Do direct R&D subsidies lead to the monopolization of R&D in the economy?*

Dirk Czarnitzki [a, c, d] and Bernd Ebersberger [b]

a) K.U.Leuven, Dept. of Managerial Economics, Strategy and Innovation, Leuven (Belgium)
b) Management Center Innsbruck (Austria)
c) Centre for R&D Monitoring (ECOOM) at K.U.Leuven (Belgium)
d) Centre for European Economic Research (ZEW), Mannheim (Germany)

October 2010

Abstract

This paper explores the impact of R&D subsidies on the concentration of R&D in an economy. First, governments are often criticized of subsidizing predominantly larger firms and thus contribute to persistence of leadership in markets and higher barriers to entry, and, hence, reduced competition eventually. Second, theoretical literature, such as endogenous growth literature, has also shown that governmental intervention in the market for R&D affects the distribution of R&D which finally affects product market concentration. We test the relationship between R&D subsidies and R&D concentration employing treatment effects models on data of German and Finnish manufacturing firms. The data and estimations allow calculating concentration indices for the population of firms for both the actual situation where some selected companies receive R&D subsidies and the counterfactual situation describing the absence of subsidies. We find that R&D subsidies do not lead to higher concentration of R&D. On the contrary, we even find that R&D concentration is significantly reduced because of subsidies. This result may be attributed to the fact that technology policy maintains special funding schemes for small and medium-sized companies. The fact that the larger companies benefit from a higher likelihood of a subsidy receipt is offset by the phenomenon that smaller firms may be completely deterred from any R&D activity if they would not receive governmental support.

Keywords: R&D Concentration, R&D Subsidies, Treatment Effects, Policy Evaluation
JEL-Classification: O31, O32, O38

Address:	Dirk Czarnitzki	Bernd Ebersberger
	K.U. Leuven	Management Center Innsbruck
	Dept. of Managerial Economics, Strategy and Innovation	
	Naamsestraat 69	Universitaetsstr. 15
	3000 Leuven	6020 Innsbruck
	Belgium	Austria
Phone:	+32 16 326 906	+43 512 2070 3135
Fax:	+32 16 326 732	+49 512 2070 3199
E-mail:	dirk.czarnitzki@econ.kuleuven.be	bernd.ebersberger@mci.edu

* We thank the MIP team of the ZEW Mannheim for making the German survey data available. Czarnitzki gratefully acknowledges financial support from the Flemish Science Foundation (grant G.0534-07N). Ebersberger gratefully acknowledges the financial support of the Finnish Ministry of Employment and the Economy and the Finnish Funding Agency for Technology and Innovation. In addition, we thank the audience at following conferences and seminars for helpful comments: Pacific Rim Innovation conference, Melbourne (Australia); XXXIII Annual Conference of Industrial Economics and Policy, Ferrara (Italy); University of Strasbourg (France); Maastricht University (The Netherlands).

1 Introduction

It is a commonly held view among academic scholars and policy makers that the market for R&D is subject to market failures which – from a societal point of view – lead to an underprovision of R&D (see Arrow 1962 for a seminal contribution). This gives rise for governmental intervention in the market for R&D. While governments typically fund basic research in public research institutions, they also subsidize R&D in the private sector through different policies. The two most prominently used tools in industrialized countries are R&D tax credit schemes and direct R&D grants to companies applying for subsidies.

For decades, the predominant question in the R&D subsidy literature has been on potential crowding-out effects. There is some concern that subsidy recipients or beneficiaries of R&D tax credits simply substitute public for private investment so that the subsidy would not generate any additional R&D efforts in the economy. In a survey about the effects of R&D tax credits Hall and van Reenen (2000) conclude that the average dollar in tax credits increases private R&D spending by another dollar. In two influential surveys about the effectiveness of direct R&D subsidies David et al. (2000) and Klette et al. (2000) criticize the methodological approach of most of the earlier econometric studies as these did not account for self-selection of firms into subsidy programs. They highlight two points. First, firms that conduct R&D intensively are more likely to apply for funding. Second, the funding agency itself is interested in maximizing the outcome of its schemes typically by "picking the winner" strategies, that is, R&D intensive firms with past visible innovation success are preferred subsidy recipients. This results in the fact that the firms that would conduct plenty of R&D anyway also receive subsidies, and a mere comparison of R&D levels of subsidy recipients and non-recipients may misleadingly assign high, positive effects to the policy measure. Consequently, scholars recently applied treatment effects models that account for such selectivity in the funding process (see e.g. Busom, 2000, Almus and Czarnitzki, 2003, Gonzales et al., 2006, Görg and Strobl, 2007, Aerts and Schmidt 2008). While there are exceptions (see Wallsten, 2000, and Lach, 2002) the majority of more recent studies reject full crowding out effects at the level of the subsidized firms. Thus, a tentative conclusion from this literature is that public R&D grants lead to higher R&D in the economy.

The more recent microeconomic literature on the effects of R&D subsidies has already gone beyond the basic question of crowding out. For instance, Branstetter and Sakakibara (2002) explore the effects of publicly sponsored R&D alliances on patenting. Czarnitzki and Hussinger (2004), Czarnitzki and Licht (2006) and Hussinger (2008) investigate how increased R&D through subsidy translates into patents and new product sales using multiple equation models. Czarnitzki et al. (2007) analyze the heterogeneous treatments effects of R&D subsidies and R&D collaborations with regard to investments and patenting. So far the literature is basically limited to estimations of "treatment effects on the treated". An exception is Czarnitzki et al. (2007) who include estimations of treatment effects on the untreated. However, no study sheds light on macroeconomic implications obtained from estimations at the firm level.[1] This paper makes a contribution into this direction. A question that – to the best of our knowledge – has never been investigated so far is the effect of R&D subsidies on the distribution of R&D effort in the

[1] Yet, Almus and Czarnitzki (2003) extrapolate their estimated treatment effects to the population of firms. They find that R&D in the Eastern German economy would have dropped by 47% in absence of public R&D support after the German re-unification.

economy. The distribution of R&D bears important insights for policy makers and academics studying market structure and/or economic growth.

