expected utility conditional on its type. There is a unique type-contingent optimal treatment if the second order condition for the Lagrangean (3) holds. Since we have linear constrains, it suffices to show that $U(R^*(s), s)$ is concave when evaluated at $s^*(\eta)$. Differentiating (2) twice, we see that $U(R^*(s), s)$ is concave if $\left(\frac{dR}{ds}\right)^2\left(\frac{\partial^2V}{\partial R^2}+\frac{\partial^2\pi}{\partial R^2}\right)+2\frac{dR}{ds}\left(\frac{\partial^2\pi}{\partial R\partial s}-g\right)+\frac{d^2R}{ds^2}\left(\frac{\partial\pi}{\partial R}+\frac{\partial V}{\partial R}-gs\right)+\frac{\partial^2\pi}{\partial s^2}<0$. The first term in the last parenthesis, $\partial\pi/\partial R$, is zero by the envelope theorem, which leaves us $f(s)$. If $f(s^*(\eta))<0$, there is a unique maximum that solves (4a-c). Since we assume that $\partial^2V/\partial R\partial\eta>0$, the optimal treatment level is increasing in η. Therefore the optimal type-contingent treatment is $\underline{s}$ for $\eta\leq\underline{\eta}$, $s^*(\eta)$ for $\eta\in(\underline{\eta},\overline{\eta})$ and $\bar{s}$ for $\eta\geq\overline{\eta}$. As a result, the applicant correctly anticipates that the expected treatment is given by (5) and makes the application decision according to (6). Because the type-contingent action of the agency in stage two is unique, the left-hand side of (6) has a unique value. In stage one the applicant either applies or does not apply, and there is thus a unique utility maximizing action in each stage of the game. *QED.*

It is rather hard to characterize when $f(s^*(\eta))<0$ without specifying functional forms. Since the second term in $f(s)$ is certainly positive if $\frac{\partial^2\pi}{\partial R\partial s}<0$, it is arguably

more likely that $f(s^*(\eta))<0$ when $\frac{\partial^2 \pi}{\partial R \partial s}>0$. This is the case in our empirical specification.

Note that $f(s^*(\eta))<0$ is only a sufficient condition for a unique equilibrium. For example, if $f(s^*(\eta))>0$, we also have a unique equilibrium where the optimal treatment level is always either the minimum or the maximum treatment and accordingly, the applicant's investment is always either $R(\underline{s})$ or $R(\bar{s})$. Given her knowledge of the agency type distribution, the applicant can again calculate the expected treatment and make optimal application decision using a rule similar to (6).

III. Finnish Innovation Policy, Tekes and Data[12]

A. Innovation Policy and Tekes

In 2001 Finland invested 3.6 per cent of GDP on R&D, which amounts to a total of 5 billion euros. Tekes is the principal public financier of private R&D in Finland.[13] 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 achieve these goals, Tekes strives to increase Finnish firms' R&D and risk-taking or "challenge" in their R&D projects. In addition to the primary criteria, Tekes has an explicit regional policy objective. Finnish regions differ greatly in their socio-economic characters, economic performance, and their R&D-intensity. E.g., some 20% of the population lives in the capital region in Southern Finland where also most economic activity and R&D takes place. As will be specified later, Tekes also treats firms fulfilling the official SME criterion differently.

Besides funding business R&D, Tekes finances feasibility studies, and R&D by public sector including scientific research. In 2001 Tekes received 2948 applications of which almost exactly 2/3 were accepted. In 2001 Tekes funding amounted to 387 million euros. The number of applications by the business sector for R&D funding was 1357 in 2001. Again, 2/3 of applications received funding. The

[12] As our application data is from Jan. 2000- June 2002, we use 2001 figures to describe the environment. This Section borrows heavily from Tanayama (2002). Tanayama spent six months in Tekes to get acquainted with the actual decision making process. Public information about Tekes can be found at http://www.tekes.fi/eng/, accessed December 20th, 2004.

[13] 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

applied funding for business R&D totaled 526 million euros while 211 million euros were granted.

Business R&D funding consists of grants, low-interest loans and capital loans. Tekes' low-interest loans have not only 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 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 (i.e., and it can be serviced only if there is distributable retained profit) 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 suggest that they do not encounter significant financial constrains, justifying our assumption A.4 (cf. footnote 7).

The application process from the submission to the final decision, which to our understanding is well known among potential applicants, has multiple stages. 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 screens the application and grades it in several dimensions by using a 6-point Likert scale from 0-5. According to Tekes civil servants, the most important dimensions in project evaluation concern the technological challenge of the project and its market risk.[14] In our data, only a 5-point scale was actually used in these dimensions. 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 (between firms) within the project, societal benefits, and the effect of Tekes funding. As mentioned above, Tekes also has an explicit regional policy objective and it takes into account whether the application comes from an SME.

---

support for innovation. See Georghiu et al. (2003) for a recent description and evaluation of the Finnish innovation policy institutions.

[14] 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'. See Tanayama (2002) for a detailed description of Tekes' screening and grading system.

Tekes' final decision is based on the proposed budget of the project before the R&D investments are made, but the actual funding is only given ex post against the incurred costs. Decision making is constrained by the rules preventing negative subsidies and very large subsidies both in relative and absolute terms. In other words, a subsidy is granted ex ante as a share of to-be-incurred R&D costs. There is an upper bound for this share: If the firm fulfils the EU SME criterion, the upper bound is 0.6, otherwise 0.5.[15] The actual funding then covers the contracted 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.[16] In terms of our model, these practices amount to $\underline{s} = 0$, $\bar{s} \in \{0.5, 0.6\}$ and a goal of setting the euro limit at $sR^*(s)$.

A part of the Tekes screening process is that it adjusts a proposed budget (both down and up) if an applicant, e.g., applies subsidies for costs that Tekes cannot cover. In practice an upward adjustment is rare and in principle occurs 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. We use this measure, which we call the 'proposed accepted investment', as our dependent variable in the R&D equation. We test the robustness of our results by using as an alternative dependent variable the R&D investment proposed by the applicant.

B. 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

[15] 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.

[16] As mentioned in footnote 6, 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 that 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.

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. The data are based on firms' official profit sheet and balance sheet statements, plus other information disclosed by the firms to public registries. We use the 1999 cross section, i.e., all firm characteristics are recorded earlier than the application data. The sample was drawn from Asiakastieto's registers 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 most likely to apply for funding from Tekes. After cleaning the data of firms with missing values, we are left with 10 944 firms. These firms form a large proportion of the population of potential applicants, and they constitute our sample of potential applicants.

Some 1000 firms from outside our sample filed applications to Tekes during the observation period. 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 have limited the count to one per firm[17] by using the first application by each firm within our observation period. 722 of these applications were accepted, i.e., received a positive subsidy share.

Table 1 displays summary statistics of our explanatory variables for potential applicants, and Table 2 conditions the statistics on the application decision, and the success of the application.

As Table 1 shows, potential applicants are heterogenous. They are on average 12 years old with 35 employees. A very high proportion are SMEs according to the official EU standard (cf. footnote 16). As explained, the SME criterion determines the upper bound of the share of the R&D costs covered by Tekes, and we therefore need

Table 1
Descriptive Statistics

	Mean	S.d.	Min.	Max.
Age, years	12.320	9.3453	1	97
# Employees	35.229	257.174	1	13451
Sales/employee, 1000 euros	164.920	2156.961	0	206875.5
Exporter	0.063	.244	0	1
SME	0.975	.157	0	1
CEO is chairman of board	0.141	.348	0	1
Board size	4.350	2.003	1	10
# past Tekes applications	0.575	3.488	0	146
Applicant	0.084	.277	0	1

NOTES: There are 10944 observations. Data sources: Asiakastieto Ltd. otherwise; for data on applications, Tekes.

to take it into account in our estimations. Sales per employee, a measure of value added, is 165 000 euros, and some 6% have exports.

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 more likely to include members with outside knowledge that may be useful either in conducting R&D (choosing among competing projects, organizing management of current projects, monitoring), or in the application process itself.

From Table 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 reports application- and Tekes-decision specific information. Some 21% of applications are rejected. The proposed projects involve on average an investment of 630 000 euros; the rejected proposals are clearly smaller with a mean of 385 000 euros. According to Tekes' rating, the projects have on average a technical

[17] Several firms in our data had multiple applications during our observation period, and taking them all into account, firms in our sample filed roughly half of all applications during our observation period.

