Cappelen, Å., Raknerud, A. & Rybalka, M. (2012). The effects of R&D tax credits on patenting and innovations. *Research Policy*, 41(2), 334-345. https://doi.org/10.1016/j.respol.2011.10.001

The effects of R&D tax credits on patenting and innovations

Ådne Cappelen, Arvid Raknerud [*], Marina Rybalka

Research Department, Statistics Norway, PO Box 8131 Dep, 0033 Oslo, Norway

Abstract
Norwegian business spending on R&D is low by OECD standards. To stimulate business R&D, in 2002 the Norwegian government introduced a tax-based incentive, SkatteFUNN. We analyze the effects of SkatteFUNN on the likelihood of innovating and patenting. Using a rich database for Norwegian firms, we find that projects receiving tax credits result in the development of new production processes and to some extent the development of new products for the firm. Firms that collaborate with other firms are more likely to be successful in their innovation activities. However, the scheme does not appear to contribute to innovations in the form of new products for the market or patenting.

Keywords: Tax credits; R&D; Patenting; Innovation; Self-selection.

[*] Corresponding author. Tel.: +47 21094445. E-mail: arvid.raknerud@ssb.no.

1. Introduction

Both economic theory and empirical evidence support the view that R&D plays a vital role in raising productivity on a sustainable basis (see, e.g., Griliches, 1992, and Romer, 1990). The social return on R&D investment is often higher than the private return to the investing firm. Thus, one can justify policy interventions if a well-designed intervention scheme can be implemented. Since the 1990s, OECD countries have tended to rely on fiscal policy incentives to promote R&D spending in the business sector. In 1996, 12 OECD countries offered tax incentives for R&D expenses; by 2004 this number had increased to 18, with Norway as one of the newcomers.

R&D incentives are designed in many different ways. Some countries offer incremental schemes targeting only increases in R&D expenses, while others have volume-based incentives. A few countries have both. Although more countries have introduced tax incentives over time, no consensus exists as to what is best practice. Evaluation of the incentives in various countries may help determine which policies or policy mixes work well.

R&D spending in the Norwegian business sector as a share of GDP is below the OECD average. To stimulate private R&D investment, the Norwegian government has traditionally used direct R&D subsidies. In 2002 this policy was supplemented with an R&D tax credit scheme — SkatteFUNN — for small and medium-sized enterprises (SMEs), which by 2003 became available to all firms.[1] SkatteFUNN provides a volume-based tax credit to firms with an R&D project that the Research Council of Norway (RCN) has approved. A tax credit of 18 percent (20 percent for SMEs) of R&D costs for the approved project is deductible from the firm's income tax, with a project cost cap roughly equal to half a million Euros. If the firm does not pay any tax or pays less tax than the tax credit, the credit is paid to the firm as if it were a grant. The scheme provides an incentive for increased R&D mainly to firms with small R&D projects for which the cap in the tax credit is not binding. Thus it lowers the *marginal*

[1] The SkatteFUNN scheme is one of the most generous tax credit systems among OECD countries according to OECD (2007, cf. Figure 5.1). The tax credit contains no sector or regional bias and is thus neutral between qualifying projects; and region, sector and the tax position of qualifying firms. Even if SkatteFUNN is available to all firms, it induced a significant increase of R&D subsidies to small firms. The official R&D statistics show that the Norwegian government spent 2.1 billion NOK (250 mill. Euro) on public subsidies to R&D in 2008 (cf. RCN, 2010). About 1 billion was given as tax subsidies through the SkatteFUNN scheme, the rest, 1.1 billion, was given as direct grants. About 40 percent of the SkatteFUNN subsidies (400 mill. NOK) were given to firms with less than 5 employees in 2008.

cost of R&D for low spenders more than for firms with large R&D expenditures. A summary of main findings from the evaluation of the scheme is given in Cappelen et al. (2010).

In this paper we study the effects of SkatteFUNN on firms' innovation activities and patenting. We analyze three types of innovations: a new (or improved) product for the firm, a new (or improved) product for the market, and a new (or improved) production process. We also have information on patent applications. We focus on the following three questions which are of general interest for research policy.

First, how is the introduction of innovations related to R&D? While R&D obviously is an important factor behind innovation, it is not the only one. The availability of high–skilled workers is another important factor. Our data reveal that formal R&D is not a necessary condition for innovation.

Second, does SkatteFUNN lead to more innovations? Hægeland and Møen (2007) find that firms receiving support through SkatteFUNN are more likely to increase their R&D investments than other firms. The question remains whether there is a causal relation between SkatteFUNN and firms' innovations.

Third, does the answer to the third question depend on type of innovation? One important reason for government intervention in the market for R&D is to create spillovers. If firms receiving subsidies mainly innovate in the form of products that are new for the firm, but not for the market (i.e., imitate other firms), it is not clear that the scheme will reach those R&D activities with the largest potential for spillovers. Thus, to assess the impact of tax credits on innovation is important from the point of view of research policy. To our knowledge, the existing literature has addressed this issue only to a limited degree.

For the analysis, we use Norwegian micro data covering firms included in the 2001 and 2004 innovation surveys (CIS3 and CIS4). These surveys contain information on the inputs and outputs of firms' innovative activities, e.g., whether firms have introduced product or process innovations and whether they have applied for a patent over the three-year period before each survey. The 2001 survey covers the three years before the introduction of SkatteFUNN (1999–2001), while the 2004 survey covers the three years following the introduction of SkatteFUNN (2002–2004). Since about 2/3 of the firms are included in both surveys (see Section 3), we are able to obtain a panel data set from these survey data. By supplementing the data with information from the (annual) R&D survey and different registers, we obtain two three-year periods with data on the innovation variables and six years of observations on R&D and other variables.

Our modeling framework is influenced by Griliches (1990), Crepon et al. (1998) and Parisi et al. (2006). The main idea in this literature is that by investing in R&D, the firm accumulates R&D capital, which plays an important role in its innovation activities. Using binary regression models, we model the probability of innovating and patenting as function of the firm's R&D capital stock at the *beginning* of each three year period, whether it participated in SkatteFUNN or not, and different firm characteristics (size, industry, share of high–skilled workers, etc.). Even if R&D *investments* are simultaneously determined with innovation activities, the timing of our R&D variable allows us to consider the R&D capital *stock* as predetermined. Moreover, access to panel data and not just repeated cross-sections gives us an opportunity to estimate models that explicitly take into account the persistence of innovation activities within firms by conditioning on past innovation and patenting activities. To identify causal effects of SkatteFUNN, we model the probability of obtaining SkatteFUNN and the probability of innovations simultaneously, while carefully examining the validity of identifying restrictions.

Our results show that the SkatteFUNN scheme contributes to the development of new production processes and, to some extent, to the development of new products for the firm. Firms that collaborate with other firms in their R&D activities are more likely to innovate. However, the scheme does not appear to contribute to innovations in the form of new products for the market or more patenting.

The structure of our paper is as follows. Section 2 describes our study in the context of the existing literature. Section 3 presents the data and provides descriptive statistics on R&D, innovations and patent applications. Section 4 discusses our identification strategy and describes the model framework, while Section 5 presents the results. Finally, Section 6 concludes and raises policy issues that are relevant for the design of tax credit schemes in general.

2. Our study in the context of the existing literature

There exists a relatively large literature evaluating the effects of public R&D subsidy programs on firms' R&D investment. The early literature is surveyd in Hall and van Reenen (2000) while Mairesse and Mohnen (2010) give an overview of later contributions. Some examples include Wallsten (2000), who examined the U.S. Small Business Innovation Research program, Lach (2000), who studied an Israeli scheme of R&D subsidies for manufacturing firms, and Almus and Czarnitzki (2003), who studied a German R&D subsidy program and Lokshin and Mohnen (2007), who analyzed a tax credit scheme in the Netherlands. Most of these studies focus on possible crowding-out effects; i.e., whether the firms substitute their private R&D investments with public R&D funding. This question has been examined in the context of the Norwegian SkatteFUNN scheme by Hægeland and Møen (2007), and the main findings are summarized in Cappelen et al. (2010). Few researchers have examined the effects on output measures, such as patents and innovations. One exception is the evaluation of the Dutch R&D subsidy program (WBSO); cf. de Jong and Verhoeven (2007), who analyzed how WBSO has influenced the proportion of turnover from sales of new products and services. They reported a significant positive effect of WBSO on their output measure but, like most other studies in this area, they did not pay attention to selection problems. As pointed out by, for instance, Klette et al. (2000) and David et al. (2002), controlling for self-selection is a prerequisite for valid inference.