It has been observed in many studies that due to the selectivity in an agency's funding decision larger firms including market leaders are more likely to receive subsidies. This may lead to persistent dominance and thus monopolization of markets. As it is a commonly held view that especially small firms suffer more from financing constraints for R&D due to imperfect capital markets (see e.g. Himmelberg and Petersen, 1994, Hall, 2002, Czarnitzki and Hottenrott, 2010), policy makers have often been criticized of their distribution practices within the public grant systems. Although many industrialized countries maintain special funding schemes for small and medium-sized firms these days,[2] there are still calls for increasing public support for small and/or young companies in high-tech sectors (see e.g. Veugelers, 2009, Schneider and Veugelers, 2010).[3] Especially for financially constrained firms with no collaterals subsidies can help to alleviate the funding gap; even more so as the receipt of subsidies can also be regarded as a signal indicating high quality of the firm's innovation efforts (Lerner, 1999, Hyytinen and Toivanen, 2005, Takalo and Tanayama, 2010). This ongoing policy discussion suggests that in the absence of small firm programs policy intervention may have led to a higher concentration of R&D in the market with possibly detrimental effects for competition.

Furthermore, the literature on endogenous economic growth emphasizes the importance of R&D as one key driver of welfare enhancement. Starting with Dasgupta

---

[2] See http://www.proinno-europe.eu/trendchart for an overview of funding schemes in EU countries.
[3] This has led to a recent directive of the European Commission requesting the implementation for subsidy schemes for YICs (Young Innovative Companies).

and Stiglitz (1980a, b) where research is the central form of competition, numerous papers have elaborated on the relationship between R&D and market structure and their impact on endogenous growth. Examples are Aghion and Howitt (1992), Aghion et al. (2001, 2005), Perotto (1999), Thompson (2001), Klette and Kortum (2004) among many others. Laincz (2005, 2009) introduces R&D subsidies to models of growth and market structures, and analyzes how R&D subsidies affect market structure and ultimately growth. This leads to highly interesting results: Because of R&D subsidies, more firms remain active in the market as the price for investment declines. Consequently, concentration is lowered by the R&D subsidies on the one hand. On the other hand, however, the effect is offset by dominant firms taking advantage of their leadership position. Incumbent leader increase their technological lead over rivals and thus create higher barriers to entry and their profits rise. This turns the net effect of R&D subsidies on market concentration to be positive. It is important to note that Laincz (2005, 2009) finds that the concentration of R&D (not product market concentration) in the economy decreases because of the subsidy. However, Laincz's (2005, 2009) subsidies are set up in the way of an R&D tax credit. The government rewards R&D by paying some percentage of total cost of R&D to all R&D-performing firms. Empirically it remains an interesting question whether grant systems that subsidize only selected R&D projects lead to higher concentration as leaders are typically preferred.

In this study, we estimate Gini indices for R&D in the German and the Finnish manufacturing sector. In particular, we use data from the third wave of the Community Innovation Survey (CIS3). The sample data can be extrapolated to the population figures, and we estimate the Gini-coefficient of R&D concentration in the economy. This reflects the *actual* situation in the economy. Furthermore, we proceed by estimating a treatment effects model in order to derive the *counterfactual* situation, that is, what firms would have invested if they had not been subsidized. This hypothetical situation is then also extrapolated to the population level of manufacturing firms. With regard to the R&D subsidies, we find that full crowding-out can be rejected both in Germany and Finland. Furthermore, it turns out that the concentration of R&D is significantly lower in the actual situation when compared to the counterfactual where no R&D subsidies would be present.

The remainder of the paper is as follows: Section 2 describes the methodology used and section 3 describes the data. The results are discussed in the fourth section, and section 5 concludes.

2 Methodology

2.1 Inequality treatment effect

Firpo (2010) discusses three different inequality treatment effects. For the discussion we denote the outcome of the firm in the case it receives funding as Y^1. In the case the firm does not receive funding we denote the outcome as Y^0. The Gini index $G(F_Y^{\cdot})$ is a function of the distribution $F_Y^{\cdot}$ of the outcome variable. If F_Y^1 denotes the distribution of the outcome variable when all observations are treated[4] and F_Y^0 denotes the distribution of the outcome variable then $\Delta = G(F_Y^1) - G(F_Y^0)$ is the *overall inequality treatment effect* (ITE). Yet, in the context of public R&D subsidies the ITE is not informative about the effect of the funding program as the situation that all firms in the economy receive

[4] For the discussion we use *treated* and *subsidized* as synonyms.

subsidies for their innovation activities is politically and fiscally not feasible. The *inequality effect on the treated* (ITT) is defines as $\Delta^{ITT} = G(F_Y{}^1{}_{|S=1}) - G(F_Y{}^0{}_{|S=1})$, where $F_Y{}^1{}_{|S=1}$ is the distribution of the actual outcome of the treated and $F_Y{}^0{}_{|S=1}$ is the distribution of the outcome of the treated when they do not receive funding. The inequality effect on the treated only analyzes the effect of the funding on a subpopulation, where the economy wide effect of the funding programs cannot be estimated.

The *current inequality treatment effect* compares the actual distribution of the outcome variable F_Y with the situation when no funding scheme is implemented $F_Y{}^0$; $\Delta^{CIT} = G(F_Y) - G(F_Y{}^0)$. Note, that the current distribution F_Y of the outcome consists of the outcome of not treated firms (S=0) and of the outcome of treated firms (S=1). The outcome distribution in the situation where no funding scheme is available $F_Y{}^0$ consists of the outcome of the not funded firms when no funding is available. It also consists of the non treatment outcome Y^0 of the treated firms (S=1).

To estimate the current inequality treatment effect of public funding for innovation we require the non-funding R&D effort of the not funded firms $Y^0|S=0$ which can be observed. We also need the R&D effort of the funded firms $Y^1|S=1$ when they receive funding. It can also be observed. In addition we need information about R&D effort of the funded firms $Y^0|S=1$ under the counterfactual situation when no funding is received. However, cannot be observed; hence, it has to be estimated.

2.2 Estimation of the counterfactual

The estimation of the counterfactual has to account for the fact that receiving funding is not a random event, which in turn leads to selection bias. The literature on the econometrics of evaluation offers different estimation strategies to correct for this selection bias (see Heckman et al., 1999, or Imbens and Wooldridge, 2009, for surveys). For cross-sectional data, popular choices for treatment effects estimations are IV regressions, control function approaches (selection models) and matching estimators. In this paper, we employ a nearest neighbor propensity score matching.[5] The advantage of the matching is that no parametric model for the R&D equation has to be specified.