Table 2
Conditional Descriptive Statistics

	Non-Applicants	Applicants	Rejected Applicants	Successful Applicants
Age	12.355	11.940	11.777	11.983
	(9.326)	(9.557)	(9.964)	(9.452)
	[10]	[10]	[9]	[10]
# Employees	21.200	189.001	101.269	212.453
	(122.282)	(775.862)	(187.503)	(866.674)
	[5]	[26]	[21]	[27]
Sales/employee	168.852	121.826	104.831	126.369
	(2252.692)	(54.996)	(94.238)	(167.307)
	[77.55]	[89.72]	[82.95]	[91.58]
Exporter	0.059	0.109	0.119	0.107
	(0.236)	(0.312)	(0.325)	(0.309)
SME	0.9860	0.850	0.855	0.849
	(0.1173)	(0.357)	(0.352)	(0.358)
CEO is chairman of board	0.141	0.149	0.176	0.141
	(0.348)	(0.356)	(0.382)	(0.349)
Board size	4.183	6.177	5.850	6.265
	(1.873)	(2.431)	(2.285)	(2.462)
	[4]	[6]	[5]	[6]
# past Tekes applications	0.247	4.163	3.228	4.413
	(1.283)	(10.657)	(10.933)	(10.576)
	[0]	[2]	[1]	[2]
Nobs.	10029	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.

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, and it is the same for successful and rejected applications.[18]

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. As some 60% of applicants only apply for a subsidy, and over 80% are only granted a subsidy, this seems a reasonable simplification. We test the robustness of our results to this definition by estimating the Tekes decision rule using only subsidies as the dependent variable. Using the pooled measure of subsidy, only 0.4% of applicants get the

[18] Since no applicant is assigned to category 5 in the market risk dimension, and only handful of applicants is assigned to category 0 in the technological risk dimension, we merge these categories with the ones next to them for estimation purposes.

Table 3
Descriptive Statistics of Tekes and Application Variables

	All Applicants	Successful Applicants	Rejected Applicants
Applied amount, euros	634294 (1254977)	700378.2 (1363460)	385790 (657539.8)
Applied for subsidy only	0.591 (0.492)	0.482 (0.500)	1.000 (0.000)
Technical challenge	2.088 (0.982) {582}	2.312 (0.872) {426}	1.474 (1.006) {156}
Risk	2.189 (0.937) {422}	2.150 (0.925) {326}	2.302 (0.937) {96}
Granted subsidy rate	-	0.316 (0.126)	-
Granted subsidy only	-	0.839 (0.600)	-
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.

maximum subsidy.[19] Successful applicants receive on average a subsidy that covers 32% of the R&D investment costs.

IV. The econometric model

A. The operationalization of the econometric model

We now operationalize the theoretical model by tailoring it to correspond to institutional details of our application explained in Section III and by making assumptions on functional forms and unobservables. It can be shown that the specified functional forms satisfy the sufficient condition for the unique equilibrium given in the proposition of Section II.

We approximate the screening process by assuming that Tekes grades each proposed project in two dimensions on a Likert scale of 5 and that this is common knowledge. The resulting 25 grade combinations are modeled by a latent regression framework. We assume that the error terms are normally distributed and uncorrelated both with each other and other unobservables of the model. Denoting the latent variable value of grading dimension *j* for application *i* by w_{ij}^* and the observed value by w_{ij}, we get:

[19] 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% .

$$w_{ij} = h \text{ if } \mu_{h-1,j} < w_{ij}^* = T_i\zeta_j + \omega_{ij} \le \mu_{h,j}$$

(7) $$h = 1,\dots,5,\ \mu_{0,j} \to -\infty,\ \mu_{1,j} = 0,\ \mu_{5,j} \to \infty$$

$$\omega_{ij} \sim \text{N}(0, 1),\ \text{cov}(\omega_{it}, \omega_{ij}) = 0 \ \ t \ne j\ .$$

where T_i is a matrix of observable characteristics of applicants, ζ_j is parameter vector to be estimated, and ω_{ij} is the unobservable applicant-specific component. Equation (7), when applied to the two evaluation dimensions, produces the probabilities of each of the 25 different outcome combinations. The application decision is then amended to

(6') $$\sum_{t=1}^{5}\sum_{k=1}^{5} p_t^C p_k^M \{\Phi(\underline{\eta} \mid t,k)\pi(R(\underline{s}),\underline{s}) + \int_{\underline{\eta}}^{\bar{\eta}} \pi(R(s(t,k,\eta)), s(t,k,\eta))\phi(\eta)d\eta$$
$$+[1-\Phi(\bar{\eta} \mid t,k)]\pi(R(\bar{s}),\bar{s})\} - K - \pi(R(0),0) \ge 0$$

where p_j^i is the probability of getting grade $j \in \{t, k\}$ in dimension $i \in \{C, M\}$ (C = technological challenge, M = market risk).

We specify applicant *i*'s objective function as

(8) $$\pi(R_i, s_i, X_i, \varepsilon_i) = \pi_{i0} + \exp(X_i\beta + \varepsilon_i)\ln R_i - (1-s_i)R_i\,,$$

where, in line with (1), s_i is the treatment and R_i is the investment. The marginal productivity of the investment is affected by observable applicant characteristics X_i, by vector β of parameters to be estimated and by ε_i, a random shock, distributed by nature, uncorrelated with the observable applicant characteristics, observed by the firm, and unobserved by the econometrician. The reservation value including other projects is embodied in π_{i0}. Although we are not able to identify π_{i0}, our cross section estimates are not affected by unobserved differences in the reservation value.

Equation (8) introduces unobservables into the applicant's objective function in an economically meaningful way. Unfortunately, it is theoretically possible that the objective function no longer satisfies $\pi \ge 0$ if π_{i0}, s_i, and ε_i are sufficiently small. This will be taken into account in the application equation later.[20]

We assume that the agency's utility form project *i* is given by

[20] We could generalize (8) 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.

(9) $U(R_i, s_i, X_i, Z_i, \varepsilon_i, \eta_i) = V(Z_i, \eta_i, R_i(s_i)) + \pi(X_i, \varepsilon_i, R_i(s_i), s_i) - gs_iR_i - F_i$,

where, similarly to (2), F_i is the sum of the fixed costs of applying and processing the application, g is the constant opportunity cost of the agency's resources, $V(.)$ is the expected agency specific utility from the project, and η_i is the random shock to it from project *i*. In line with the theoretical model the shock is assumed to be distributed by nature, uncorrelated with applicant characteristics, observed by the agency, and unobserved by the applicant and the econometrician. Compared to (2) and (8), the new term in (9) is Z_i, a matrix of observable applicant characteristics that affect the agency specific utility from the project. It may contain the same elements as X_i.

We solve the model backwards as in Section II. In stage three, the applicant optimizes (8) with respect to investment R_i. This yields

(10) $\ln R_i = X_i\beta - \ln(1-s_i) + \varepsilon_i$.

In stage two, the agency chooses a subsidy to maximize (9), taking (10) into account. To arrive at an estimable model we therefore need to specify the effect of R_i on $V(.)$. In the theoretical model we formalized A.2. by assuming that $\partial V/\partial R = E[\partial V/\partial R] + \eta$. We now take that $E[\partial V/\partial R_i] = Z_i\delta$ where δ is a vector of parameters to be estimated. As a result,

(11) $\partial V/\partial R_i = Z_i\delta + \eta_i$.[21]

A convenient implication of (11) is that

(12) $V = (Z_i\delta + \eta_i)R_i + \text{constant}$.

Equation (12) permits that the effect of the applicant's investment that is agency specific can be decreasing in the level of investment. For example, it is possible that some R&D projects exhibit negative externalities while being privately profitable. Equation (12) considerably facilitates dealing with double-censoring and sample selection. Moreover, it reduces the informational requirements for implementation of optimal subsidy decisions. To determine an optimal subsidy, the agency needs to know nothing about the applicant's objective function. However, the remaining informational requirements, in particular the fact that the agency should know $V()$, may still be challenging in practice.

[21] We empirically test whether $\partial V/\partial R$ is a function of R and cannot reject the null hypothesis that it is not.

By using (8), (10) and (11), the agency's unconstrained decision rule can be written as

(13) $s_i = 1 - g + Z_i\delta + \eta_i$.

As a result, the probability that an applicant gets the minimum subsidy is $\Phi(\underline{s}+g-1-Z_i\delta)$ and the probability of getting the maximum subsidy is $1-\Phi(\bar{s}+g-1-Z_i\delta)$.