In general, when evaluating the treatment effects of a public policy program, one must address a counterfactual question: what is the change in the outcome variable given treatment, say $Y(1)$, compared to the potential outcome in the case of non-treatment, $Y(0)$? We cannot answer this question just by regressing the observed outcome, Y (e.g., the number of patent applications), on a dummy for whether the unit is treated ($D = 1$) or non-treated ($D = 0$). The reason is that observed and unobserved variables that affect the dependent variable, Y, may also affect the outcome of D; i.e., we may have a case of self-selection into the program. In the example with SkatteFUNN, firms that already are engaged in R&D activities have a larger probability of applying for R&D subsidies than other firms; i.e., there is "selection based on observables". There may also be "selection based on unobservables". For example, the decision to apply for SkatteFUNN may be based on the (unobserved) probability of success for already ongoing projects. Ignoring selection problems may lead to seriously biased estimates of causal effects. Indeed, our study shows that the way we control for self-selection has a major impact on the results.

One approach to the self-selection problem is propensity score matching, which is based on the assumption that there exists a vector of exogenous covariates, X, such that $Y(0)$ and $Y(1)$ are independent of D given X (conditional independence). According to a result in Rosenbaum and Rubin (1983), a *treated* firm and a *nontreated* firm can be matched if they have identical probability of participating in the program, given X. That is, they can be considered as if they were equal in all other respects (except for an additive error term). The difference in Y can then be calculated for all *matched pairs* and the average value of these differences is a valid estimator of the average treatment effect among the treated. This idea was applied by Almus and Czarnitzki (2003). However, this approach depends on the assumption of an additive error term in the equation for Y (the outcome equation), and therefore is not appropriate when Y is a categorical variable. Even more fundamentally, we will argue that the conditional independence ("selection on observables") assumption is not valid for the Norwegian SkatteFUNN scheme.

To control for "selection on unobservables", Busom (2000) applies a two-stage or control function approach based on Heckman (1976, 1979). In her approach, an equation for whether firms participate in a public funding program is estimated in the first stage, from which suitable variables ("control functions") are computed and included as additional regressors when estimating the effects of the program in the second stage. Selection based on unobservables is also essential in the model of Crepon et al. (1998), often referred to as the CDM model. The CDM model assumes that there exists an unobserved variable for each firm that determines the firm's decision to undertake R&D or not. If the unobserved variable exceeds some threshold, the firm will undertake a positive amount of R&D. Because R&D is an endogenous explanatory variable, it is replaced by its predicted value obtained in the first stage of the CDM model. In the second stage, various innovation variables (patenting, share of innovative sales out of total sales, etc.) are modeled conditional on the R&D activity of the firm. In addition, variables such as capital intensity, firm size, industry dummies, and demand pull and technology push factors are used as covariates. Although Crepon et al. (1998) are not concerned with policy evaluation *per se*, they address selectivity issues that are highly relevant also in an evaluation context. As we discuss later, it is likely that access to various forms of external funding is endogenous to the firm and may impair the estimated effects of policy interventions.

Our approach can be seen as combining the CDM model with the control-function methodology of Busom (2000), although there are some notable differences. First, access to panel data enables us to condition on lagged innovation and patenting variables, both in the equation determining participation in SkatteFUNN and in the outcome equations for innovations and patenting. This enables us to take into account the possibility of persistence in R&D and innovation, cf. Hall et al. (1986) and results surveyed by Mairesse and Mohnen (2010). Second, because we use nonlinear binary-regression models in the second stage, the standard control function approach cannot be applied directly as in Busom (2000). Instead we consider a modification of this approach, building on the likelihood based framework developed in Dagsvik et al. (2011). In fact, our approach applies two of the suggestions made by Mairesse and Mohnen (2010) on how to make progress in innovation analyses; we use panel data and merge innovation surveys with a rich dataset based on administrative registers. Without these two additions our econometric approach would not be feasible.

3. Data and descriptive statistics on R&D, innovations and patents

3.1 Data and variables construction

For the analysis, we use Norwegian micro data on firms included in the 2001 and 2004 innovation surveys, covering the three-year periods 1999–2001 and 2002–2004, respectively. These data are collected by Statistics Norway as a part of the annual R&D survey (we refer to them as *R&D statistics*). They contain detailed information on firms' R&D and innovation activities, including total R&D expenditures (divided into intramural R&D and extramural R&D services), whether the firm has introduced a new product (for the firm or for the market) or a process innovation, and whether it has applied for a patent over the corresponding three year period (henceforth referred to as a *subperiod)*. The sample for the survey is selected using a stratified method for firms with 10–50 employees, whereas all firms with more than 50 employees are included. The strata are based on industry classification (NACE codes) and firm size. Each survey contains about 5000 firms. By supplementing these data with information from the Tax Register, the Register of Employers and Employees (REE) and the

National Education Database (NED), we obtain a panel of firms with observations on innovation outcomes, R&D and other variables over a time span of six years. Table 1 presents an overview of the main variables and the data sources applied in our study.

Table 1.
Overview over variables and data sources.

Variable	Definition	Data source(s)
inpdt	1 if firm has introduced a new product for the firm in the given subperiod, 0 else	R&D statistics
inmar	1 if firm has introduced a new product for the market in the given subperiod, 0 else	R&D statistics
inpcs	1 if firm has introduced a new production process in the given subperiod, 0 else	R&D statistics
d_patent	1 if firm has applied for a patent in the given subperiod, 0 else	R&D statistics
R	R&D investment in the given year	R&D statistics
d_R	1 if $R > 0$ in at least one year in the given subperiod, 0 else	R&D statistics
RK	R&D capital stock at the end of the given year	R&D statistics
rk	R&D capital intensity: *RK*/man-hour	R&D statistics/REE
coopf	1 if firm cooperated with another firm in R&D in the given subperiod, 0 else	R&D statistics
coopu	1 if firm cooperated with a university or research institute in R&D in the given subperiod, 0 else	R&D statistics
d_SFS	1 if the firm obtained SkatteFUNN subsidies in at least one year in the given subperiod, 0 else	Tax register
ac	Share of man-hours worked by employees with academic education (17 or more years of education)	REE/NED

Note: Subperiod includes the last three years.

The first three variables in Table 1, *inpdt, inmar* and *inpcs*, are dummy variables indicating whether the firm has introduced a new product for the firm, a new product for the market or a new production process (during the corresponding subperiod), while the dummy variable *d_patent* indicates whether the firm has applied for a patent during the same period. All these variables are measures of how innovative the firm is and are considered as dependent variables in our analysis.

R&D investment, R, is annual R&D investment as it is reported in the questionnaire, deflated by the R&D deflator used in the national accounts[2]. The dummy variable d_R indicates whether the firm has invested in R&D in at least one year during the last three years.

The volume of R&D capital stock at the end of a given year t, RK_t, is computed by the perpetual inventory method using a constant rate of depreciation ($\delta = 0.15$). That is:

$$RK_t = (1-\delta)RK_{t-1} + R_t,\ t = 1,2,\dots.$$

Following Hall and Mairesse (1995), the benchmark for the R&D capital stock at the beginning of the observation period for a given firm, RK_0, is calculated as if it were the result of an infinite R&D investment series, R^*_{-t}, $t = 0,1,2\dots$, with a fixed presample growth rate $g = 0.05$. That is:

$$RK_0 = R^*_1/(g+\delta) \text{ with } R^*_{-t+1} = (1+g)R^*_{-t},\ t = 0,1,2,\dots,$$

(cf. equation (5) in Hall and Mairesse, 1995). We can interpret R^*_t as the equilibrium growth path for the firm's R&D investments. Hall and Mairesse (1995) use the estimator $R^*_1 = R_1$, which, however, is vulnerable to measurement errors if the actual investment series R_t for $t \le 1$ deviates from the constant growth rate assumption near $t = 1$ (more distant historic values carry less weight due to discounting). We instead apply a more robust estimator by averaging investment observations in the neighborhood of $t = 1$: $R_1^* = 1/T\sum_t R_t$. Here T is the number of observations on the given firm, whereas the summation is over all t for which data are available[3]. This estimator is obviously less influenced by random fluctuations in the observed investment series. R&D capital intensity, rk, is calculated as the R&D capital stock per man-hour, where the latter is the sum of all man-hours in the firm.

The two variables, *coopf* and *coopu*, indicate if the firm cooperated with another firm (*coopf*) or with a university or research institute (*coopu*) when carrying out R&D during the last three

[2] More than 60 percent of total R&D expenditures are made up of labor costs.

[3] We use all available data on the firm's R&D investments from the biannual 1993–1999 R&D surveys, and the annual 2001–2005 R&D surveys. Although the sample we analyze in this paper is restricted to 1999–2004, we also utilize the out-of-sample data to estimate the initial R&D capital stock, if available for the given firm.

years. The dummy variable *d_SFS* indicates whether the firm obtained a SkatteFUNN subsidy in at least one year during the corresponding subperiod. It may have a positive value only for the second subperiod, 2002–2004, because SkatteFUNN was introduced in 2002.