The counterfactual outcome of treated firms is constructed from a control group of non-treated firms. The matching relies on the intuitively attracting idea to balance the sample of program participants and comparable non-participants. Remaining differences in the outcome variable between both groups are then attributed to the treatment. Initially the counterfactual cannot simply be estimated as average outcome of the non-participants, because $E(Y^0|S=1) \neq E(Y^0|S=0)$ due to the possible selection bias. The subsidized firms and non-subsidized firms are expected to differ, except in cases of randomly assigned measures in experimental settings. Rubin (1977) introduced the conditional independence assumption (CIA) to overcome the selection problem, that is, participation and potential outcome are independent for firms with the same set of exogenous characteristics X. Phrased differently, the selection only occurs on observables:

$$Y^0 \perp S \mid X \tag{1}$$

If this assumption is valid, it follows that

[5] In the appendix, we also present the results of parametric treatment effects estimations. The choice of method does not influence our result on the treatment effect on the treated. We cannot apply IV techniques, however, as we do not have convincing instrumental variables available.

$$E(Y^0 \mid S=1, X) = E(Y^0 \mid S=0, X) \qquad (2)$$

The outcome of the non-participants can be used to estimate the counterfactual outcome of the participants in case of non-participation provided that there are no systematic differences between both groups. The average treatment effect can be written as

$$E(\Delta_{TT}) = E(Y^1 \mid S=1, X=x) - E(Y^0 \mid S=0, X=x) \qquad (3)$$

Conditioning on X takes account of the selection bias due to observable differences between participants and non-participants. In nearest neighbor matching for each one of the subsidized firms one picks the most similar firm from the potential control group of non-subsidized firms. In addition to the CIA, another important precondition for consistency of the matching estimator is common support. It is necessary that the control group contains at least one sufficiently similar observation for each treated firm. In practice, the sample to be evaluated is restricted to common support. However, if the overlap between the samples is too small the matching estimator is not applicable.

As one often wants to consider more than one matching argument, one has to deal with the "curse of dimensionality". If we employ a lot of variables in the matching function, it will become difficult to find appropriate controls. Rosenbaum and Rubin (1983) suggested to use a propensity score as a single index and thus to reduce the number of variables included in the matching function to just one. Therefore a probit model is estimated on the dummy indicating the receipt of subsidies S. The estimated propensity scores are subsequently used as matching argument. Lechner (1998) introduced a modification of the propensity score matching ("hybrid matching") as it is often desirable to include additional variables in the matching function. In this case, instead of a single X (the propensity score), other important characteristics may be employed in the matching function. The following matching protocol summarizes the empirical implementation of the matching procedure used in this paper.

Table 1: The matching protocol

Step 1	Specify and estimate a probit model to obtain the propensity scores $\hat{P}(X)$.
Step 2	Restrict the sample to common support: delete all observations on treated firms with probabilities larger than the maximum and smaller than the minimum in the potential control group. (This step is also performed for other covariates that are possibly used in addition to the propensity score as matching arguments.)
Step 3	Choose one observation from the subsample of treated firms and delete it from that pool.
Step 4	Calculate the distance between this firm and all non-subsidized firms in order to find the most similar control observation. As we match on the propensity score, we use a Euclidian distance. In case multiple matching arguments are used a standard choice is the computation of a Mahalanobis distance.
Step 5	Select the observation with the minimum distance from the remaining sample. (Do not remove the selected controls from the pool of potential controls, so that it can be used again.)
Step 6	Repeat steps 3 to 5 for all observations on subsidized firms.
Step 7	Using the matched comparison group, the *average effect on the treated* can thus be calculated as the mean difference of the matched samples (Note that the same observation may appear more than once in that group. As we perform sampling with replacement to estimate the counterfactual situation, an ordinary t-statistic on mean differences is biased, because it does not take the appearance of repeated observations into account. Therefore, we have to correct the standard errors in order to draw conclusions on statistical inference. We follow Lechner (2001) and calculate his estimator for an asymptotic approximation of the standard errors.).
Step 8	Now having available the distribution of the current R&D effort F_Y and the estimated distribution of the R&D effort in the case of no funding schemes F_Y^0 the current inequality treatment effect Δ^{CIT} can be estimated.

2.3 Estimation of Concentration Index

In order to test the impact of R&D subsidies not only on R&D spending among treated firms, but also on the concentration of R&D in the economy, we are interested in deriving measures of inequality, in particular the Gini coefficient (or the Lorenz curve respectively). We would like to compute an actual Lorenz curve based on the F_Y - the situation where R&D subsidies are in place, and a counterfactual Lorenz curve based on the distribution of R&D effort F_Y^0 where we simulate the absence of subsidies using our treatment effects results.

The literature on statistical inference for measures of inequality is still developing and so far there seems to be no commonly used approach. For instance, Mills and Zandvakili (1997) discuss the advantage of bootstrapping methods versus analytically derived asymptotic variance estimates. Deriving asymptotic variances of inequality measures can turn out to be quite cumbersome or computationally intensive. For example, jackknife procedures for deriving a standard error of a Gini coefficient were proposed by, among others, Sandström et al. (1985, 1988). Karagiannis and Kovacevic (2000) and Ogwang (2000) discuss ways of reducing the computational burden associated with the jackknife approximation to a level where this method can be applied even with very large samples. Ogwang (2000) also introduces a regression-based interpretation of the Gini coefficient which Giles (2004) used to derive standard errors without using jackknife procedures. He argues that the standard error of the OLS regression can directly be used for the computation of the standard error of the Gini index itself (see also Ogwang, 2004). Giles (2004) also discusses the usefulness of the regression-based approach with respect to hypothesis testing. Modarres and Gastwirth

(2006), however, conduct simulations and find that the standard errors of the Giles (2004) and Ogwang (2002, 2004) procedures are quite inaccurate. They recommend returning to the complex or computationally intensive methods. Giles (2006) and Ogwang (2006) do not disagree with the Modarres-Gastwirth criticism.

Consequently, we prefer an application of the percentile bootstrap method as suggested by Mills and Zandvakili (1997). Bootstrapping the probability intervals of the Gini coefficient has several advantages compared to using analytical standard errors. As the Gini coefficient is a non-linear function of a random variable, the interval estimates available from asymptotic theory may not be accurate and the small sample properties are not known. Furthermore, analytical standard errors cannot account for the fact that the Gini coefficient lies in the [0,1] interval, and thus may lead to estimations beyond the theoretical bounds. The bootstrap, in contrast, provides intervals drawn from a small sample distribution and takes into account the bounds of the Gini index. Mills and Zandvakili (1997) have shown that the percentile bootstrap method performs well in a variety of applications.

In addition, the percentile bootstrap method also allows us to incorporate sampling weights easily, and also accounts for the fact that our counterfactual Lorenz curve will be partly based on estimated data which leads to larger sample variation.