As to stage one, the applicant decides whether or not to apply according to (6'). Our functional form assumptions together with our assumptions on unobservables create the possibility that utility maximizing investments lead to a negative utility. This would mean that the applicant prefers not to invest. This possibility can distort the decision rule. Because the possibility arises only if π_{i0}, s_i, and ε_i are all sufficiently small, and because there is a way out of this complication at the cost of a slight discontinuity at $R_i=0$, we proceed by using (6').[22]

We specify the fixed costs of applying as

(14) $K_i = \exp(Y_i\theta + \upsilon_i)$

where Y_i is a matrix of observable applicant characteristics, θ is a vector of parameters to be estimated and υ_i is a random cost shock, distributed by nature, uncorrelated with observable applicant characteristics, observed by the firm, and unobserved by the econometrician.

By using (8), (10) and (14) and the fact that $\underline{s} = 0$ as explained in Section III, (6') can after some algebra be rewritten in logarithmic form as

(6'')

$$X_i\beta - Y_i\theta - \ln\sum_{t=1}^{5}\sum_{k=1}^{5} p_t^C p_k^M \left\{ \int_{\underline{\eta}}^{\bar{\eta}} \ln\left(1-s(t,k,\eta)\right)\phi(\eta)d\eta - \left(1-\Phi(\bar{\eta}\,|\,t,k)\right)\ln(1-\bar{s}) \right\} \geq \nu_i - \varepsilon_i$$

Our econometric model can thus be summarized by *the screening equations* (7), *the Tekes decision rule*

(13') $s_i^* = (1-g) + Z_i\delta + \eta_i$, with observations $s_i = 0$ if $s_i^* \leq 0$ and $s_i = \bar{s}$ if $s_i^* \geq \bar{s}$,

---

[22] The theoretical way out would be to utilize the facts that i) $R_i = \exp(X_i\beta + \varepsilon_i)$ without subsidies and ii) $x\ln x - x$ is a convex function with a unique minimum of -1 at $x = 1$. By introducing a small discontinuity at R=0 into the applicant utility function by including the indicator function $1(R>0)$, one would ensure that an applicant always invests a positive amount, and that the increase in (expected) utility from investing is nonnegative. This change in the utility function would yield the current decision rule after subtracting one (euro) from the constant to get the true constant of the cost of application function. The corrected decision rule could be estimated using a simulation estimator. Our estimates show that negative profits from investment in the case of not applying are extremely unlikely (of the order 10^{-22}).

and the applicant's decision rule, i.e., *the investment equation*

(10’) $\ln R_i^* = X_i \beta - \ln(1-\bar{s}) + \varepsilon_i$.

We observe $\ln R_i = d_i \ln R_i^*$, where d_i is an indicator function determined by the application rule, i.e., *the application equation*, which completes the econometric model:

(6”’)

$$d_i = 1\left[X_i \beta - Y_i \theta - \ln \sum_{t=1}^{5} \sum_{k=1}^{5} p_t^C p_k^M \left\{ \int_{\underline{\eta}}^{\bar{\eta}} \ln\left(1 - s(t,k,\eta)\right)\phi(\eta)d\eta - \left(1 - \Phi(\bar{\eta} \mid t,k)\right)\ln(1-\bar{s}) \right\} \geq \nu_i - \varepsilon_i \right]$$

B. Statistical Assumptions, Identification and Estimation

We now explain our statistical assumptions, how identification takes place, and how we estimate the model. Our econometric model contains three unobservables, ε, η and υ . As mentioned, they are assumed to be uncorrelated with the observed applicant characteristics. In estimating the model by ML, we further impose

A.10 a) $\nu = (1+\rho)\varepsilon + \nu_0$, b) $\eta \perp \varepsilon$, c) $\eta \perp \nu_0$, d) $\varepsilon \perp \nu_0$

e) $\eta \sim N(0,\ \sigma_\eta^2)$ f) $\varepsilon \sim N(0,\ \sigma_\varepsilon^2)$ g) $\upsilon_0 \sim N(0,\ \sigma_{\nu_0}^2)$.

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 (υ) affecting the application cost. As A.10a) shows, there is no restriction on the correlation between υ and ε. Whether the agency decision rule (13’) suffers from sample selection depends on the correlation between the three error terms. Assumptions A.10b) and c) rule such correlations out. Although a priori these are strong assumptions, their validity is ultimately an empirical question, and we therefore test the validity of them below.[23] A.10e)-g) may be relaxed when we use semi- or nonparametric estimation methods.

The first estimation steps are the ordered probit the screening equations (7). By using the estimates we can calculate the predicted probability that a submitted application gets a particular grade in the two evaluation dimensions. Given our

[23] Because (13’) involves no sample selection problem, we do not need to multiply it by d_i, the indicator function determined by (6’’’). Note that without A.10b) we would have to employ a simulation estimator even if we imposed distributional assumptions on the error terms.

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 (13'). In estimation we use the actual values for the grades from the evaluation of each project. Tekes decision rule identifies δ, i.e., the effect of observed applicant and project characteristics on agency specific utility derived from the project. If we impose assumptions A.10b) and A.10c), we can estimate (13') using a double-hurdle Tobit model without correcting for selection. To test whether A.10b) and c) hold, we estimate a sample selection double-hurdle Tobit and test for the significance of the Mills ratio term. We also use an alternative, more flexible, approach of nonparametrically estimating (13') by a two-limit version of Powell's (1984) conditional least absolute deviations estimator.

After estimating the agency's screening equation and its decision rule, we calculate the effect of subsidies on the applicant's expected profits. We then estimate the investment and selection equations using both ML and the non-parametric approach suggested by Das, Newey, and Vella (2003).[24] The first stage is the selection equation (6'''). It allows us to identify how the observed applicant characteristics affect the fixed costs of application. Our theoretical model suggests a form for the error term in the selection equation and, as a result, we identify the correlation between v_i and ε_i when using ML. There is no need for a variance normalization as long as we, following theory, constrain the coefficient of the summand $\ln \sum_{t=1}^{5} \sum_{k=1}^{5} p_t^C p_k^M \left\{ \int_{\underline{\eta}}^{\bar{\eta}} \ln\left(1 - s(t,k,\eta)\right)\phi(\eta)d\eta - \left(1 - \Phi(\bar{\eta} \mid t,k)\right)\ln(1-\bar{s}) \right\}$ to unity.[25]

An implication of our theoretical model and its empirical counterpart is that an applicant strictly prefers reporting a proposed budget based on a maximum subsidy per cent over reporting any smaller amount, and is indifferent between

---

[24] Manski (1989) compares relative merits of the two 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 approach. Here theory comes to our aid, as it suggests an exclusion restriction that can be utilized both in parametric and nonparametric estimation.

[25] This implication of our theoretical model cannot be tested. If we imposed the standard variance normalization, the coefficient of the term would be $1/\sigma_{\upsilon}$ instead of unity.

proposing that amount and any larger amount.[26] Consequently, we can use the data on proposed budgets to estimate the investment equation (10') where we have inserted $\bar{s}$ into the equation. Correcting for selection bias using the application equation (6'''), we obtain consistent estimates of β that determine the effect of the observable applicant characteristics on the marginal profitability of the R&D-investment. To obtain consistent standard errors in the application and investment equations, we bootstrap the whole model ((6'''), (7), (10') and (13')) both when using ML and when using the semi/nonparametric estimator.

Note also what we cannot identify. We cannot identify separately *g*, the opportunity cost of government funds, and the constant in δ. Nor can we identify *V*(), as (13') cannot be integrated to a unique number. Welfare analysis is nonetheless possible, because our functional form assumptions ensure that all projects will be carried out irrespective of the subsidy decision. Thus each project will produce the fixed component in *V*() regardless of whether it is subsidized or not. [27] We are also unable to identify the agency's screening costs (F_i-K_i). This will result in an upward bias in the welfare calculations if these costs are significant. Finally, in the nonparametric estimation of the selection and investment equations, the parameters of the application cost function cannot be identified.[28]

Our model implies that the applicant's best-reply function, $\frac{dR_i}{ds_i} = \frac{\exp(X_i'\beta + \varepsilon_i)}{(1-s_i)^2}$, is increasing in treatment and is heterogenous both with respect to observables and the unobservable profitability shock.

---

[26] Too see this, recall first that the applicant does not know the Tekes' type (A.2) and the subsidy share is bounded above at $\bar{s}$ (A.7). As mentioned in Section IIIA, there is also an euro limit to the ex post reimbursements which is based on the proposed budget. Then note that since $\partial\pi/\partial s>0$ by (8), 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 even if 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.