For each firm, we distinguish between two types of employees: those with academic education (corresponding to a Master or PhD level of education) and those without. We assume that the former group is particularly relevant to the R&D activity in the firm. The variable *ac* is defined as the number of man-hours worked by employees with academic education divided by the total number of man-hours in the firm.

The 2001 and 2004 innovation surveys include 3896 and 4655 firms, respectively. After merging them and excluding firms with incomplete information on the variables of interest, we obtain a reduced unbalanced panel of 2476 firms. Keeping only firms included in both surveys gives us a balanced panel of 1689 firms.

Table 2.
Mean values of key variables.

	1999–2001	2002–2004
Number of firms	1689	1689
rk	0.13 (0.35)	0.13 (0.38)
ac	5.9%	6.4%
Share of firms with *d_R = 1*	39.8%	53.6%
rk | *d_R = 1*	0.27 (0.51)	0.24 (0.50)
ac| *d_R = 1*	9.3%	9.0%
Share of firms with *all SFS* > 0	–	10.5%
Share of firms with *d_SFS = 1*	–	36.2%

Note: Standard deviation in parenthesis; all means are over three year periods; *all SFS > 0* indicates that the firm obtained SkatteFUNN subsidy each year during the corresponding three-year period.

3.2 Descriptive statistics on R&D, innovations and patents

For our sample of firms, Table 2 reports the mean (and standard deviation) for R&D capital intensity, *rk*, and the share of man-hours worked by employees with academic education, *ac*, in the two subperiods. We compute these measures both for the total sample of firms and for the subsample of firms that had some R&D activity during the relevant subperiod (*d_R = 1*).

Not surprisingly, the share of man-hours worked by employees with academic education, *ac*, is higher for the firms that invested in R&D. Moreover, more firms invested in R&D (*d_R =*

1) in the second subperiod; i.e., about 40 percent in 1999–2001 versus about 54 percent in 2002–2004. Turning to the SkatteFUNN variables, we see that 10.5 percent of the firms in the sample received a tax subsidy each year in 2002–2004 (*all SFS* > 0). The share of firms who received a subsidy at least once during 2002–2004 (*d_SFS = 1*) is much higher; i.e., 36.2 percent.

Table 3.
The share of innovative firms by the type of innovation and patenting, percent.

	1999–2001	2002–2004
New product for the firm (*inpdt*)	41.2	36.7
New product for market (*inmar*)	18.6	19.5
New production process (*inpcs*)	33.3	22.4
Applied for a patent (*d_patent*)	13.9	13.9
inpdt | *d_R* = 1	81.0	61.6
inmar | *d_R* = 1	39.4	33.9
inpcs | *d_R* = 1	62.2	38.4
patent | *d_R* = 1	28.7	23.3
inpdt | *d_SFS* = 1	–	65.8
inmar | *d_SFS* = 1	–	38.6
inpces | *d_SF*S = 1	–	39.9
d_patent | *d_SFS* = 1	–	24.6

Note: *n* = 1689 firms in each subperiod.

Table 3 gives information about the firms' innovation and patenting activities in 1999–2001 and 2002–2004. The first group of rows reports, separately for each subperiod, the share of innovative firms by different types of innovations and patent application. The next two groups of rows report the share of innovative firms in the two subsamples – the firms carrying out some R&D activity (*d_R = 1*) and the SkatteFUNN firms (*d_SFS = 1*), respectively. One can see that the most frequent type of innovation for Norwegian firms is a new product for the firm, followed by a new production process. Less than 20 percent of these firms innovate in the form of a new product for the market and only about 14 percent apply for a patent. Firms with positive R&D investments and SkatteFUNN firms are more innovative than other firms in both subperiods. There is less innovation during 2002-2004 than during 1999-2001. This could be the result of the economic downturn that Norway experienced in the aftermath of the ICT-bubble in 2001, leading to unfavorable demand conditions. It is consistent with the original CDM-model of Crepon et al. (1998) that demand conditions affect innovation.

4. Econometric approach

The main purpose of our analysis is to estimate the causal effect of SkatteFUNN on the probability of the various types of innovations. In many empirical analyses, e.g., using the CDM model, innovations are considered as a type of output of a production function, with R&D capital as one of the input factors. The impact of a tax credit on the innovation output of the firms is then assumed to depend on two things: (i) how much the tax credit increases R&D investment (or R&D intensity) compared to a hypothetical situation without any tax credit, and (ii) whether this (marginal) investment leads to an innovation or not. Ideally, we should be able to observe for each firm both (i) and (ii). In practice, we only observe whether a firm participates in SkatteFUNN or not, and whether it innovates or not.

In principle, we could model the joint relation between R&D subsidies, R&D investment and innovation output. This is done in, for instance, Czarnitzki and Licht (2006) and Hussinger (2008), who examine the impact of subsidies on R&D expenditures in a first stage of a two-stage procedure, and then the impact of the estimated incremental R&D on innovative output in the second stage. Our approach, however, is more modest aiming only to identify the *marginal* effect of the tax credit scheme on the innovation output. Nevertheless, we do take R&D intensity into account, since we condition on the firm's R&D capital stock at the *beginning* of the subperiod (which is determined before the tax credit program starts). We believe that a marginal model is adequate in our situation for several reasons. First, our data clearly reject the assumption of a recursive structure underlying a standard two-stage-procedure: many firms report innovative output without carrying out formal R&D. [4] Thus positive R&D investment (as measured in the data) is not a necessary condition for innovations. Secondly, the relation between R&D subsidies and (incremental) R&D is notoriously hard to measure and it is far from obvious that a joint model (with two outcome equations instead of one) would give more accurate answers to our questions. Thirdly, an increase in R&D expenditures as a result of subsidies may not fully reflect a quantity increase. For example, it has been documented in Norway that subsidies led to increasing wages of R&D workers in the industries mostly affected by the SkatteFUNN program and also to re-

[4] The Norwegian business sector as a whole reported innovation costs to be 22 bill. NOK in 2004, whereas total R&D investments were 13 bill. NOK. Clearly, firms regard innovation activities and expenditures as comprising a wider set of activities than R&D.

labeling of ordinary operating costs as R&D expenditures.[5] Finally, it is an interesting policy question *per se* whether innovative output increases as a consequence of the R&D subsidy program: If a positive effect on innovations can be identified, reported increases in reported R&D expenditures cannot be driven solely by re-labeling and price effects.

We will now turn to our formal models. The dependent variable, Y_t, is binary and takes the value 1 ("success") or 0 ("failure"). With reference to the variable definitions in Table 1, we will study four types of innovations: (i) $Y_t =$ "*d_patent*", (ii) $Y_t =$ "*inpdt*", (iii) $Y_t =$ "*inmar*", and (iv) $Y_t =$ "*inpcs*". In the case (i), $Y_t = 1$ means that the firm applied for at least one patent in subperiod *t*. In the cases (ii)–(iv), $Y_t = 1$ means that the corresponding type of innovation occurred in subperiod *t*. The time index *t* takes two values: $t = 1$ refers to the subperiod 1999–2001 (the three years before the introduction of SkatteFUNN), while $t = 2$ refers to the subperiod 2002–2004 (the three years following the introduction of SkatteFUNN in 2002).

Our empirical analyses are based on three different versions of binary regression. The first version should be considered as merely providing descriptive statistics. Here Y_t is regressed on a vector of explanatory variables, X_t:

(1) $$\Pr(Y_t = 1 \mid X_t) = \frac{1}{1 + \exp(-X_t\beta)}.$$

That is, (1) is a logit model, where $\Pr(Y_t = 1 \mid X_t)$ denotes the probability of "success" for a given firm in period *t* given X_t, where X_t is a (row) vector of independent variables (covariates) and β is the corresponding (column) vector of regression coefficients. The vector X_t contains dummies for subperiod 1 or 2; intervals for the number of employees: [10,20), [20,50), [50,100) and ≥ 100; industry (construction, retail trade, transportation services, other services and different manufacturing groups)[6]; whether the firm obtained

---

[5] See http://www.ssb.no/skattefunn/rapp_200802_en.pdf.

[6] All industry definitions and the definition of 11 manufacturing groups are based on NACE-codes, SN2002.

SkatteFUNN subsidies (*d_SFS*); and whether the firm cooperates with other firms (*coopf)* and/or a university or research institute (*coopu*) when carrying out R&D. Through the interaction terms $d\_SFS \times coopf$ and $d\_SFS \times coopu$, we also allow the effect of SkatteFUNN to depend on whether the firm has such cooperation or not. Finally, X_t contains two continuous regressors: the share of total man-hours by employees with academic education (*ac*) and R&D capital per man-hour (*rk*).