Most recently, Davidson (2009) suggested an approximation procedure for obtaining reliable standard errors of the Gini index, and also suggests using a bootstrap percentile t-method which provides an asymptotic refinement. Unfortunately, this requires the calculation of the standard error as Davidson suggests, and he does not provide information on how to incorporate sampling weights. Admittedly, Bhattacharya

(2007) derives asymptotic inference for stratified and clustered data, but as Davidson (2009: 30) points out, this "[…] is however not at all easy to implement". Therefore, we stick to the percentile bootstrap method as proposed by Mills and Zandvakili (1997).

For our application, we consequently compute the Lorenz curve, or Gini index respectively, for stratified random samples. This is straightforward as it amounts to a calculation of the Gini coefficient with clustered data. The actual Gini index, $G(F_Y)$, is computed based on the data of actual R&D employment as it is in the sample. In order to derive the counterfactual Gini index of inequality, $G(F_Y^0)$, we simply replace the actual R&D figure Y^1 of the treated firms (S=1) , with the counterfactual situation, $Y^0 = Y^1 - \Delta_{TT}$, which we obtain from the treatment effects estimation. As discussed above, we use the actual sample data for both the actual and the counterfactual R&D effort to compute the Gini index.

The only remaining choice to be made is the number of bootstrap replications, B. We follow Davidson and MacKinnon (2000) who suggest to choose B so that $\alpha(B+1)$ is an integer. Thus, for a significance level of $\alpha = 0.05$, it is recommended to choose at least $B = 399$. Then the critical values for the hypothesis test $\Delta^{CIT} = G(F_Y) - G(F_Y^0) = 0$ are given by the 10th and 390th value (lower and upper $\alpha/2$ quantiles) of the distribution of ${}_b\Delta^{CIT}$ (with $b = 1,\ldots,B$).

3 Data

3.1 Data source

The data basis for this analysis is the Community Innovation Survey (CIS). The CIS, launched in 1991 jointly by Eurostat and the Innovation and SME Program, aims at improving the empirical basis of innovation theory and policy at the European level by surveying innovation activities at the enterprise level in the Member States' economies. The CIS surveys collect firm-level data on innovation activities across member states by means of largely harmonized questionnaires. Thus the data are comparable on the European scale. The surveys are based on a stratified random sample of companies within the respective economies. In this analysis we use the third wave of the Community Innovation Survey (CIS3) for Germany and Finland which covers the years 1998 to 2000.[6]

The concepts used in the Community Innovation Survey closely reflect the definitions of the Oslo Manual (OECD 1997) and thus provide a good coverage to build indicators that can be used to analyze the effects of government intervention into corporate innovation. The analysis below will not only draw upon the Community Innovation Survey, but also construct a patent stock indicator capturing the corporate knowledge accumulation and patenting history of individual firms. We utilize data on patent applications with the Finnish and German patenting offices.

3.2 Variables

We analyze the effect of governmental intervention on the number of R&D employees in the firm (*RD*) and the R&D intensity (*RDINT*) measured by the fraction of R&D employees of all employees. The latter is not used for the calculation of the Gini index in the second step, but included for information only as this is the standard variable used by most of the previous studies. The innovation surveys contain information

[6] Detailed information about the design of the survey and the collection of the data can be found in Eurostat (2004).

whether the firm has received government funding. We utilize the dummy variable (*SUBS*) on whether individual firms received public funding for their innovation activities from international, national or regional sources. As control variables we use the size of the firm measured by the log of the number of employees (*LNEMP*), the export share of turnover (*EXQU*) and the accumulated knowledge measured by the number of discounted patents per employee (*PATSTOCK/EMPL*). Additionally we capture essential characteristics of the firm by dummy variables, whether the firm is part of a corporate group (*GROUP*), whether this corporate group is headquartered internationally (*FOREIGN*) and whether the firm has been established in the three years prior to the observation (*EST*). Sectoral characteristics are captured by the Herfindahl index of sales concentration at the NACE 3-digit level (*HHI*)[7], the level of R&D in the industry (*INDRDN*) and 10 industry dummies. Being targeted by special funding programs firms from Eastern Germany are identified through the dummy variable (*EAST*) in the German data set.

To arrive at estimates for the whole population of firms the sampling weights supplied with the survey data are used for the whole analysis.

3.3 Descriptive statistics

The variables used in the analysis are summarized in Table 2. The overall data set contains 1,000 observations in the Finnish sample and 1,403 observations in the German sample.

[7] Due to data availability the analysis of the Finnish sample uses the CR10 concentration ratio.

Table 2 Descriptive statistics of the used samples

	Finland (N=1,000)				
	Not Funded		Funded		
Variable	Mean	Std. Err.	Mean	Std. Err.	
RD	0.540	0.084	7.829	1.874	***
RDINT	0.006	0.001	0.057	0.006	***
LNEMP	3.461	0.039	3.908	0.079	***
PSTOCK/EMPL	0.004	0.001	0.017	0.003	***
EXQU	0.329	0.009	0.397	0.017	***
GROUP	0.324	0.022	0.383	0.034	
FOREIGN	0.071	0.011	0.087	0.015	
EST	0.074	0.013	0.109	0.022	
HHI	0.442	0.007	0.478	0.012	**
INDRND	1.033	0.171	3.529	0.932	***
P(*SUBS*)	0.214	0.005	0.337	0.014	***
	Germany (N=1,403)				
	Not Funded		Funded		
Variable	Mean	Std. Err.	Mean	Std. Err.	
RD	2.026	0.161	7.476	0.647	***
RDINT	0.023	0.002	0.090	0.007	***
LNEMP	3.761	0.035	4.121	0.078	***
PSTOCK/EMPL	0.011	0.002	0.015	0.002	*
EXQU	0.138	0.008	0.194	0.016	***
GROUP	0.222	0.016	0.312	0.036	**
FOREIGN	0.053	0.007	0.054	0.013	
EST	0.017	0.004	0.028	0.015	
HHI	39.547	2.367	42.185	4.543	
INDRND	2.076	0.119	3.085	0.253	***
EAST	0.115	0.009	0.309	0.031	***
P(*SUBS*)	0.157	0.004	0.277	0.012	***

Note: *** (**, *) indicates a significance level of 1% (5%, 10%). Weights used. *P*(*SUBS*) is the estimated probability of receiving subsidy. The probit regression for this estimation is reported in Table 3.

4 Results

Before we analyze the effects of the public subsidies on the mean R&D employees and the mean innovation intensity we report the results of the matching approach.

4.1 Matching

Table 2 compares the mean of the exogenous and the endogenous variables before matching. We observe significant differences in the R&D effort measured by the number of R&D personnel (*RD*) and by the R&D intensity (*RDINT*). We also observe that the funded companies are significantly larger (*LNEMP*), have higher patent stock per employee (*PSTOCK/EMPL*) and a higher export share (*EXQU*) than the not funded companies in the raw data set. Additionally, funded firms are more likely to be found in sectors with a higher concentration and a higher R&D intensity.