[27] This is strictly speaking not true given the application decision rule suggested by theory. It, however, holds in the data, and would hold in the model if we introduced a discontinuity into the applicant's utility function as discussed in footnote 23.

[28] In standard treatment effect models researchers are not interested in these parameters. As our results below demonstrate, however, they should not be overlooked in the evaluation of treatment programs.

C. Treatment effects

The literature on treatment effects (see e.g. Blundell and Costa-Dias, 2002) emphasizes the effects of the treatment on potential applicants. In our case the effect of the subsidy is essentially given by (6’’’), which shows how it differs from the standard treatment effect by taking into account the cost of applying. In our case the expected treatment effect is also heterogenous if (un)observable applicant characteristics affect the marginal profitability of investment or the cost of applying.

An implication of our model is that a subsidy has effects on the agency beyond those on the applicant. Furthermore, if the agency were a benevolent social planner, $V()$ would capture all general equilibrium effects of a treatment beyond those appropriated by the applicant, and consequently the joint effect of the treatment on the agency and the applicant would constitute the social treatment effect.

We will thus calculate gross and net treatment effects, where the former refers to the standard calculation that does not take into account application costs. We further divide the treatment effects into private (firm), agency, and joint treatment effects, where the private gross treatment effect on the treated is the usually calculated one, agency treatment effects are the change in agency specific utility that is due to the treatment, and joint refers to the sum of private and agency treatment effects.

V. Estimation results

We include the following firm characteristics into all estimation equations: age, the log of the number of employees, sales per employee, an SME dummy, 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 include the squares of the continuous variables in our specification when estimating the investment and application equations. We also include region[29] and industry dummies. In the reported specifications, we use a slightly different set of explanatory variables in the screening equations and the Tekes decision rule on the one hand, and the selection and investment equations on the other.[30] We report the

[29] We divide Finland into five regions: Southern, Western, Eastern, Northern and Central Finland. Of these, Eastern and Northern Finland are the least developed.

[30] We did try interactions between firm characteristics and industry and region dummies. We used LR-tests to narrow down the set of explanatory variables in each equation. This was done in order to speed up computation of the bootstrap. The second order terms were excluded from the screening equations and the Tekes decision rule based on LR-tests.

results from the screening equations in the Appendix as they are not of material interest here.

A. Tekes decision rule

In Table 4 we report the estimation results concerning the Tekes decision rule. By using ML (column one) 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 (p-value 0.13) coefficient. Firm size obtains a positive and significant (at 10% level) coefficient. A possible interpretation is that in Tekes' view, moving an otherwise identical R&D project into a larger firm creates larger positive externalities e.g. through bigger employment rents. As against the 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 8.5 percentage points. The corporate governance variables and the number of previous applications have no effect.

We relegate the details of the coefficients of industry dummies to the Appendix. 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.

Another finding not reported in Table 4 (again, see the Appendix) is that regional aspects seem to influence Tekes' decision making: Firms in Eastern and Central Finland obtain subsidies that are 7-10 percentage points higher than they would obtain if they were 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. This finding is perhaps not robust as only 2% of our sample firms come from Northern Finland.

Table 4
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	-.018	-.020**	-.019
	[-.041 .005]	[-.039 -.001]	[-.048 .009]
Technical challenge	.100***	.094***	.120**
	[.076 .124]	[.074 .113]	[.090 .150]
Age	-.001	.0003	-.001
	[-.003 .001]	[-.0017 .0023]	[-.004 .002]
Log employment	.0164*	.024***	.031***
	[-.003 .036]	[.008 .040]	[.007 .055]
Sales/ employment	.000036	.000034	.000036
	[-.000136 .000276]	[-.000083 .000151]	[-.00017 .000243]
SME	.085**	.068*	.093*
	[-.001 .170]	[-.003 .138]	[-.011 .197]
Parent company	.006	.016	.014
	[-.040 .053]	[-.023 .055]	[-.043 .070]
# previous applications	-.001	-.002	-.003
	[-.006 .004]	[-.006 .002]	[-.009 .003]
CEO also chairman	.001	-.018	-.013
	[-.053 .055]	[-.064 .028]	[-.080 .055]
Board size	-.007	-.0001	-.009
	[-.017 .003]	[-.0084 .0082]	[-.021 .003]
Exporter	-.042	-.016	-.079*
	[-.107 .023]	[-.069 .038]	[-.161 .002]
Constant	-.060	-.103	-.197**
	[-.217 .098]	[-.233 .028]	[-.393 -.001]
σ_η	.189***	-	.225***
	[.173 .206]		[.203 .247]
Nobs.	379	379	379
LogL.	-18.636	-	-91.763
Wald	0.000	-	0.000
Linearity 1	0.690	-	-
Linearity 2	0.313	-	-
Sample selection	0.941	-	-

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.
Li nearity 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 selection = the p-value of a t-test on the Mills ratio from estimating a sample selection model where the first stage is 1(apply).
***, **, 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 euros, see main text) divided by accepted proposed investment. In column (3), the dependent variable is the subsidy (in euros) divided by the accepted proposed investment.

The above results are obtained under the assumptions A.10b) and A.10c), which maintain that the error in the Tekes decision rule uncorrelated with the errors in the investment and selection equations. To test these assumptions, we ran a first stage probit selection equation[31] and re-estimated the Tekes decision rule by inserting the Mills ratio into it. The Mills ratio obtained small negative (less than 0.2 in absolute

value) and very imprecisely estimated (p-values > 0.9) coefficients in all of the several specifications that we tried. This suggests that our assumptions A.10b) and A.10c) of no correlation are reasonable.

We also tested our assumption that *V*(), the agency specific utility, is linear in applicant investment. 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. After incorporating such terms into the Tekes decision rule, we could not reject the Null of (joint) insignificance of the terms.

We also estimated the Tekes decision rule by a two-limit version of Powell's (1984) conditional least absolute deviations (CLAD) estimator.[32] This allows for nonparametric estimation of (two-limit) censored regressions. The results are shown in column two of Table 4. The results are relatively close to those obtained using Tobit ML. The only significant 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 the Central Finland is no more significant. There are some relatively large differences between the insignificant coefficients, though.

Finally, 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 reveals that our results are not driven by our definition of the dependent variable.[33]

B. Cost of application function

In Table 5 we report the estimates of the application cost function (equation (14)).[34]

---

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

[32] 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.

[33] 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.

Table 5
Application Cost Function Results

Variable	Value
Age	.013
	[-.019 .273]
Age sq.	-7.375e-06
	[-.004 .0004]
Log of employment	-.381
	[-3.884 .125]
Ln(emp) sq.	.050
	[-.015 .418]
Sales/employee	.002**
	[.0004 .015]
Sales/emp. Sq.	-1.986e-07
	[-8.84e-07 1.61e-06]
SME	.236
	[-.609 3.750]
Parent company	-.127
	[-2.488 .226]
# Previous applications	-.221**
	[-3.877 -.019]
# Prev appl. sq.	.002**
	[.0002 .028]
CEO is chairman	-.326
	[-1.308 .222]
Board size	-.101
	[-1.406 .028]
Exporter	-.736**
	[-6.685 -.029]
Constant	11.830***
	[10.404 14.638]
Nobs	10751
LogL.	-18.636
Wald (d.f. 29) (p-value)	0.000

NOTES: Reported numbers are coefficient and [95% confidence interval]. Statistics refer to the probit 1st stage regression from the results of which the cost function coefficients have been backed out. 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 * denote significance at 1, 5, and 10% level.

We find that only few firm characteristics affect application costs. Age, size, and board size have no statistically significant effects. SMEs, firms where their CEO is also the chairman of the board, and parent companies do not have different costs. However, sales per employee increase application costs. One interpretation is thatfirms producing high value added products have complicated R&D projects based on soft information that are laborious to write down. Another is that the opportunity

[34] We only present results from the model where the log of proposed accepted investment was the dependent variable in the 2nd stage investment equation as results using the log of proposed investment yielded essentially identical results.

costs of the effort of making and promoting an application most likely are far greater than the direct monetary costs of filling in and filing it, and that firms with high value current production have higher opportunity costs of applying. Exporters have lower costs, maybe because they are relatively more experienced in dealing with government bureaucracy than non-exporting firms.

The number of past applications has a nonlinear effect, first decreasing and then, after 141 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 35%. One prior application decreases costs by 20% and four by 58%. It seems that learning by doing is going on. Given that our data is cross sectional, however, it is possible that instead of being attributed to path-dependence, the results are generated by unobserved heterogeneity.