The date of the variables in X_t refers, in most cases, to the beginning of subperiod *t*. For example, *rk* in subperiod $t = 1$ refers to the R&D capital per man-hour in 1999. We choose this dating to reduce the potential endogeneity problem that occurs if the right-hand side variables can be adjusted as a consequence of changes in the dependent variable, Y_t. For example, if the firm increases its R&D investment or initiate a joint research project with another firm as a consequence of an innovation. The one exception to this dating convention regards the dummy variable *d_SFS*, which is one if the firm received SkatteFUNN subsidies *during* the three year period *t* and zero otherwise.

Valid statistical inference in the framework of equation (1) requires that *d_SFS* is independent of Y_t given the covariates, X_t (i.e., conditional independence). This assumption will not be fulfilled if the firms that *independently of* SkatteFUNN would have had the highest number of patents or innovations also have the highest probability of participating in SkatteFUNN (given the observable exogenous variables). There are certainly reasons to believe that this is the case. Another shortcoming of specification (1) is that it does not take into account the persistence of innovation activities at the firm level; that the dependent variable, Y_t, may depend on Y_{t-1} (given the explanatory variables X_t).[7] For example, we expect, *ceteris paribus*, that the probability of innovating in the second subperiod ($Y_2 = 1$) is larger for firms that innovated in the first subperiod ($Y_1 = 1$) than for those that did not ($Y_1 = 0$).

In our second econometric specification, we attempt to address these concerns by assuming that

[7] Earlier studies have shown that innovation is a quite persistent feature of firms cf. Peters (2007).

(2) $$\Pr(Y_2 = 1 \mid X_2, Y_1) = \frac{1}{1 + \exp\{-(X_2\beta + Y_1\alpha_1 + (d\_SFS \times Y_1)\alpha_2)\}}.$$

Thus, we model the probability of success in subperiod 2 (the probability that $Y_2 = 1$) conditional on the explanatory variables in subperiod 2, X_2, *and* the dependent variable in subperiod 1, Y_1. In addition, we include an interaction effect between *d_SFS* and Y_1, $d\_SFS \times Y_1$. This interaction term allows the effect of SkatteFUNN to depend on whether the firm had patents/innovations in the previous subperiod (i.e., it depends on Y_1). The validity of specification (2) rests on the assumption that the firms that obtained SkatteFUNN in period 2 constitute a randomized sample from the population of all firms given the value of the independent variables X_2 *and* the lagged value of the dependent variable, Y_1. In other words, the SkatteFUNN firms are allowed to be a self-selected subsample based on the previous outcome of the dependent variable.

Even if the conditional independence assumption behind (2) is weaker than for (1), it may not be valid. For example, it is invalidated if the decision to apply for SkatteFUNN (or to accept the application) is based on the (unobserved) probability of success of already ongoing projects. Thus, there may be reverse causality that cannot be accounted for simply by conditioning on the lagged dependent variable or other observed variables.

To account for reverse causality, we propose a third, more elaborate, econometric specification, that generalizes (2) in two substantial ways. First, *d_SFS* is allowed to be endogenous in the sense that an unobserved variable that influences Y_2 may also influence the outcome of *d_SFS*. Second, we incorporate unobserved heterogeneity in the effects of SkatteFUNN by letting the coefficients of *d_SFS* — inclusive of the interaction terms — to be random coefficients (and hence vary across firms). This second feature of the model is motivated by a consideration of how SkatteFUNN may change a firm's incentives for carrying out R&D activities: One way is, obviously, by reducing the marginal costs of R&D. Another way is by improving the liquidity of the firm. In the latter case, the tax credit may finance R&D investments that would have been profitable also in the absence of the scheme (see Hall (2002) for a discussion of the importance of financing constraints for R&D investments). The fact that most government revenue loss due to SkatteFUNN is in the form of payment of grants to firms that are not in a tax position, suggests that a firm's liquidity situation is important for whether it chooses to participate in the program or not[8]. In fact, firms that are not in a tax position are significantly over-represented among the SkatteFUNN firms, even when we control for a broad set of covariates, including industry dummies, firm age, operating margins and value added per man-hour. Since a firm's tax position is likely to be positively correlated with its access to funds, firms with poor liquidity seem to be over-represented among those who participate in SkatteFUNN.

The above discussion motivates our choice to incorporate heterogeneous treatment effects in the third version of our econometric model. Then the average treatment effect on the treated — the average effect of SkatteFUNN on the firms that participate in the scheme — can be considered as a weighted average of 0 (the contribution from participants whose behavior is not affected) and the average effect among participants who increase their R&D as a result of the tax credit. On the other hand, firms that do not participate ("the control group") do so because the scheme does not give them sufficient incentives to change their behavior. The latter group is clearly also heterogeneous, comprising e.g. firms whose demand for R&D is locally inelastic (possibly due to fixed costs or indivisibilities in R&D investments), and firms that do not bother to participate in SkatteFUNN even if they may be eligible for tax credits.[9]

To present the third version of our model, let X^* be a continuous latent index, such that $d\_SFS = 1$ if the value of X^* is larger than 0:

(3)
$$d\_SFS = \begin{cases} 1 & \text{if } X^*>0 \\ 0 & \text{else} \end{cases} .$$

Furthermore, assume that

(4)
$$X^* = Z^{(1)}\gamma_1 + \varepsilon^{(1)},$$

where $Z^{(1)}$ is a row-vector of exogenous or predetermined variables that determines the probability that $d\_SFS = 1$, γ_1 is a fixed vector of (unknown) parameters, and $\varepsilon^{(1)}$ is normally

[8] Since 2002, roughly 75 percent of the subsidies have been given as grants and this share has been a very stable feature of the scheme irrespective of business cycle fluctuations.

[9] In fact, about 41 percent of firms that reported positive R&D expenditure in 2004 did not apply for SkatteFUNN tax credits. This suggests that many firms do not bother to apply, even if they are eligible for tax credits.

distributed with expectation 0 and variance 1. The equations (3) and (4) thus represent a standard probit model for the binary dependent variable *d_SFS*. In our specification of $Z^{(1)}$ we include all variables in X_t except the (endogenous) variables that involve *d_SFS*. In addition, we include regional dummies, the lagged innovation variable, Y_1, and a binary variable *taxposition*, which is 1 if the firm is in a tax position (i.e., pays taxes) in the beginning of subperiod 2 (i.e., in 2002). The variables in $Z^{(1)}$ are thus as follows:

$Z^{(1)}$ = *(1, dummies for employment intervals, industries and regions, ac, rk, coopf, coopu, taxposition,* Y_1*).*

The outcome of the innovation variable Y_2, is also assumed to be determined by a probit model:

(5) $$Y_2 = \begin{cases} 1 & \text{if } Y^* > 0 \\ 0 & \text{else,} \end{cases}$$

where

(6) $$Y^* = Z^{(2)}\gamma_2 + S(\beta + \eta) + \varepsilon^{(2)}.$$

In equation (6), $Z^{(2)}$ includes the same variables as $Z^{(1)}$, except *taxposition* and the regional dummies. That is

$Z^{(2)}$ = *(1, dummies for employment intervals and industries, ac, rk, coopf, coopu,* Y_1*).*

Furthermore, the vector *S* contains the endogenous variables that involve *d_SFS*:

$S = (d\_SFS,\ d\_SFS \times coopf,\ d\_SFS \times coopu,\ d\_SFS \times Y_1)$.

Moreover, γ_2 and β are fixed regression coefficients, η is a random coefficient vector with expected value 0, $\varepsilon^{(2)}$ is an additive error term with expectation 0 and variance 1. The vector $(\varepsilon^{(1)}, \varepsilon^{(2)}, \eta')'$ is assumed to have a multinormal distribution with expectation 0 and where η is independent of $(\varepsilon^{(1)}, \varepsilon^{(2)})$ and $\varepsilon^{(1)}$ and $\varepsilon^{(2)}$ both have variance 1 (to obtain identification).

Model (3)–(6) generalizes (2) in two important ways. First, the additive error term, $\varepsilon^{(1)}$, in the equation that determines participation in SkatteFUNN, is allowed to be correlated with the additive error term, $\varepsilon^{(2)}$, in the equation that determines the latent innovation variable Y^*. Secondly, we allow for heterogeneity in the effects of the endogenous explanatory variables through the vector of random slope coefficients, η.

Our *a priori* hypothesis (which will later be tested) is that the variable *taxposition* and the regional dummies do affect the participation in SkatteFUNN, but not the probability of innovations *per se*. Thus they can be excluded from $Z^{(2)}$. The main rationale for the exclusion restriction with regard to *taxposition* is that, while financing constraints may restrict a firm's innovation activities and its R&D investment in particular, it is not a production factor *per se*. That is, given the vector of productive inputs at the *beginning* of the three-year period (including R&D capital stock), information about the firm's tax position should not help us to predict its innovation probability during the corresponding subperiod. With regard to the regional dummies, there is little reason a priori to believe that the regional heterogeneity has a direct effect on the probability of success of the research projects. The validity of our proposed exclusion restrictions is examined empirically in Section 5, where we conduct a "placebo-experiment", using data for two 3-year subperiods prior to SkatteFUNN.