Table 3: Probit regression of the public funding dummy (weights used)

	Finland			*Germany*		
SUBS	Coef.	Std. Err.		Coef.	Std. Err.	
LNEMP	0.268	0.054	***	0.171	0.053	***
PSTOCK/EMPL	12.423	2.398	***	12.902	4.206	***
$(PSTOCK/EMPL)^2$	-16.892	4.397	***	-70.443	26.908	***
EXQU	0.552	0.212	***	0.172	0.288	
EAST	-	-		0.819	0.110	***
GROUP	-0.205	0.138		0.127	0.162	
FOREIGN	-0.236	0.164		-0.427	0.187	**
EST	0.251	0.188		0.296	0.443	
HHI	0.365	0.348		0.000	0.001	
INDRND	0.005	0.005		0.036	0.030	
CONS	-2.357	0.313	***	-2.448	0.304	***
No. of obs.	1,000			1,403		
Wald χ^2 (df^+)	100.07 ***			123.37 ***		
McFadden R^2	0.102			0.121		
Sum of weights	3,900			49,808		

Note: *** (**, *) indicates a significance level of 1% (5%, 10%). Regressions include 10 sector dummies not reported here, which are jointly significant. + Regression for Finland: df=19; regression for Germany: df=20.

Overall the Finnish (German) sample contains 1,000 (1,403) companies, of which 349 (343) receive public funding. Table 3 reports the probit regressions of the funding dummies. The results show that larger companies are more likely to receive public funding (*LNEMP*). For previous experience in successful invention (*PSTOCK/EMPL*) we find an inverse U-shape. When calculating the peak of the curve, however, it turns out that it is located at the right end of the patent per employee distribution for both countries. Thus we basically find a positive effect on the likelihood to receive funding, but with decreasing marginal returns to experience in inventive activity. Foreign ownership reduces the likelihood of receiving funding. In contrast to the German sample the effect is not significant in the Finnish sample.

We use the propensity score to receive public funding as the matching criterion; for each of the subsidized companies we match one company from the subsample of not-subsidized companies which is most similar in terms of the propensity to receive funding and which operates in the same sector. To achieve common support 13 (15) companies have to be dropped in the Finnish (German) data set. The weights of the matched non-funded companies (matched controls) are chosen such as to ensure that the sum of the weights of the matched controls equals the sum of the weights of the funded firms.

To illustrate the effect the matching procedure has on the composition of the sample we report the mean of the exogenous company characteristics before and after the matching procedure. Table 4 displays the mean of the exogenous characteristics. It reveals that the matching procedure succeeded in balancing the sample of not funded firms. On the average the sample of the funded companies and the sample of the not funded companies are the same with respect to the exogenous characteristics.

Table 4: Means of the exogenous variables after matching (weights used)

	Finland (N=674)			
	Counterfactual (not funded - matched)		Actual (funded)	
	Mean	Std.	Mean	Std.
LNEMP	3.994	0.081	3.890	0.079
PSTOCK/EMPL	0.016	0.003	0.014	0.003
EXQU	0.380	0.017	0.393	0.016
GROUP	0.487	0.038	0.378	0.034
FOREIGN	0.101	0.017	0.089	0.015
EST	0.085	0.016	0.111	0.023
HHI	0.495	0.012	0.475	0.012
INDRND	3.399	0.956	2.958	0.880
P(SUBS)	0.328	0.013	0.330	0.014
	Germany (N=656)			
	Counterfactual (not funded - matched)		Actual (funded)	
	Mean	Std.	Mean	Std.
LNEMP	4.193	0.082	4.109	0.079
PSTOCK/EMPL	0.013	0.002	0.015	0.002
EXQU	0.197	0.020	0.195	0.016
EAST	0.345	0.032	0.297	0.030
GROUP	0.318	0.032	0.309	0.037
FOREIGN	0.083	0.019	0.055	0.013
EST	0.039	0.015	0.025	0.015
HHI	43.441	4.039	42.30	4.633
INDRND	3.329	0.284	3.123	0.258
P(SUBS)	0.282	0.011	0.271	0.012

Note: *** (**, *) indicates a significance level of 1% (5%, 10%). The sum of weights in the Finnish (German) sample is 1,870 (16,060).

4.2 Treatment effects

Table 6 reports the mean effect of the public subsidy on the innovation input of the treated firms measured by the R&D intensity and measured by the number of R&D employees.

Table 6: Mean effect of the public subsidy on the R&D intensity and the R&D employees

Finland (N=674)	*RDINT*	*RD*
Countercatual (not funded - matched)	0.015	2.266
Actual (funded)	0.053	5.603
Δ_{TT}	0.038***	3.337***
Germany (N=656)	*RDINT*	*RD*
Countercatual (not funded - matched)	0.028	3.290
Actual (funded)	0.090	7.361
Δ_{TT}	0.062***	4.071***

Note: *** (**, *) indicates a significance level of 1% (5%, 10%). The sum of weights in the Finnish (German) sample is 1,870 (16,060).

For both the Finnish and the German sample we find significantly positive mean treatment effects for the funded firms. On the average the public subsidies induce an increase of the R&D input. This finding rejects complete crowding out of private R&D efforts by public subsidies.

4.3 Effects on the distribution

Having established the effect on the mean of the innovation input, we now turn to the effect public subsidies have on the distribution of innovation effort investigating the current inequality treatment effect Δ^{CIT} based on the estimated distribution of the number of R&D employees.

Table 8 summarizes the effects of the public subsidies on the distribution of innovation activity. It shows that with the subsidies the innovation activities are more equally distributed over the sample of companies than they would have been the case without the subsidies. Subsidies reduce the Gini coefficient of R&D effort in the Finnish manufacturing from 0.9413 to 0.9182. In the German manufacturing subsidies induce a reduction in the Gini coefficient from 0.890 to 0.8450. The reduction in the concentration of R&D effort is visualized by the Lorenz curves in Figure 1 and Figure 2.

For the German sample the distribution of the bootstrapped Gini-coefficients indicates that subsidies induce reduction in concentration which is significant at the 1 percent level of significance. Yet, the initial 399 bootstrapping repetitions suggest in the Finnish sample that the reduction in concentration is significant slightly above the 5 percent level of significance. With 799 bootstrapping repetitions we find it to be clearly significant at the 5 percent level.