C. Investment equation

Recall that our investment equation (10') identifies the effects of exogenous variables on marginal profitability of (log) R&D. In Table 6 we report two specifications: The first one is the same as the one used for the cost of application function. The estimates are rather imprecise, with only two reported variables carrying a significant coefficient.[35] Firms with higher value-added have higher marginal profitability of R&D whereas it is lower in firms with CEOs as chairmen.

In the specification without the quadratic terms, sales per employee and the CEO as chairman continue to carry significant coefficients. In addition, we find that larger firms, measured by the log of the number of employees, have higher marginal profitability of R&D. Henderson and Cockburn (1996), the only other study besides us that employs project level data, report a similar result.

To test the robustness of our results, we estimated the model using Das, Newey and Vella's (2003) nonparametric sample selection estimator. The results, presented in column three of Table 6, are in line with the ML estimates: Most coefficients are within the ML 95%confidence intervals. This suggests to us that our ML distributional assumptions are not biasing the parameter estimates. The propensity score carries a negative coefficient as expected (significant at 12.5% level).

---

[35] Several industry and region dummies carried significant coefficients, too.

Table 6
R&D Investment Function Results

Variable	(1) ML Dep. var. accepted proposed investment	(2) ML Dep. var. accepted proposed investment	(3) DNV Dep. var. accepted proposed investment	(4) ML Dep. var. proposed investment
Age	-.005	.002	.0001	-.005
	[-.027 .008]	[-.005 .006]	[-.030 .025]	[-.029 .006]
Age sq.	.0002	-	.0002	.0001
	[-.00003 .0005]		[-.0002 .0005]	[-.00007 .0004]
Log of employment	-.077	.042**	-.024	-.130
	[-.191 .234]	[.012 .134]	[-.362 .327]	[-.290 .203]
Ln(emp) sq.	.015	-	-.001	.022
	[-.022 .030]		[-.039 .036]	[-.017 .043]
Sales/ employee	.001***	0.0008***	.001**	.001*
	[.0001 .002]	[.0006 .001]	[.0003 .003]	[-.0002 .002]
Sales/emp. sq.	-1.95e-07	-	-2.9e-07	-1.53e-07
	[-7.31e-07 1.29e-06]		[-1.01e-06 1.33e-06]	[-6.10e-07 1.58e-06]
SME	-.258	-.281	-.011	-.063
	[-.561 .202]	[-.434 .011]	[-.766 .815]	[-.500 .350]
Parent company	.020	.066	-.091	-.035
	[-.134 .262]	[-.026 .210]	[-.438 .236]	[-.186 .173]
# Previous applications	-.047	-.006	-.295	-.047
	[-.061 .020]	[-.012 .001]	[-.748 .174]	[-.069 .004]
# Prev appl. sq.	.0003	-	.002	.0003
	[-.0003 .0005]		[-.005 .011]	[-.0001 .0006]
CEO is chairman	-.182*	-.198**	-.158	-.107
	[-.362 .003]	[-.336 -.069]	[-.368 .066]	[-.278 .080]
Board size	-.008	.008	-.065	.007
	[-.031 .049]	[-.005 .046]	[-.207 .086]	[-.021 .066]
Exporter	-.255	-.198	-.398	-.118
	[-.400 .037]	[-.301 .001]	[-.849 .162]	[-.280 .173]
Propensity score	-	-	-13.363[a]	-
			[-28.604 3.440]	
Constant	13.234***	12.401***	-	13.002***
	[10.920 13.638]	[11.224 12.475]		[10.965 13.428]
Nobs.	722	722	688	915
Wald (d.f.)	0.000	0.000	0.000	0.000
$\ln(1-\bar{s})$ coefficient (p-value)	0.158 (0.181)			-0.718 (0.740)

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. The SME dummy was excluded from the test regressions due to collinearity with $\ln(1-\bar{s})$.
***, **, *, and [a] denote significance at 1, 5, 10, and 15% level.

Table 7
Covariance Structure Results

Variable	Value
σ_{ε} Standard deviation of the investment equation shock	1.120*** [.834 1.256]
σ_{η} Standard deviation of the Tekes specific utility (=V()) shock	.189*** [.173 .206]
$\sigma_{\upsilon 0}$ Standard deviation of the uncorrelated part of the application cost function shock	.456*** [.111 12.552]
ρ Measure of the variance share of ε in υ	1.485*** [1.052 11.010]
$\rho_{\varepsilon\upsilon}$ Correlation between ε and the application equation error term	-.766*** [-.879 -.153]

NOTES: Reported numbers are coefficient and [95% confidence interval]. For all but σ_{η}, these 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.

In line with Das, Newey and Vella (2003), we interpret that there is evidence in favor of normal disturbances, because cross-validation (CV) suggests that no higher order terms of the propensity score are needed. [36]

Finally, we estimated the investment equation using the R&D investment proposed by the applicant as an alternative dependent variable. The results, presented in column four, are close to those in column one.[37] The one notable difference is that the coefficient of the CEO as chairman variable, although close in value, is no longer statistically significant. It thus seems that the definition of the dependent variable is not driving the results.

[36] We used the same trimming and transformation as Das, Newey and Vella (2003). The transformation gives exact sample selection correction for Gaussian disturbances. The trimming explains the difference in 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. CV indicated that we should include up 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). Only the 1st order propensity score variable obtained a significant coefficient, and CV confirmed that we only should use the 1st order propensity score. CV-values are reported in the Appendix. We used a Gram-Schmidt ortho-normalization for the 3rd and 4th order terms in both stages.

[37] The results using the restricted specification are close to those reported in column two.

D. Covariance structure

As Table 7 shows, we are able to identify precisely the variances of all error terms, and the covariance between the unobservables in the selection and investment equations. The coefficient determining the variance share of the unobservable of the investment equation in the unobservable of the application cost function (equation (14)) obtains a value of 1.5. 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. Similar to the finding that sales per employee increase application costs, it could be that projects with higher marginal profitability of R&D are more complicated involving tacit knowledge and are therefore more difficult to describe in an application. Moreover, the application costs are essentially opportunity costs, which should be higher for projects with higher marginal profitability of R&D.

VI. Implications of the results

We can utilize the structure of our model to back out a number of figures. In this section we report the following: i) the expected subsidy, ii) the expected marginal profitability of R&D, iii) the expected discounted profits from R&D with and without the expected subsidy, iv) the expected application costs, v) the increase in Tekes specific expected discounted utility from granting the actual or expected subsidy. This is the (gross) treatment effect on the agency, vi) the increase in expected discounted gross and net profits due to expected and actual subsidy. These are the expected (prior to application) and actual[38] private gross and net treatment effects on both the treated and the non-treated, and vii) the increase in gross and net expected discounted joint welfare due to actual subsidies. These are the joint gross and net treatment effects.

We exploit the information on unobservables that the covariance structure and the selection equation yield. The indicator function in (6’’’) takes value one for applicants and zero for non-applicants. We can then calculate values of the unobservables conditional on the value of the indicator function. For the R&D investment equation, we use the results in columns one and four of Table 6.[39] Those produced using the proposed R&D investment are presented in square brackets.

We report medians in Tables 8 and 9. The median expected subsidy is 16.6 per

[38] Actual in the sense of the treatment being realized. Naturally, these are still expected discounted effects.

[39] Using the results in column (2) made no essential difference.

Table 8
Implications of the Model

Entity	(1) Dep. var. log of accepted proposed investment	(2) Dep. var. log of proposed investment
Expected subsidy rate, all firms	.166	Same as in column (1)
Same, applicants	.195	Same as in column (1)
Same, non-applicants	.163	Same as in column (1)
Expected marginal profitability of log R&D, applicants	45228.72	62708.76
Same for non-applicants	199844.4	427592.4
Expected discounted profits from R&D w/o subsidies, euros, applicants	487485.9	673571.1
Same for non-applicants	2455946	5487330
Increase expected discounted gross profits due to expected subsidies, euros, applicants	12000.34	17231.23
Same for non-applicants	45882.42	98874.66
Increase expected discounted gross profits due to actual subsidies, euros, applicants	11079.47	16449.4
Expected application cost, euros, applicants	7657.862	11106.59
Same for non-applicants	326129.1	1567572
Increase in expected discounted net profits due to expected subsidies, euros, applicants	3208.659	5174.192
Increase in expected discounted net profits due to actual subsidies, euros, applicants	4217.539	6857.951
Tekes specific expected discounted utility (=V()) from the projects w/o subsidies, euros, applicants	17611.46	25183.04
Same for non-applicants	71308.45	152482.8

NOTES: Reported numbers are medians. Gross (Net) profits refers to gross (net) of application costs. The figures are calculated assuming g = 1.2 (g= shadow cost of public funds).

cent (19.5 for applicants, 16.3 for non-applicants), while the median actual subsidy is28 per cent. The distribution of expected subsidies is centered low, with a long right tail. The marginal profitability of log R&D expected by non-applicants is larger than by applicants by a factor of four. This is due to the positive correlation of the marginal profitability and application cost shocks: Applicants have smaller shocks and therefore lower profitability. The difference between the median marginal profitability expected by non-applicants and applicants is much smaller, if the application decision is not used to obtain information on the unobservable application costs. Similarly, the expected discounted profits on the non-applicants' projects are 2.5 [5.5] million euros whereas they are only on 0.5 [0.7] million euros on the applicants' projects.