If Y^* was a directly observable variable, Models (3)–(6) would be a special case of the general ordered probit model analyzed by Dagsvik et al. (2011), and estimation could be carried out by the maximum likelihood methods derived there. However, because Y^* is not observable — we only observe the binary variable Y_2 — the methods in Dagsvik et al. (2011) must be modified. Our approach consists of two steps. In step one we estimate the probit-model (3)–(4) in the traditional way. In step two, given the estimates of γ_1 from the first step, we reformulate (6) in a way similar to that in Dagsvik et al. (2011)[10]:

(7) $$Y^* = S\beta + Z^{(2)}\gamma_2 - \lambda(d\_SFS)\theta + \varepsilon^*,$$

[10] Equation (7) corresponds to equation (20) in Dagsvik et al. in the special case with $\rho = 0$: That is, no correlation between the random slope coefficients η and $\varepsilon^{(1)}$.

where the error term ε^* has conditional expectation equal to 0 given $Z^{(2)}$ and *S*. In (7), ρ is an unknown parameter vector, θ is an unknown parameter, and $\lambda(d\_SFS)$ is an inverse Mills ratio (see, e.g., Heckman, 1976):

$$(8) \qquad \lambda(d\_SFS) = \begin{cases} \dfrac{\phi(-Z_t^{(1)}\gamma_1)}{\Phi(-Z_t^{(1)}\gamma_1)} & \text{if } d\_SFS = 0 \\ \dfrac{-\phi(-Z_t^{(1)}\gamma_1)}{1-\Phi(-Z_t^{(1)}\gamma_1)} & \text{if } d\_SFS = 1. \end{cases}$$

In (8), $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal probability density and cumulative distribution functions, respectively. The parameter vector γ_1 is replaced by the corresponding estimate obtained in the first step when applying (8). By postulating a logistic distribution for ε^*, we obtain a logit model for Y_2. The corresponding logit estimators of γ_2 and β can be interpreted as pseudo maximum likelihood estimators, cf. Gourieroux et al. (1984).

5. Estimation results

5.1 Which firms participate in the SkatteFUNN scheme?

We consider first the estimation of the SkatteFUNN-participation equation (4). We next use the results of this estimation to calculate the function $\lambda(\cdot)$ defined in (8), to estimate the innovation equation (7). The problems we address in this subsection have many similarities with Blanes and Busom (2004), who study participation in R&D subsidy programs for Spanish manufacturing firms. Like them, we include measures of the skill of the employees, lagged R&D activity, firm size and financing constraints as explanatory variables in the participation equation.

The results in Table 4 show that the regional dummies *South*, *West*, and *Mid-Norway* are significant and have a positive sign (the coefficients of the capital region Oslo and the category of less than 10 employees[11] are 0 *a priori*). The results are in line with a long tradition in Norway of taking regional considerations into account when deciding on policies, with a bias in favor of remote and sparsely populated regions. Although SkatteFUNN is open to all firms, firms that apply for SkatteFUNN funding do so through the regional offices of Innovation Norway, which may differ in their ability to support firms in their application writing. Note, however, that Innovation Norway is only helping firms through the application process, and has no authority to approve the project.[12]

From Table 4 we also observe a significant *negative* relation between SkatteFUNN and the variable *taxposition*. A plausible explanation, as suggested in our discussion in Section 4, is that participation in SkatteFUNN is motivated by the liquidity situation of the firm: If the firm is not in a tax position, the tax credit will be given as a grant and thus increases the firm's cash holdings. SkatteFUNN is a more easily accessible source of cash than ordinary research grants, which was also the intension of the scheme from the start. Another explanation could be that *taxposition* is negatively correlated with the firm's tax planning abilities, including knowledge about and active use of tax reducing schemes such as SkatteFUNN.[13]

Finally, the results in Table 4 show that the dummy for lagged innovation activities, Y_I, has a huge impact on participation in SkatteFUNN. Y_I is clearly the most significant variable in the table, so there is a strong persistence in the firms' innovation activities. The share of man-hours worked by persons with academic education, *ac,,* and both types of cooperation, *coopf* and *coopu*, have both large impacts on the participation probability. There is considerable heterogeneity across different industries, with firms from some of the manufacturing industries having, *ceteris paribus*, the highest probability of participating in SkatteFUNN.

[11] Firms with less than 10 employees may be included in the survey if they had at least 10 employees at the time of the sample selection, but they make up less than two percent of the observations.

[12] The project plan must be approved by the Research Council of Norway, while the final R&D expenditures have to be approved by the tax authorities, based on their own judgments and a statement from the applicant's auditor.

[13] The latter interpretation was suggested to us by an accountant who had experience in auditing firms subsidized by the SkatteFUNN scheme. His experience with these firms was that they had somehow specialized in applying for government funds and that many of them had acquired skills in managing the bureaucracy involved.

Table 4.
Probit estimates for the probability that *d_SFS* = 1.

Explanatory variables ($Z^{(1)}$)	Estimate	Std. Dev.
10–19 employees (d)	–0.36	0.23
20–49 employees (d)	–0.08	0.22
50–99 employees (d)	–0.10	0.22
≥ 100 employees (d)	–0.27	0.21
Construction (d)	0.38	0.26
Retail trade (d)	0.07	0.24
Transport (d)	–0.30	0.19
Services (d)	0.55	0.18
Industry A (Nace 15–16) (d)	0.65	0.18
Industry BC (Nace 17–19) (d)	0.64	0.23
Industry D (Nace 20) (d)	0.62	0.23
Industry E (Nace 21–22) (d)	0.16	0.19
Industry FGH (Nace 23–25) (d)	0.91	0.21
Industry I (Nace 26) (d)	0.48	0.24
Industry J (Nace 27–28) (d)	0.69	0.19
Industry K (Nace 29) (d)	1.09	0.21
Industry L (Nace 30–33) (d)	0.80	0.20
Industry M (Nace 34–35) (d)	0.73	0.19
Industry N (Nace 36–37) (d)	1.20	0.23
East coast (d)	0.10	0.12
East inland (d)	0.11	0.17
South (d)	0.64	0.12
West (d)	0.41	0.12
Mid-Norway (d)	0.39	0.15
North (d)	0.29	0.18
ac	1.18	0.35
rk	0.12	0.07
coopf (d)	0.30	0.10
coopu (d)	0.39	0.12
taxposition (d)	–0.16	0.07
Y_1 (lagged innovation) (d)	0.44	0.08
constant	1.19	0.26

Note: *n* = 1689; (d) indicates binary variable. There are seven main regions in Norway, where the capital region Oslo is made the reference; less than 10 employees is the reference category for the employment variable.

5.2 Testing the validity of exclusion restrictions

A key role in our strategy for identifying causal effects is assigned to variables that are contained in $Z^{(1)}$, but not in $Z^{(2)}$. These exclusion restrictions contribute to identifying the effects of participating in SkatteFUNN: For a given value of $Z^{(2)}$, variation in the excluded variables contributes to *exogenous* variation in *d_SFS*. Our hypothesis is that *taxposition* and the regional dummies satisfy these criteria. This identification requires, of course, that regional location and whether a firm is in a tax position or not, has no direct effect on the innovation probability, but do affect the SkatteFUNN participation probability.

In view of the great importance attributed to the exclusion restrictions for identifying causal effects, it is a main feature of our data that we can actually test, at least partially, the validity of these restrictions. To do so, we utilize the innovation survey of 1999, which contains information about the outcome variables *inpdt* and *inpcs* (but not *d_patent* and *inmar*) for the three years *up to* 1999 — henceforth denoted subperiod *0*. Thus we are able to estimate $\Pr(Y_1 = 1 \mid X_1, Y_0)$. That is, the conditional logit model defined in (2), with Y_1 as the outcome variable conditional on (X_1, Y_0). Since there was no SkatteFUNN program before period *2*, this estimation is equivalent to estimating the $\Pr(Y^* > 0 \mid Z^{(2)})$ defined in (7), with $\beta = \gamma = \xi = 0$. Moreover, when considering Y_1 as the outcome variable, we can include *taxposition* and all the regional dummies in $Z^{(2)}$ and test whether the corresponding parameters are zero. This can be considered as a "placebo experiment". The results are shown in Table 5.

Table 5.
Testing exclusion restrictions regarding *taxposition* and the regional dummies. Dependent variable: Y_1. Subperiod *t*=0 corresponds to three year period up to 1998.