Table 8 Effects of public funding on the distribution of innovative activity

Gini coefficient	*Finland*
Actual (with funding) $G(F_Y)$	0.9182
Counterfactual (without funding) $G(F_Y{}^0)$	0.9413
$G(F_Y) - G(F_Y{}^0) = \Delta^{CIT}$	-0.0231**
95% CI (Δ^{CIT})	[-0.0362, 0.001]
95% CI (Δ^{CIT})	[-0.0392;-0.001]$^+$

	Germany
Actual (with funding) $G(F_Y)$	0.8450
Counterfactual (without funding) $G(F_Y{}^0)$	0.8900
$G(F_Y) - G(F_Y{}^0) = \Delta^{CIT}$	-0.0451***
95% CI (Δ^{CIT})	[-0.0614;-0.0278]

Note: *** (**, *) indicates a significance level of 1% (5%, 10%). Confidence interval and mean based on 399 bootstrap replications; $^+$Confidence interval based on 799 bootstrap replications.

Figure 1 Lorenz-curve for the distribution of the R&D personnel (Finland)

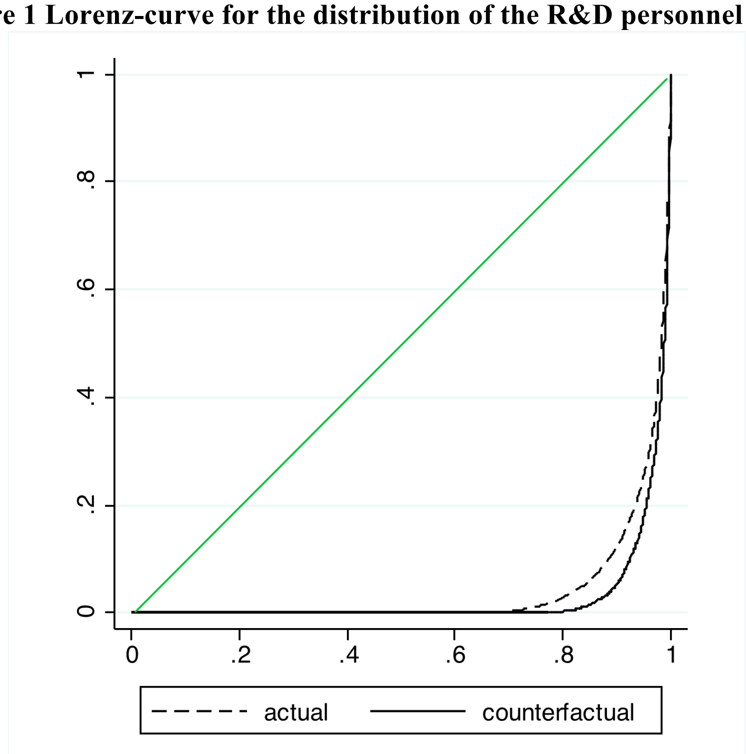
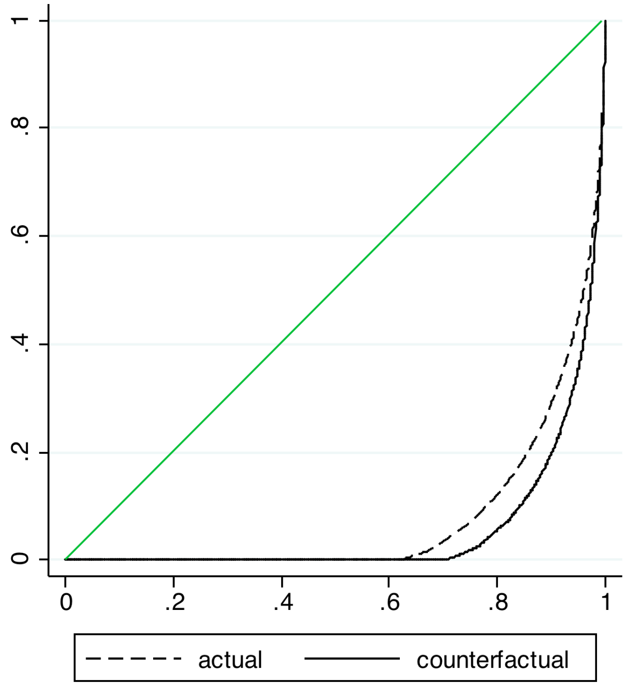


Figure 2 Lorenz-curve for the distribution of the R&D personnel (Germany)



5 Conclusion

This study has investigated the effects of R&D subsidies on corporate R&D in the Finnish and German manufacturing sector. In addition to calculating a treatment effect on the treated, we proposed to use this estimation for further macroeconomic implications. In particular, we were interested in the effect of subsidies on concentration of R&D in the business sector. As new growth theory has suggested, we find that public funding of R&D effort decreases the concentration of R&D in the economy. This, however, is by no means trivial. In contrast to the theoretical models that assume subsidies in form of R&D tax credits such that each R&D performer is reimbursed a certain percentage of its total R&D budget, we investigate two countries that do not employ R&D tax credits but subsidize innovation activity through a systems of direct grants. As the subsidies are based on proposal submissions of the firms and expert evaluations, such subsidy instruments are much more selective than R&D tax credits used in the theoretical literature.

We find that the German and Finnish subsidy systems do not lead to higher R&D concentration, moreover our results suggest that R&D concentration is significantly reduced by the subsidies. This gives rise to interesting policy conclusions. First, this result may be attributed to the fact that governments maintain special programs for small and medium-sized firms nowadays. These programs allow small firms to conduct R&D with the result that more firms are active in R&D. Financial constraints in the credit market may foreclose such investments otherwise. This effect can be seen in the Lorenz curves in Figures 1 and 2. In the counterfactual situation describing the absence of R&D policy, many more firms would not conduct R&D. The actual number of R&D performing companies would be reduced by 23% (33%) in Germany (Finland) if there had not been subsidies. This illustrates the fact that subsidies are effective in inducing R&D activities in firms which otherwise would not have performed R&D activities.

This fact outweighs the observation that larger firms are more likely to receive subsidies with respect to the concentration. Therefore, we also conclude that subsidies do not contribute to the long-run monopolization of R&D in the economy.

Of course, our study is not without limitations. We employ a nearest neighbor matching to estimate the treatment effect on the treated using two cross-sectional databases from Finland and Germany. It would certainly be interesting to apply more advanced methods such as conditional difference-in-difference estimation that would allow controlling for unobserved heterogeneity of firms. Unfortunately, we do not have panel data available for our study which limits the scope for using alternative estimation techniques. For further research it would also be interesting to apply quantile treatment effects in our context (see e.g. Abadie et al., 2002, Chernozhukov and Hansen, 2005, Firpo, 2007). However, the methods proposed by Abadie et al. (2002) and Chernozhukov and Hansen (2005) would require instrumental variables. Unfortunately, we do not have variables in our data that are convincing candidates for instruments, i.e., variables that determine the subsidy receipt but do in turn not impact on R&D.