Turning to the effects of treatments, the expected subsidies increase the applicants' median profits gross of application costs substantially less than the non-applicants' gross profits (12 000 [17 000] euros vs. 46 000 [99 000] euros). They increase the applicants' expected discounted net profits by 3 000 [5 000] euros, whereas the actual subsidies increase them by 4 000 [7 000] euros. Applicants also have considerably smaller median costs of application (7 700 [11 100])

Table 9
Implications of the Model c'ed

Entity	(1) Dep. var. log of accepted proposed investment	(2) Dep. var. log of proposed investment
Increase in expected discounted Tekes specific utility (=V()) from granting the expected subsidy, euros, applicants	15688.89	16820.58
Same for non-applicants	13181.76	14913.33
Increase in expected discounted Tekes specific utility (=V()) from granting the actual subsidy, euros, applicants	4155.703	6310.829
Increase in expected discounted joint gross welfare due to expected subsidies, euros, applicants	29740.05	35917.49
Increase in expected discounted joint net welfare due to expected subsidies, euros, applicants	20870.01	23129.74
Increase in expected discounted joint net welfare due to actual subsidies, euros, applicants	8605.993	13613.86
Tekes-specific utility's share of increase in gross welfare due to estimated subsidies	.564	.503
Tekes-specific utility's share of increase in net welfare due to estimated subsidies	.589	.766
Private rate of return on R&D w/o subsidies, applicants	9.719	10.739
Same for non-applicants	11.205	12.837
Joint rate of return on R&D w/o subsidies, applicants	12.547	12.151
Same for non-applicants	12.651	13.012
Joint rate of return on expected subsidies, gross, applicants	2.568	2.222
Joint rate of return on expected subsidies, gross, non-applicants	1.286	1.154
Joint rate of return on actual subsidies, gross, applicants	1.153	1.140
Joint rate of return on expected subsidies, net, applicants	1.716	1.459
Joint rate of return on actual subsidies, net, applicants	.785	.778

NOTES: Reported numbers are medians. Gross (Net) profits refers to gross (net) of application costs. The figures are calculated assuming g = 1.2 (g= shadow cost of public funds). Non-applicant application costs do not include unobservables, and are hence downward biased.

than non-applicants (330 000 [1 570 000] euros). We find that the applicants' projects generate a Tekes specific median expected discounted utility (w/o subsidies) of 18 000 [25 000], the corresponding utility from non-applicants' projects being 71 000 [152 000].[40] Using expected subsidies, applicants' and non-applicants' projects result in almost the same median increase in the Tekes specific utility (16 000 [17 000] and 13 000 [15 000] euros). The increase due to actual subsidies is 4 000 [6 000] euros. The median joint gross (net) treatment effect (welfare increase) is 30 000 [36 000] (21 000 [23 000]) euros for applicants using expected subsidies. Tekes captures 56% [50%] (59% [77%]) of the median increase in joint gross (net) welfare from those projects that are submitted.

We report the expected private and joint rates of return to R&D in Table 9.

[40] The calculations are based on the assumption that the shadow cost of taxes, *g*, is 1.2. Kuismanen (2000) estimates the dead-weight loss of existing Finnish taxation to be 15% using labor supply models. Both the constant of integration and the fixed costs of screening applications are ignored.

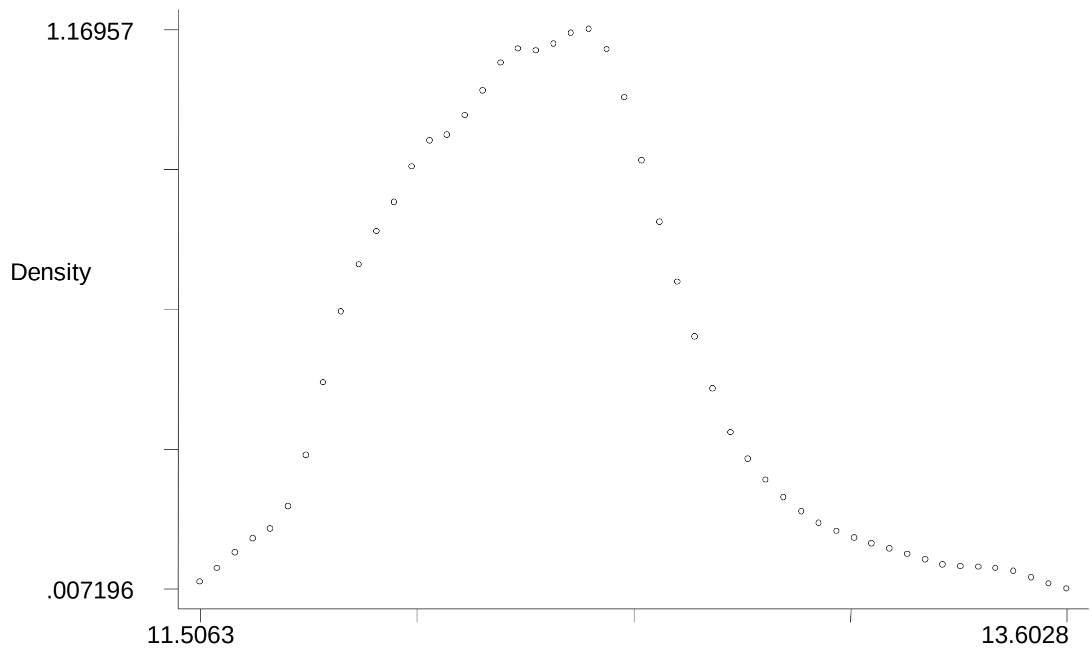

Figure 1: Private expected rates of return to R&D w/o subsidies, non-applicants

The estimated private returns are very high for applicants (median close to 900% [1000%]), and even higher for non-applicants.[41] Joint returns are appreciably higher, but the differences are dominated by the very high private returns. The private returns may seem too high for comfort even keeping in mind that these figures are based entrepreneurs' and firms' plans rather than on realizations. The relative dominance of private returns is more understandable, because Tekes operates in a small open economy from which most of the consumer surplus and firm spillovers flow abroad. [42] If Tekes is maximizing domestic welfare, it should ignore those effects, yielding the finding that private returns constitute a large part of joint returns. The distribution of private and, hence, joint returns to R&D, is skewed for non-applicants (see Figure 1), confirming earlier results (Pakes, 1986, Scherer and Harhoff, 2000).

Finally, we have calculated joint rates of return on subsidies (that on actual subsidies can only be calculated for applicants).[43] The median gross and net joint rate

[41] Even the minimum is high (750%) for applicants (0% for non-applicants).
[42] The literature on R&D, geography and trade (see e.g. Eaton and Kortum, 2002) finds that much of the spillovers are international.
[43] The joint rate of return is defined as the sum of agency specific utility and firm profits divided by subsidy amount in euros, where the subsidy amount in euros equals subsidy times the expected R&D investment, conditional on the subsidy.

of returns on the actual subsidy for applicants are 1.15 [1.14] and 0.79 [0.78]. The corresponding figures on the expected subsidy are 2.57 [2.22] and 1.72 [1.46] for applicants. For non-applicants, the median gross return on the expected subsidy is 1.28 [1.14]. The joint rate of return on the subsidy program is 9%, ignoring the opportunity cost of taxes. Returns using actual subsidies are lower because some firms get zero subsidies (no applicant expects to get zero), and some who would have generated very high returns if they had received expected subsidies, received lower subsidies and therefore generate lower returns.

The private and therefore also joint treatment effects, conditional on expected subsidies, are substantially lower for the applicants, while the agency treatment effects and joint rates of return are similar for applicants and non-applicants. The reason why applicants' projects are submitted to Tekes is that their application costs are much lower than those of the projects that are not submitted. Some privately and jointly profitable projects have very high private opportunity costs of applying. The results suggest that the average joint rate of return could be much higher if the positive correlation between application costs and marginal profitability could be lowered, because then the society would reap the large increases in private treatment effects.