Explanatory variables ($Z^{(1)}$)	New product for the firm (d)		New production process (d)	
	Est.	Std. Dev.	Est.	Std. Dev.
coopf (d)	0.68	0.29	0.86	0.32
coopu (d)	0.54	1.14	0.64	0.36
East coast (d)	0.36	0.28	–0.26	0.29
East inland (d)	–0.41	0.42	–0.47	0.53
South (d)	–0.46	0.29	–0.17	0.30
West (d)	–0.19	0.27	–0.65	0.41
Mid–Norway (d)	–0.24	0.38	–0.64	0.95
North (d)	–1.00	0.40	–1.46	0.68
ac	0.30	1.05	1.59	1.09
rk	2.00	0.50	0.40	0.27
Y_0 (d)	1.97	0.20	1.45	0.25
taxposition (d)	–0.08	0.17	–0.09	0.21

Note: *n* = 1770; (d) indicates binary variable. Size and industry dummies are included in the analyses, but the corresponding parameter estimates are not reported here.

We see from Table 5 that all the regional dummies, *except North*, are clearly insignificant in both innovation equations and can be excluded. This result is also confirmed when carrying out a joint likelihood ratio test. However, *North* cannot be excluded according to these results, and we decided to include *North* as an explanatory variable in our subsequent analyses, thus extending the variable list in $Z^{(2)}$. Moreover, Table 4 shows that *South, West and Mid-Norway* are relevant, and therefore contribute to exogenous variation in *d_SFS*.

It is also confirmed in Table 5 that *taxposition* has no significant effect on the probability of innovating, with an estimate close to zero in both equations. Thus, whereas *taxposition* can be excluded from the innovation equation in view of the results in Table 5, it is clearly a relevant variable in view of the results in Table 4.

5.3 Estimation results for innovations and patents

Tables 6–9 present the results for the four different dependent variables and three different versions of our econometric model. The tables focus on the most important explanatory variables. We do not report results for control variables such as firm size and industry. Our estimates regarding these variables are in line with results well established elsewhere in the empirical literature: large firms (with many employees) have a higher probability of patenting or innovating than other firms, cf. Griffith et al. (2006), while manufacturing industries are those with the highest probability of having patent applications and innovations. Based on the results in Table 5, where the exclusion restriction regarding the dummy *North* was firmly rejected, this variable is included among the explanatory variables and reported in Tables 6-9. The result regarding *North* is in line with the findings reported in Table 5: the propensity to innovate is generally lower in the most northern region.

Table 6 shows the results for patents. We first note that the estimated coefficient of the time dummy of the second period is lower than for the first one in the (standard) logit model. This corresponds to an approximately 10 percent drop in total patent applications in Norway between these two subperiods. In the standard logit model, the share of employees with academic education, *ac*, and R&D capital intensity, *rk*, are also highly significant explanatory variables. This result is typical in the literature; cf. Crepon et al. (1998), Parisi et al. (2006), and Griffith et al. (2006).

Table 6.
Binary regression. Dependent variable: Y_t = "d_patent" (patents).

Explanatory variables	logit		conditional logit		conditional logit with selection	
	Est.	Std. Dev.	Est.	Std. Dev.	Est.	Std. Dev. [a]
1999–2001 (d)	–3.45	0.40	–	–	–	–
2002–2004 (d)	–3.55	0.40	–3.55	0.51	–4.12	0.77
North (d)	–1.55	0.67	–1.05	0.57	–0.84	0.86
ac	2.39	0.47	0.75	0.88	0.51	0.94
rk	0.80	0.11	0.37	0.16	0.27	0.22
coopf (d)	0.62	0.13	0.43	0.22	0.49	0.24
coopu (d)	0.80	0.14	0.61	0.26	0.80	0.27
d_SFS (d)	0.24	0.23	0.44	0.30	1.10	1.18
d_SFS × *coopf* (d)	0.31	0.24	0.24	0.30	0.23	0.31
d_SFS × *coopu* (d)	0.15	0.22	0.10	0.29	0.05	0.29
d_SFS × Y_{t-1} (d)	–	–	0.07	0.35	0.16	0.40
Y_{t-1} (d)	–	–	2.54	0.26	2.54	0.31
$Corr(\varepsilon^{(1)}, \varepsilon^{(2)})$	–	–	–	–	0.12	(0.05)

Note: Logit n = 2467; conditional logit n = 1527; (d) indicates binary variable; dummies for size and industry are included as regressors, but the corresponding parameter estimates are not reported here.
[a] Std. Dev. calculated conditional on the step-1 estimates.

The coefficient of *d_SFS* is not significant in any of the three model specifications reported in Table 6. However, a joint hypothesis of whether *all* the parameters involving SkatteFUNN-variables are zero is rejected in the standard logit model, with a p-value of 0.01. The conditional logit model that allows *d_SFS* to be endogenous (cf. the column named "conditional logit with selection") gives estimates for the effect of SkatteFUNN that are less significant than in the simple conditional logit model. For example, a joint test that *d_SFS* has zero effect on patenting yields a p-value of 0.63. Thus there is no evidence that the SkatteFUNN scheme affects the probability of patent applications. We also see that by conditioning on lagged innovations, Y_{t-1}, much of the explanatory power of *rk* and *ac* disappears compared to the standard logit model. Our result on the significance of persistence in patenting contradicts results in Geroski et al (1997) and Malerba and Orsenigo (1999) who find no persistence in patenting. On the other hand, the dummies for cooperation *coopf* and *coopu* are highly significant in all the three model variants, especially cooperation with a research institute (*coopu*).

We see that the estimated correlation between $\varepsilon^{(1)}$ and $\varepsilon^{(2)}$, i.e., the residual terms of the SkatteFUNN participation equation and the patent application equation, is positive (0.12) and significant at the 5 percent level. Thus there is clear evidence of self-selection and endogeneity. The corresponding (positive) correlation is even stronger in the case of the innovations reported in Tables 7–9, ranging from 0.22 to 0.34.

Table 7.
Binary regression. Dependent variable: Y_t = "inpdt" (new product for the firm).

Explanatory variables	logit		conditional logit		conditional logit with selection	
	Est.	Std. Dev.	Est.	Std. Dev.	Est.	Std. Dev.[a]
1999–2001 (d)	–1.44	0.28	–	–	–	–
2002–2004 (d)	–2.07	0.28	–2.93	0.47	–2.72	0.53
North (d)	–1.61	0.52	–1.21	0.47	–1.12	0.52
ac	1.12	0.45	1.60	0.71	2.30	0.82
rk	0.70	0.17	–0.01	0.14	–0.02	0.17
coopf (d)	1.19	0.12	0.47	0.21	0.53	0.22
coopu (d)	0.67	0.14	0.23	0.26	0.27	0.28
d_SFS (d)	1.52	0.17	1.68	0.22	1.48	0.86
$d\_SFS \times coopf$ (d)	0.19	0.21	0.48	0.25	0.44	0.27
$d\_SFS \times coopu$ (d)	–0.03	0.23	0.17	0.27	0.07	0.29
$d\_SFS \times Y_{t-1}$ (d)	–	–	–0.92	0.28	–1.20	0.31
Y_{t-1} (d)	–	–	1.91	0.17	2.32	0.25
$Corr(\varepsilon^{(1)}, \varepsilon^{(2)})$	–	–	–	–	0.34	(0.06)

Note: Logit n = 2467; conditional logit n = 1484; (d) indicates binary variable; dummies for size and industry are included as regressors, but the corresponding parameter estimates are not reported here.

[a] Std. Dev. calculated conditional on the step-1 estimates.

Examining the results for innovations reported in Tables 7–9, one can observe a general pattern for all three types of innovations. From the standard logit model we find that the share of academics (*ac*) and R&D capital per man-hour (*rk*) are highly significant variables. In the two conditional logit models, however, we find no significant impact of *rk* on the probability of innovating. The explanation may be that the value of the lagged dependent variable Y_{t-1} also incorporates the effect of *rk,* because these variables are highly correlated, as is evident from the standard logit model. Again, we find that the coefficient of Y_{t-1} is highly significant and positive and of a similar magnitude in the two conditional models. For all innovation types, we find that cooperation with another firm (*coopf*) is a significant explanatory variable in contrast to the case of patenting, where *coopu* (cooperation with a research institute) is more influential. This difference may reflect that innovations in general are "closer to the market" than patenting, where academically oriented collaboration is more important, cf. Mairesse and Mohnen (2010) who report similar findings in their survey.

Table 8.
Binary regression. Dependent variable: Y_t = "inmar" (new product for the market).