References

Abadie, A., J.D. Angrist, and G.W. Imbens (2002), Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings, *Econometrica* 70(1), 91–117.

Aerts, K., and T. Schmidt (2008), Two for the price of one? Additionality effects of R&D subsidies: A comparison between Flanders and Germany, *Research Policy* 37(5), 806-822.

Arrow, K.J. (1962), Economic welfare and the allocation of resources for invention, in: R.R. Nelson (ed.), *The rate and direction of inventive activity: economic and social factors*, Princeton, N.J., 609-625.

Aghion, P., and P. Howitt (1992), A model of growth through creative destruction, *Econometrica* 60(2), 323-351.

Aghion, P., C. Harris, P. Howitt, and J. Vickers (2001), Competition, imitation, and growth with step-by-step innovation, *Review of Economic Studies* 68(3), 467-492.

Aghion, P., N. Bloom, R. Blundell, R. Griffith and P. Howitt (2005), Competition and innovation: an inverted U relationship, *Quarterly Journal of Economics* 120(2), 701-728.

Almus, M., and D. Czarnitzki (2003), The effects of public R&D subsidies on firms' innovation activities: the case of Eastern Germany, *Journal of Business and Economic Statistics* 21(2), 226-236.

Bhattacharya, D. (2007), Inference on inequality from household survey data, *Journal of Econometrics* 123, 674-707.

Branstetter, L.G., and M. Sakakibara (2002), When Do Research Consortia Work Well and Why? Evidence from Japanese Panel Data, *American Economic Review* 92(1), 143-159.

Busom, I., (2000), An Empirical Evaluation of the Effects of R&D Subsidies, *Economics of Innovation and New Technology* 9, 111–148.

Chernozhukov, V., and C.B. Hansen (2005), An IV Model of Quantile Treatment Effects, *Econometrica* 73(1), 245–261.

Czarnitzki, D., B. Ebersberger and A. Fier (2007), The Relationship between R&D Collaboration, Subsidies and R&D Performance: Empirical Evidence from Finland and Germany, *Journal of Applied Econometrics* 22(7), 1347-1366.

Czarnitzki, D., and H. Hottenrott (2010), R&D investment and financing constraints of small and medium-sized firms, *Small Business Economics*, forthcoming.

Czarnitzki, D., and K. Hussinger (2004), The Link between R&D Subsidies, R&D Spending and Technological Performance, ZEW Discussion Paper No. 04-56, Mannheim.

Czarnitzki, D., and G. Licht (2006), Additionality of Public R&D Grants in a Transition Economy: The Case of Eastern Germany, *Economics of Transition* 14(1), 101-131.

Dasgupta, P., and J.E. Stiglitz (1980a), Uncertainty, industrial structure and the speed of R&D, *Bell Journal of Economics* 11, 1–28.

Dasgupta, P., and J.E. Stiglitz (1980b), Industrial structure and the nature of innovative activity, *Economic Journal* 90, 266–293.

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

Davidson, R. (2009), Reliable inference for the Gini index, *Journal of Econometrics* 150, 30-40.

Davidson, R., and J.G. MacKinnon (2000), Bootstrap tests: How many bootstraps?, *Econometric Reviews* 19, 55-68.

Eurostat (2004), *Innovation in Europe Results for the EU, Iceland and Norway,* Luxembourg.

Firpo, S. (2007), Efficient Semiparametric Estimation of Quantile Treatment Effects, *Econometrica* 75(1), 259–276.

Firpo, S. (2010), Identification and Estimation of Distributional Impacts of Interventions Using Changes in Inequality Measures, IZA Discussion Paper Series 4841, Bonn.

Giles, D.E.A. (2004), Calculating a standard error for the Gini coefficient: some further results, *Oxford Bulletin of Economics and Statistics* 66, 425-433.

Giles, D.E.A. (2006), A cautionary note on estimating the standard error of the Gini index of inequality: Comment, *Oxford Bulletin of Economics and Statistics* 68, 395-396.

Görg, H., and E. Strobl (2007), The Effect of R&D Subsidies on Private R&D. *Economica* 74, 215-234.

González, X., J. Jaumandreu, and C. Pazó (2004), *Barriers to Innovation and Subsidy Effectiveness*, *Rand Journal of Economics* 36(4), 930-950.

Gonzalez, X., and C. Pazo (2008), Do public subsidies stimulate private R&D spending? *Research Policy 37*(3), 371-389.

Hall, B.H. (2002), The Financing of Research and Development, *Oxford Review of Economic Policy* 18(1), 35-51.

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

Heckman, J.J., R.J. Lalonde and J.A. Smith (1999), The Economics and Econometrics of Active Labour Market Programs, in: A. Ashenfelter and D. Card (eds.), *Handbook of Labour Economics*, Amsterdam: North Holland, Vol. 3, 1866-2097.

Himmelberg, C., and B. Peterson (1994), R&D and Internal Finance: A Panel Study of Small Firms in High-Tech Industries, *Review of Economics and Statistics* 76, 38-51.

Hussinger, K. (2008), R&D and Subsidies at the Firm Level: An Application of Parametric and Semi-Parametric Two-Step Selection Models, *Journal of Applied Econometrics* 23, 729-747.

Hyytinen, A., and O. Toivanen (2005), Do financial constraints hold back innovation and growth? Evidence on the role of public policy, *Research Policy* 34(9), 1385-1403.

Imbens, G.W., and J.M. Wooldridge (2009), Recent Developments in the Econometrics of Program Evaluation, *Journal of Economic Literature* 47(1), 5–86.

Karagiannis, E., and M. Kovacevic (2000), A method to calculate the jackknife variance estimator for the Gini coefficient, *Oxford Bulletin of Economics and Statistics* 62, 119–122.

Klette, T.J., and S. Kortum (2004), Innovating firms and aggregate innovation, *Journal of Political Economy* 112(5), 986-1018.

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

Lach, S., (2002), Do R&D subsidies stimulate or displace private R&D? Evidence from Israel, *Journal of Industrial Economics* 50, 369–390.

Laincz, C.A. (2005), Market structure and endogenous productivity growth: how do R&D subsidies affect market structure?, *Journal of Economic Dynamics and Control* 29, 187-223.