VII. A counterfactual experiment

As mentioned in III.A, Tekes' tries to encourage 'risk-taking' in R&D projects. Our understanding of this objective is that Tekes encourages firms to take technically more challenging projects; this is verified by our estimation results that suggest that a one unit increase on the Likert scale regarding Tekes' perception of a project's technological challenge yields a 10 percentage points higher subsidy rate. We explore how potential applicants would behave if a one unit increase yielded a 20 percentage points higher subsidy rate.

As such a change would increase expected subsidies, all the firms already applying would continue to apply. They would receive a higher subsidy and, consequently, invest more in R&D. Moreover, some firms not applying under the current Tekes' decision rule would be induced to apply. We found that the median application probability for our sample's non-applicants is 16.4% after the policy change. The increase in expected discounted profits would double for both applicants to 24 000 euros and non-applicants to 88 000 euros. The median joint rate of return on

non-applicants' projects, gross of application costs, would increase by ten percentage points.

IIX. Conclusions

To gain understanding of how an R&D subsidy program works, we build and estimate a structural model of a treatment program. In our theoretical model an applicant decides whether or not to apply for a treatment and, subsequently, invests to maximize her utility, irrespective of whether or not she receives the treatment. The public agency running the program screens the application and provides a treatment to maximize its own utility, where the applicant's utility is an argument. The treatment is a continuous decision variable of the agency.

In formulating an econometric model, the theory yields a restriction that allows us to identify a sample selection model, an economic interpretation of the parameters of the model, predictions concerning the covariance structure of the model, and a generalization of the standard treatment effects. Both the theoretical and econometric models highlight the information the agency needs in its decision making. In our parameterization of the econometric model, the agency need know nothing about the utility function of the applicant to optimally choose a treatment. The econometric model provides a natural test of treatment effect heterogeneity, and enables the estimation of the role of observables and unobservables in the various treatment effects that have been studied in the literature.

Our empirical application concerns R&D subsidies in Finland. Such subsidies are a primary tool of industrial policy in the OECD countries. In Finland, where tax credits of R&D investments are not allowed, innovation policy hinges on subsidies. Their allocation is also centralized in Finland, giving a unique opportunity to evaluate their effects. We have access to all subsidy applications during two and a half years, which we match to balance sheet and other information on over 14 000 firms.

We find that subsidies increase the expected discounted profits of the firms by 3 000-5 000 euros in the median. The increase in agency specific utility has a median of 16-17 000 euros. The agency captures 50-55% of the increase in the utility that is caused by the subsidy.

Besides the effect of R&D subsidies, our structural model and estimations allow us to estimate the preferences of the agency running the subsidy program and to conduct counterfactual policy experiments. Our findings suggest that if a firm is

classified as an SME, it receives higher subsidies but otherwise the subsidy rate is increasing in the size of the firm. There is also some evidence that firms located in a less-developed area receive larger subsidies. The findings partially reflect the agency's focus on SMEs and regional policy. Our counterfactual experiment shows that doubling the effect of the perceived technical challenge of a project on the subsidy rate activates current non-applicants to apply with a median probability of 16%.

We also uncover the determinants of R&D and its rate of return by using project level data, which contributes to the so far inconclusive discussion on the R&D-size relationship that is primarily based on industry and firm level data. Our results support the side of the debate claiming that marginal profitability of R&D is directly related to firm size. We also find that marginal profitability of R&D is negatively affected if the CEO of a firm also acts as the chairman of its board, and positively by sales per employee. Private and joint returns to R&D are very high and their distribution skew. The median gross and net joint returns on subsidies are 15% and -21%. The joint return on the subsidy program is 9%.

In a departure from most of the literature, we can recover application costs. In our case at least this departure, made possible by the structural modeling of the application process, has a first-order effect. Taking application costs into account lowers the estimated rates of return by 70-90 percentage points. We also find a positive correlation between the unobservables of the application cost function and the marginal profitability of R&D, suggesting that firms with higher marginal profitability of R&D are less likely to apply for subsidies. The suggestion is further reinforced by our finding that sales per employee increase application costs. In line with these observations, non-applicants' application costs, expected discounted profits, and rates of return turn out to be higher than the applicants'. Our results suggest that the short-cut of ignoring application costs is recommendable neither in the research of treatment effects nor in practical policy making.

REFERENCES

Blanes, J., and Busom. I., 2003, Who participates in R&D subsidy programs? The case of Spanish manufacturing firms, forthcoming, Research Policy.

Blundell, R., and Costa-Dias, M., 2002, Alternative approaches to evaluation in empirical microeconomics, Portuguese Economic Journal, 1, 91-115.

Blundell, R., and MaCurdy, T., 1999, Labor supply: A review of alternative approaches, in Ashenfelter, O., and Card, D. (eds.), Handbook of Labor Economics, 3, Amsterdam: North-Holland.

Che, Y.-K., 1995, Revolving doors and the optimal tolerance for agency collusion, RAND Journal of Economics, 26, 378-397.

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

Cohen, W., and Levin, R., 1989, in Schmalensee, R., and Willig, R. (eds.), Handbook of Industrial Organization, 2, Amsterdam: North-Holland

Das, M., Newey, W., and Vella, F., 2003, Nonparametric estimation of sample selection models, Review of Economic Studies, 70, 33-58.

David, P., Hall, B., and 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.

Davidson, C., and Woodbury, S., 1993, The displacement effect of unemployment bonus programs, Journal of Labor Economics, 20, 380-388.

Dehejia, R., 2005, Program evaluation as a decision problem, Journal of Econometrics, 125,, 141-173.

Eaton, J., and Kortum. S., 2002, Technology, geography, and trade, Econometrica, 70, 1741-1779.

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

Hall, B. H., 2002, The financing of research and development, Oxford Review of Economic Policy, 18, 35-51.

Heckman, J., 1997, Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations, Journal of Human Resources, 32, 441-462.

Heckman, J., Heinrich, C., and Smith, J., 1997, Assessing the performance of performance standards in public bureaucracies, American Economic Review, Papers and Proceedings, 87, 389-395.

Heckman, J., Lalonde, R., and Smith, J., 1999, The Economics and Econometrics of Active Labor Market Policies, in Ashenfelter, O., and Card, D. (eds), Handbook of Labor Economics, Amsterdam: North-Holland.

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

Heckman, J., and 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., and Smith, J., 1995, Assessing the case for social experiments, Journal of Economic Perspectives, 9, 85-110.

Heckman, J. and 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., and 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.

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

Imbens, G., 2000, The role of propensity score in estimating dose-response functions, Biometrika, 87, 706-710.

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

Klette, T., Møen, J., and 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.

Laffont, J.-J., and Tirole, J., 1991, A Theory of Incentives in Procurement and Regulation, Cambridge, Ma: MIT Press.

Lechner, M., 2001, Identification and estimation of causal effect of multiple treatments under the conditional independence assumption, in Lechner, M., and Pfeiffer, F. (eds), Econometric Evaluation of Labour Market Policies, 43-58, Heidelberg: Physica/Springer.

Maddala, G., 1983, Limited-dependent and qualitative variables in econometrics, Econometric Society Monograph No. 3., Cambridge

Manski, C., 1989, The anatomy of the selection problem, Journal of Human Resources, 24, 343-360.

Manski, C., 2000, Identification problems and decision under ambiguity: Empirical analysis of treatment response and normative analysis of treatment choice, Journal of Econometrics, 95, 415-442.

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.

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

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

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

Pretschker, U., 1998, Public support to industrial R&D efforts, STI Review, 21, 91-104.

Reiss, P., and Wolak, F., 2004, 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., and 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.

Scherer, F., and Harhoff, D., 2000, Technology policy for a world of skew-distributed outcomes, Research Policy, 29, 559-566.

Tanayama, T., 2002, Selective public business funding – the case of Tekes, mimeo.

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.

Yatchew, A., 1998, Nonparametric regression techniques in economics, Journal of Economic Literature, 36, 669-721.

APPENDIX A

In this Appendix, we report the ordered probit estimation of the Tekes grading process; 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; industry and region dummy descriptive statistics and their coefficients for the estimated equations; and the cross-validation figures for the 1st and 2nd stage DNV estimations.

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.