Explanatory variables	logit		conditional logit		conditional logit with selection	
	Est.	Std. Dev.	Est.	Std. Dev.	Est.	Std. Dev.[a]
1999–2001 (d)	–3.21	0.31	–	–	–	–
2002–2004 (d)	–3.05	0.32	–3.67	0.47	–3.52	0.55
North (d)	–1.96	0.46	–1.36	0.58	–1.46	0.72
ac	1.49	0.38	1.34	0.62	1.73	0.73
rk	0.01	0.03	0.15	0.14	0.18	0.15
coopf (d)	0.87	0.11	0.57	0.19	0.61	0.50
coopu (d)	0.60	0.13	0.43	0.23	0.72	0.46
d_SFS (d)	1.25	0.16	1.51	0.20	1.10	0.84
$d\_SFS \times coopf$ (d)	0.10	0.18	0.13	0.21	0.11	0.24
$d\_SFS \times coopu$ (d)	0.10	0.19	0.15	0.22	0.16	0.25
$d\_SFS \times Y_{t-1}$ (d)	–	–	–1.16	0.29	–1.30	0.34
Y_{t-1} (d)	–	–	2.09	0.21	2.39	0.27
$Corr(\varepsilon^{(1)}, \varepsilon^{(2)})$	–	–	–	–	0.25	(0.05)

Note: Logit n = 2467; conditional logit n = 1484; (d) indicates binary variable; dummies for size and industry are included as regressors, but the corresponding parameter estimates are not reported here.

[a] Std. Dev. calculated conditional on the step-1 estimates.

The SkatteFUNN dummy is clearly significant in the standard logit and conditional logit models, but the results with regard to *d_SFS* are ambiguous in the model where it is allowed to be endogenous. Then the coefficient of *d_SFS* is significant at the 5 percent level in the case of a new production process and 10 percent level in the case of a new product for the firm. The interaction effect between *d_SFS* and cooperation with another firm ($d\_SFS \times coopf$) is significant at the 10 percent level both in the conditional logit model and the conditional model with selection in the case of a new product for the firm, but is far from being significant with regard to the other types of innovations. The estimated interaction effect $d\_SFS \times Y_{t-1}$ is negative and significant in both the conditional models: if the firm was innovating *before* getting a SkatteFUNN subsidy, the effect of SkatteFUNN is weaker. Likelihood ratio tests of the joint hypothesis that all the variables involving *d_SFS* have corresponding coefficients equal to zero (4 degrees of freedom in the standard logit model, and 5 degrees of freedom in the conditional logit models) were clearly rejected in all the model variants, with p-values close to zero.

Table 9.
Binary regression. Dependent variable: Y_t = "inpcs" (new production process).

Explanatory variables	logit		conditional logit		conditional logit with selection	
	Est.	Std. Dev.	Est.	Std. Dev.	Est.	Std. Dev.[a]
1999–2001 (d)	–1.71	0.27	–	–	–	–
2002–2004 (d)	–2.44	0.28	–2.57	0.45	–3.71	0.54
North (d)	–1.62	0.40	–1.06	0.52	–1.60	0.72
ac	0.60	0.39	0.65	0.58	1.42	1.13
rk	0.12	0.09	–0.09	0.12	0.12	0.14
coopf (d)	0.85	0.11	0.47	0.18	0.61	0.20
coopu (d)	0.63	0.13	0.43	0.21	0.47	0.26
d_SFS (d)	1.27	0.16	1.44	0.19	1.88	0.87
$d\_SFS \times coopf$ (d)	–0.01	0.18	0.15	0.21	0.13	0.24
$d\_SFS \times coopu$ (d)	–0.14	0.19	–0.10	0.22	0.15	0.25
$d\_SFS \times Y_{t-1}$ (d)	–	–	–0.98	0.24	–1.30	0.34
Y_{t-1} (d)	–	–	1.81	0.17	2.31	0.25
$Corr(\varepsilon^{(1)}, \varepsilon^{(2)})$	–	–	–	–	0.22	(0.04)

Note: Logit n = 2467; conditional logit n = 1484; (d) indicates binary variable; dummies for size and industry are included as regressors, but the corresponding parameter estimates are not reported here.
[a] Std. Dev. calculated conditional on the step-1 estimates.

From the highly significant coefficient estimate of the lagged dependent variable, Y_{t-1}, for all four types of innovation (see Tables 6–9) we can infer that innovation is a rather persistent characteristic of a firm. This is illustrated by the estimated conditional probabilities of innovation in Table 10 (based on the results reported in Tables 6–9). For example, the conditional probability of patenting in subperiod t, given that the firm applied for a patent in subperiod $t–1$, is 0.54 (S.E. = 0.05) if $d\_SFS = 0$ and 0.74 (S.E. = 0.03) if $d\_SFS = 1$. The corresponding numbers when $Y_{t-1} = 0$, i.e., the firm did not apply for patents in subperiod $t-1$, are much smaller: 0.04 (S.E. = 0.01) and 0.10 (S.E. = 0.01), respectively. For other types of innovation we observe the same pattern. Apparently, Table 10 shows that SkatteFUNN has a huge positive effect on the probability of innovating. However, the pairwise differences obtained by comparing $d\_SFS = 0$ with $d\_SFS = 1$ *cannot* be interpreted as effects of SkatteFUNN, because they reflect a gross effect — firms with $d\_SFS = 1$ have, on average, different values for the other explanatory variables in X_t, which have been "marginalized out" to obtain the estimated conditional probabilities reported in Table 10.[14]

[14] If Z denotes the vector of variables that are "marginalized out", then, for example:
$$\Pr(Y_t = 1 \mid Y_{t-1} = 1, d\_SFS = 0) = \sum_z \Pr(Y_t = 1 \mid Y_{t-1} = 1, d\_SFS = 0, Z = z) \Pr(Z = z \mid Y_{t-1} = 1, d\_SFS = 0) .$$

Let us now look at the (more interesting) estimates for the *partial effects* of SkatteFUNN on the different types of product and process innovations. We will restrict this part of the analysis to the model with selection specified in Equations (5)–(8), where *d_SFS* is specified as an endogenous variable, and to two types of innovations: a new product for the firm and a new production process. For only these types of innovations, we clearly reject the joint hypothesis that all the variables involving *d_SFS* have zero coefficients.

Table 10.
Estimated probabilities of innovation conditional on previous innovation activity and participation in SkatteFUNN.[a]

	Type of innovation			
Probability	Patent (d)	New product for the firm (d)	New product for the market (d)	New production process (d)
$\Pr(Y_t = 1 \mid Y_{t-1} = 1, \text{d\_SFS} = 0)$	0.54	0.65	0.61	0.56
$\Pr(Y_t = 1 \mid Y_{t-1} = 1, \text{d\_SFS} = 1)$	0.74	0.90	0.79	0.74
$\Pr(Y_t = 1 \mid Y_{t-1} = 0, \text{d\_SFS} = 0)$	0.03	0.16	0.10	0.14
$\Pr(Y_t = 1 \mid Y_{t-1} = 0, \text{d\_SFS} = 1)$	0.10	0.66	0.48	0.49

Note: (d) indicates binary variable
[a] Based on conditional logit with selection.

Estimates of the partial effects of SkatteFUNN subsidies on the probability of a new product for the firm are presented in part (1) of Table 11. We see that a significant positive effect is found only for firms that did not innovate in the previous subperiod *and* cooperated with another firm. For example, for a representative firm with such cooperation and no innovation in the previous subperiod (and average values of all other variables), the effect of a change in the value of *d_SFS* from 0 to 1is given by a logit coefficient of 1.92 (S.E. = 0.90).[15] This change increases the probability of a new product for the firm by 0.26 from 0.16 (see Table 7 and Table 11). In the case of cooperation with both, i.e., another firm and a research institute, this probability increases by 0.27. On the other hand, if the firm, *ceteris paribus*, had such an innovation in the previous subperiod, the effect of SkatteFUNN becomes insignificant.

The partial effects of SkatteFUNN subsidies on the probability of a new production process are presented in part (2) of Table 11. We see that these effects are highly dependent on the lagged dependent variable, Y_{t-1}. For a firm with no cooperation and no process innovations in the previous subperiod, the estimated effect is given by a logit coefficient of 1.88 (S.E. =

[15] Formally, this is the change in the log odds $\ln(p/(1-p))$, where p is the probability of innovation.

0.87), which is significant at the 5 percent level. For a representative firm, this means an increase in the probability of a process innovation equal to 0.23 (from 0.14) as a result of the SkatteFUNN subsidy (see Table 9 and Table 11). If the firm also cooperated with another firm or a research institute the probability of a process innovation increases by 0.24 (and by 0.26 in the case of cooperation with both another firm and a research institute). On the other hand, if the firm, *ceteris paribus*, had innovations in the previous subperiod, none of the partial effects are significant.