Laincz, C.A. (2009), R&D subsidies in a model of growth with dynamic market structure, *Journal of Evolutionary Economics* 19(5), 643-673.

Lechner, M. (1998), *Training the East German labour force: microeconometric evaluations of continuous vocational training after unification*, Heidelberg.

Lechner, M. (2001), Identification and estimation of causal effects of multiple treatments under the conditional independence assumption, in: M. Lechner and F. Pfeiffer (eds.), *Econometric evaluation of active labor market policies*, Heidelberg, 43-58.

Lerner, J. (1999), The Government as Venture Capitalist: The Long-Run Impact of the SBIR Program, *Journal of Business* 72(3), 285-318.

Mills, J.A., and S. Zandvakili (1997), Statistical inference via bootstrapping for measures of inequality, *Journal of Applied Econometrics* 12, 133-150.

Modarres, R, and J.L. Gastwirth (2006), A cautionary note on estimating the standard error of the Gini index of inequality, *Oxford Bulletin of Economics and Statistics* 68, 385-390.

Ogwang, T. (2000), A convenient method of computing the Gini index and its standard error, *Oxford Bulletin of Economics and Statistics* 62, 123–129.

Ogwang, T. (2004), Calculating a standard error for the Gini coefficient: some further results: Reply, *Oxford Bulletin of Economics and Statistics* 66(3), 435-437.

Ogwang, T. (2006), A cautionary note on estimating the standard error of the Gini index of inequality: Comment, *Oxford Bulletin of Economics and Statistics* 68, 391-393.

Perotto, P. (1999), Cost reduction, entry, and the interdependence of market structures and economic growth, *Journal of Monetary Economics* 43(1), 173-195.

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

Rubin, D.B. (1977), Assignment to treatment group on the basis of a covariate, *Journal of Educational Statistics* 2, 1–26.

Sandström, A., J.H. Wretman and B. Walden (1985), Variance estimators of the Gini coefficient: simple random sampling, *Metron* 43, 41–70.

Sandström, A., J.H. Wretman and B. Walden (1988), Variance estimators of the Gini coefficient: probability sampling, *Journal of Business and Economic Statistics* 6, 113–120.

Schneider, C., and R. Veugelers (2010), On young innovative companies: why they matter and how (not) to policy support them, *Industrial and Corporate Change* 19 (4), 969-1007.

Takalo, T., and T. Tanayama (2010). Adverse selection and financing of innovation: is there a need for R&D subsidies? *Journal of Technology Transfer* 35(1), 16-41.

Thomson, P. (2001), The microeconomics of an R&D based model of endogenous growth, *Journal of Economic Growth* 6(4), 263-283.

Veugelers, R (2009), A lifeline for Europe’s young radical innovators, *Bruegel Policy Brief* 2009(1), Brussels.

Wallsten, S.J. (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.

Appendix

In addition to the matching estimator presented in the main body of the text, we also experimented with the application of parametric treatment effects models based on the well-known sample selection model. In contrast to matching this also controls for selection on unobservable. A convincing application, however, requires an exclusion restriction, i.e. a variable that determines the subsidy receipt but not R&D. In our case, we do not have convincing variable in our data that could be justified as exclusion restriction on theoretical basis. Therefore, we could only follow an empirical approach and searched for a variable that is significant in the subsidy equation but not in the R&D equation. We found that the square of the patent stock per employee can serve as exclusion restriction in this case. As this is not fully convincing because we only rely on the non-linearity of the functional form (in addition to the non-linearity of the mills ratio), we relegated the results to the appendix. However, as the results in the following table show, full crowding out can also be rejected when selection on unobservables is taken into account.

Table 9: Parametric treatment effects models

	FINLAND						GERMANY					
	RD			*RDINT*			*RD*			*RDINT*		
	Coef.	S.E.		Coef.	S.E.		Coef.	S.E.		Coef.	S.E.	
LNEMP	2.100	0.266	***	-0.001	0.002		3.351	0.319	***	-0.010	0.002	***
PSTOCK/EMPL	3.339	4.100		0.242	0.104	**	27.012	7.447	***	0.081	0.057	
EXQU	4.322	0.715	***	0.010	0.009		1.506	1.093		0.037	0.013	***
GROUP	0.363	0.327		0.010	0.004	**	1.170	0.529	**	0.003	0.006	
FOREIGN	-0.844	0.815		-0.004	0.007		-1.982	1.133	*	-0.012	0.007	*
GRUEND	-0.545	0.416		-0.002	0.005		1.511	0.719	**	0.013	0.016	
HHI	-2.792	1.631	*	-0.009	0.016		-0.002	0.003		0.000	0.000	
INDRND	-0.003	0.021		0.001	0.001	**	-0.006	0.191		0.001	0.002	
EAST							-0.745	0.328	**	0.004	0.005	
SUBS	3.731	0.494	***	0.050	0.006	***	4.694	0.738	***	0.065	0.008	***
CONS	-7.221	1.196	***	0.000	0.009		-12.705	1.296	***	0.042	0.009	***
	SUBS EQUATION						*SUBS EQUATION*					
LNEMP	0.267	0.055	***	0.261	0.054	***	0.178	0.054	***	0.168	0.054	***
PSTOCK/EMPL	12.115	2.509	***	12.768	2.612	***	13.271	4.244	***	12.967	4.237	***
PSTOCK/EMPL2	-17.722	5.084	***	-19.355	5.415	***	-72.435	27.412	***	-69.875	26.966	**
EXQU	0.549	0.211	***	0.553	0.209	***	0.173	0.286		0.176	0.287	
GROUP	-0.217	0.137		-0.209	0.137		0.123	0.162		0.126	0.164	
FOREIGN	-0.226	0.165		-0.221	0.164		-0.428	0.186	**	-0.425	0.188	**
GRUEND	0.267	0.192		0.274	0.190		0.301	0.447		0.297	0.444	
HHI	0.336	0.343		0.346	0.336		0.000	0.001		0.000	0.001	
INDRND	0.005	0.005		0.005	0.005		0.036	0.030		0.037	0.030	
EAST	-			-			0.823	0.111	***	0.823	0.110	***
CONS	-2.510	0.313	***	-2.488	0.306	***	-2.480	0.307	***	-2.439	0.305	***
NOBS	990			1000			1403			1403		
WALD χ^2	129.1	***		169.06	***		270.63	***		225.09	***	
WALD $\rho=0$	17.23	***		5.99	**		6.11	**		1.13		

Note: *** (**, *) indicates a significance level of 1% (5%, 10%). 10 industry dummies are included in the regressions. In the regression of R&D personnel for the Finnish data set the 1% largest R&D performing firms are excluded.