A.1. The evaluation equations

Table A.1
Estimation of the Evaluation Equations

Variable	Technical Challenge	Risk
Age	.003 [-.007 .013]	-.0042379 [-.0164625 .0079868]
Log Employees	.008 [-.076 .092]	-.0536393 [-.1538962 .0466177]
Sales/employee	.001*** [.0002 .002]	-.0008665* [-.0017846 .0000516]
SME	-.101 [-.476 .274]	.0600485 [-.3851782 .5052751]
Parent Company	-.002 [-.206 .202]	-.1378355 [-.3769572 .1012863]
# Previous Applications	.021* [-.003 .044]	-.0189169 [-.045992 .0081582]
CEO is chairman	-.247** [-.487 -.006]	-.0118448 [-.2940517 .270362]
Board size	.078 [.034 .121]	.0331881 [-.0160126 .0823889]
Exporter	.170 [-.114 .454]	.2292716 [-.1084814 .5670247]
Nobs.	582	422
LogL.	-753.92882	-528.7958
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.

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 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 region 2 (Western Finland) also decreases the probability of being classified as high risk.

A.2. Descriptive statistics of the applicant samples

Table A.2 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 region 3 (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 A.2
Descriptive Statistics of Different Applicant Samples

Variable	All Applicants	Applicants with strictly positive accepted proposed investment	Applicants for whom grades in both evaluation dimensions are observed
Age	11.940 (9.557)	11.983 (9.452)	11.425 (8.961)
Log Employees	3.416 (1.787)	3.469 (1.786)	3.213 (1.684)
Sales/employee	121.826 (154.996)	126.369 (167.307)	120.252 (128.096)
SME	.850 (.357)	.849 (.358)	.879 (.327)
Parent company	.510 (.500)	.525 (.500)	.478 (.500)
# Previous applications	4.163 (10.657)	4.413 (10.576)	2.765 (4.545)
CEO is chairman	.149 (.356)	.141 (.349)	.174 (.380)
Board size	6.177 (2.431)	6.265 (2.462)	6.090 (2.367)
Exporter	.109 (.312)	.107 (.309)	.116 (.321)
Food	.0350 (.184)	.037 (190)	.032 (.175)
Paper	.0514 (.221)	.051 (.221)	.037 (.189)
Chemicals	.0317 (.175)	.035 (.183)	.026 (.160)
Rubber	.0623 (.242)	.061 (.239)	.061 (.239)
Metals	.0787 (.269)	.080 (.272)	.069 (.253)
Electric	.101 (.301)	.108 (.311)	.106 (.308)
Radio and TV	.040 (.197)	.039 (.193)	.047 (.213)
Other manufacturing	.0929 (.290)	.091 (.288)	.087 (.282)
Telecoms	.009 (.093)	.010 (.098)	.003 (.051)
Data processing	.207 (.405)	.197 (.398)	.259 (.438)
R&D	.148 (.355)	.147 (.354)	.129 (.336)
Region 2	.321 (.467)	.321 (.467)	.351 (.478)
Region 3	.115 (.319)	.125 (.331)	.058 (.234)
Region 4	.085 (.279)	.079 (.270)	.087 (.282)
Region 5	.022 (.146)	.019 (.138)	.029 (.168)
Nobs.	915	722	379

A.3. Descriptive statistics of the industry and region dummies for the whole sample

Table A.3
Descriptive Statistics of the Industry and Region Dummies for the Sample

Indicator	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)
Region 1 (Southern Finland)	.453 (.498)
Region 2 (Western Finland)	.386 (.487)
Region 3 (Eastern Finland)	.078 (.268)
Region 4 (Central Finland/Oulu region)	.061 (.240)
Region 5 (Northern Finland/Lapland)	.023 (.149)

Notes: there are 10944 observations.

A.4. Coefficients of industry and region dummies

Table A.4
Estimated Industry and Region Dummy Parameters

Variable	Tekes Decision Rule Table 4			Application Cost Function Table 5	R&D Investment Function Table 6			
Column	(1)	(2)	(3)		(1)	(2)	(3)	(4)
Food	.246***	.241***	.312***	.325	-.524***	-.606***	-.518*	-.522***
	[.122 .370]	[.137 .345]	[.163 .461]	[-.965 13.121]	[-.987 -.240]	[-1.00 -.269]	[-.968 .025]	[-.904 -.179]
Paper	-.017	.018	.0003	.085	.183	.013	.144	.183
	[-.140 .106]	[-.080 .116]	[-.1488 .1494]	[-1.169 1.913]	[-.335 .361]	[-.349 .343]	[-.395 .808]	[-.208 .525]
Chemicals	.094	.052	.132	.979	.163	.267	.232	.163
	[-.039 .228]	[-.060 .164]	[.029 .292]	[-.318 12.204]	[-.196 .789]	[-.170 .753]	[-.573 .889]	[-.413 .723]
Rubber	.012	.080	.008*	.228	.080	.099	.109	.080
	[-.084 .108]	[-.002 .162]	[-.111 .126]	[-.662 2.052]	[-.195 .434]	[-.213 .407]	[-.214 .542]	[-.267 .441]
Metals	.004	.013	-.014	.369[a]	.404	.231	.289	.403**
	[-.089 .095]	[-.063 .089]	[-.128 .100]	[-.217 2.842]	[-.0416 .512]	[-.067 .472]	[-.127 .708]	[.012 .658]
Electric	-.046	-.006	-.052	-.192	.254	.167	-.078	.254**
	[-.128 .036]	[-.076 .063]	[-.153 .050]	[-3.618 .597]	[-.066 .541]	[-.030 .540]	[-.678 .593]	[.019 .648]
Radio and TV	-.029	.011	-.001	.473	.603***	.621***	.486*	.603**
	[-.137 .078]	[-.077 .100]	[-.131 .128]	[-3.211 1.477]	[.238 1.184]	[.247 1.183]	[-.066 1.287]	[.082 1.197]
Other manufacturing	-.019	.013	-.016	.281	.206	-.050	.0002	.205
	[-.107 .069]	[-.060 .086]	[-.123 .092]	[-.574 3.803]	[-.353 .267]	[-.379 .217]	[-.391 .460]	[-.201 .472]
Telecoms	-	-	-	1.056[a]	.602	.514	.888*	.602
				[-.154 5.572]	[-.053 1.200]	[-.084 1.08]	[-.221 2.095]	[-.111 1.188]
Data processing	-.066*	-.028	-.058	-.432	.209	.172	-.199	.210**
	[-.140 .008]	[-.090 .033]	[-.150 .034]	[-5.276 .360]	[-.0797 .475]	[-.029 .484]	[-.917 .552]	[.017 .585]
R&D	.007	.049	.024	.178	.096	-.075	-.071	.096
	[-.073 .087]	[-.018 .117]	[-.075 .122]	[-.353 3.593]	[-.301 .243]	[-.286 .229]	[-.353 .251]	[-.178 .377]
Region 2 (Western Finland)	.018	.026	.019	.362*	.235**	.153**	.147*	.236***
	[-.028 .064]	[-.012 .065]	[-.038 .075]	[-.022 2.083]	[.0164 .336]	[.012 .328]	[-.011 .321]	[.090 .424]
Region 3 (Eastern Finland)	.096**	.088**	.145***	-.196	-.462**	-.374**	-.539**	-.462***
	[.007 .185]	[.014 .162]	[.037 .252]	[-.854 2.069]	[-.603 -.039]	[-.553 -.059]	[-.980 -.030]	[-.622 -.102]
Region 4 (Central Finland/Oulu region)	.069*	.031	.102**	.096	.062	-.034	-.175	.062
	[-.006 .145]	[-.030 .092]	[.010 .193]	[-.595 1.049]	[-.277 .272]	[-.246 .255]	[-.600 .242]	[-.193 .372]
Region 5 (Northern Finland/Lapland)	-.031	-.026	-.014	.168	.096	.281	.245	.096
	[-.158 .095]	[-.121 .070]	[-.170 .142]	[-2.891 1.272]	[-.056 .710]	[-.027 .715]	[-.188 .702]	[-.171 .507]

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. Region 1 (Southern Finland) is our base region.

A.5. Cross-validation

In the table below, we present the cross-validation figures for the application and the investment equations. Cross-validation figures were calculated using equation (2.22) in Yatchew (1998).

Table A.5
Cross-validation of the Application and R&D Investment Equations

Specification	Application Equation	R&D Investment Equation
Linear term	0.0595	0.7961
+2nd power	0.0602	0.7982
+2nd and 3rd power	0.0586	0.8006
+2nd -4th power	0.0635	0.8039
+ 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.