Table 11.
Partial effects of SkatteFUNN on the probability of innovating (new product for the firm and process innovation).[a]

Conditional on	(1) New product for the firm (d)				(2) New production process (d)			
	Share of obs.	Estimated logit	Std. Dev.	Changes in probability	Share of obs.	Estimated logit	Std. Dev.	Changes in probability
$Y_{t-1}=0, coopf=0, coopu=0$	0.48	1.48	0.86	0.20	0.52	1.88	0.87	0.23
$Y_{t-1}=0, coopf=1, coopu=0$	0.02	1.92	0.90	0.26	0.03	2.01	0.90	0.24
$Y_{t-1}=0, coopf=0, coopu=1$	0.22	1.55	0.91	0.21	0.18	2.03	0.91	0.24
$Y_{t-1}=0, coopf=1, coopu=1$	0.04	1.99	0.95	0.27	0.04	2.16	0.94	0.26
$Y_{t-1}=1, coopf=0, coopu=0$	0.03	0.28	0.91	0.06	0.06	0.58	0.93	0.14
$Y_{t-1}=1, coopf=1, coopu=0$	0.01	0.72	0.95	0.16	0.02	0.71	0.96	0.17
$Y_{t-1}=1, coopf=0, coopu=1$	0.11	0.35	0.96	0.08	0.08	0.73	0.97	0.18
$Y_{t-1}=1, coopf=1, coopu=1$	0.07	0.79	1.00	0.17	0.06	0.86	1.00	0.21

Note: (d) indicates binary variable. The change in probability is calculated for a representative firm in the corresponding group.

[a] Based on conditional logit with selection.

6. Conclusions

In this paper, we have studied how the Norwegian R&D tax credit scheme, SkatteFUNN, has affected firms' innovation activities. Our results imply that the SkatteFUNN scheme contributes to the development of new production processes and to a lesser extent to new products for the firm. Firms that collaborate with other firms are more likely to have successful innovations. However, the scheme does not appear to contribute significantly to innovations in the form of new products for the market or to patenting.

Our findings are based on an elaborate empirical model which controls for self-selection and endogeneity in a two-equation setup: one SkatteFUNN participation-equation and one innovation outcome-equation. The model also allows for firm specific heterogeneity in the effects of participating in the tax credit scheme. One important advantage of our data is the possibility to test the exclusion restrictions used in order to identify causal effects of

SkatteFUNN on innovations. For that purpose we used the innovation survey from the 3-year period preceding SkatteFUNN.

Our analyses demonstrated that controlling for self-selection has a huge impact on the conclusions. If we had based our inference on models that do not take self-selection into account, thus implicitly assuming selection based on observables, the conclusions would have been much more positive with regard to the effects of the tax credit scheme.

The finding that the SkatteFUNN scheme stimulates innovations in the form of new products for the firm and new production processes, but not major product innovations, suggests that the scheme does not stimulate innovations that may create significant spillovers among the firms. If we take into account that the majority of firms that are subsidized through the scheme are SMEs, our results may not come as a surprise. The scheme targets firms with little or no previous record of R&D activities as the *marginal* cost of R&D is lowered by the tax credit only for firms whose R&D spending would be below the cap in the absence of the scheme. On the other hand, for firms that usually undertake R&D of some size, the scheme mainly offers a (fairly small) subsidy to *average* R&D. Since the probability of a process innovation *relative* to that of a product innovation (i.e., a new product to the market) is highest in the first group, it is not surprising that the scheme favours process innovations over product innovations.

For stimulation of major product innovations, other research policy instruments such as grants from the Research Council of Norway are probably more relevant. It has been argued, cf. Narula (2002), that the Norwegian innovation system has been built around large firms engaged in traditional, more mature resource intensive industries. Smaller firms operating outside these industries have less access to outside funding and the tax credit scheme could be interpreted as an attempt to reduce this lock-in. More generally, the idea that governments should adapt a mix of policies to foster innovation has been suggested by Mohnen and Röller (2005).

Acknowledgments

We are grateful to Jon Hekland, Jarle Møen, Terje Skjerpen and two anonymous referees for many helpful comments. We acknowledge financial support from the Norwegian Research Council (the "BITE" project).

References

Almus, M., Czarnitzki, D., 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.

Blanes, J.V., Busom, I., 2004. Who participates in R&D subsidy programs? The case of Spanish manufacturing firms. Research Policy 33, 1459–1476.

Busom, I., 2000. An empirical evaluation of the effects of R&D subsidies. Economics of Innovation and New Technology 9, 111–148.

Cappelen, Å., Fjærli, E., Foyn, F., Hægeland, T., Møen, J., Raknerud, A., Rybalka, M., 2010. Evaluation of the Norwegian R&D tax credit scheme. Journal of Technology Management and Innovation 5, 97-109.

Crepon, B., Duguet, E., Mairesse, J., 1998. Research, innovation and productivity: An econometric analysis at the firm level. Economics of Innovation and New Technology 7, 115–158.

Czarnitzki, D., Licht, G., 2006. Additionality of public R&D grants in a transition economy: The case of Eastern Germany. Economics of Transition 14, 101-131.

Dagsvik, J.K., Hægeland, T., Raknerud, A., 2011. Estimating the returns to schooling: A likelihood approach based on normal mixtures. Forthcoming in Journal of Applied Econometrics (DOI: 10.1002/jae.1172).

David, P.A., Hall, B.H., Toole, A.A., 2000. Is public R&D a complement or substitute for private R&D? A review of the econometric evidence. Research Policy 29, 497–530.

De Jong, J.P.J., Verhoeven, W.H.J., 2007. WBSO Evaluation 2001–2005: Impacts, target group, reach and implementation. Research Series 07 OI 35. Den Haag: Ministry of Economic Affairs.

Geroski, P., van Reenen, J., Walters, C., 1997. How persistently do firms innovate? Research Policy 26, 33-48.

Gourieroux, C., Montfort A., Trognon, A., 1984. Pseudo maximum likelihood methods: Theory. Econometrica 52, 681–700.

Griffith, R., Huergo, E., Mairesse, J., Peters, B., 2006. Innovation and productivity across four European countries. Oxford Review of Economic Policy 22, 483–498.

Griliches, Z., 1990. Patent statistics as economic indicators: A survey. Journal of Economic Literature 28, 1661–1707.

Griliches, Z., 1992. The search of R&D spillovers. Scandinavian Journal of Economics 94, S29–47.

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

Hall, B.H., Griliches, Z., Hausmann, J. A., 1986. Patents and R&D: Is there a lag? International Economic Review 27, 265-283.

Hall, B.H., Mairesse, J., 1995. Exploring the relationship between R&D and productivity in French manufacturing firms. Journal of Econometrics 65, 263–293.

Heckman, J.J., 1976. The common structure of statistical models of truncation: Sample selection and limited dependent variables and a simple estimator for such models. Annals of Economic and Social Measurement 5, 475–492.

Heckman, J.J., 1979. Sample selection bias as a specification error. Econometrica 47, 153–161.

Hussinger, K, 2008. R&D and subsidies at the firm level: An application of parametric and semiparametric two-step selection models. Journal of Applied Econometrics 23, 729-747.

Hægeland, T., Møen, J., 2007. Input additionality in the Norwegian R&D tax credit scheme. Reports 2007/47, Oslo: Statistics Norway.

Klette, T.J., Møen, J., Griliches, Z., 2000. Do subsidies to commercial R&D reduce market failures? Microeconometric evaluation studies. Research Policy 29, 471–497.

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

Lokshin, B., Mohnen, P., 2007. Measuring the effectiveness of R&D tax credits in the Netherlands. UNU-MERIT Working paper No. 2007-025.

Malerba, F., Orsenigo, L., 1999. Technology entry, exit and survival: An empirical analysis of patent data. Research Policy 28, 643-660.

Mairesse, J., Mohnen, P., 2010. Using innovation surveys for econometric analysis. In: Hall, B.H., Rosenberg, N. (Ed.), Handbook of the Economics of Innovation, Elsevier, Amsterdam, pp. 1130-1155.

Mohnen, P., Röller, L-H., 2005. Complementarities in innovation policy. European Economic Review 49, 1431-1450.

Narula, R., 2002. Innovation systems and 'inertia' in R&D location: Norwegian firms and the role of systemic lock-in. Research Policy 32, 1-22.

RCN, 2010. Det norske forsknings- og innovasjonssystemet – statistikk og indikatorer (Norwegian research and innovation systems – statistics and indicators), NOU.

OECD, 2007. Economic Surveys: Norway. Volume 2007/2. OECD, Paris.

Parisi, M.L., Schiantarelli, F., Sembenelli, A., 2006. Productivity, innovation and R&D: Micro evidence for Italy. European Economic Review 50, 2037–2061.

Peters, B., 2009. Persistence of innovation: Stylised facts and panel data evidence. Journal of Technology Transfer 34, 223-43.

Romer, P., 1990. Endogenous technological change. Journal of Political Economy 98, 71–101.

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

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.